GUIDELINES FOR SPEECH-ACCESSIBLE HTML FOR DRAGON® NATURALLYSPEAKING® AND DRAGON MEDICAL

NUANCE

CONTENTS

OVERVIEW

HTML provides great flexibility in designing documents and can be used to create an almost endless variety of document styles and formats. This document identifies specific techniques and design approaches that help make HTML documents and applications conducive to use with the Dragon® NaturallySpeaking®, Dragon Medical 360 | Network Edition, and Dragon Medical Practice Edition (referred to collectively in this document as Dragon) software and Microsoft® Internet Explorer (version 5 and higher) or Mozilla Firefox® (version 2). It also points out specific techniques that can hinder this goal.

In general, speech support can be either explicit or implicit. In the explicit case, a web developer incorporates speech support directly into the document using the ActiveX controls in the Dragon API. In the implicit case, the end user takes advantage of the support for Internet Explorer or Mozilla Firefox in the Dragon program, and can view and navigate HTML documents that are not explicitly speech-enabled.

This document addresses the latter case, by listing guidelines for authoring HTML documents that work well with the Dragon program. The intent is to help web designers create HTML documents and applications that users can navigate and enter text into intuitively and conveniently, though perhaps not exclusively, by speech.

In many cases, the guidelines are similar to those for making HTML documents usable by text-only browsers or screen readers for the visually impaired, except that the suggestions apply only to the mechanisms for user interaction, and not to the presentation of the content.[1] The guidelines apply only to those elements with which users can interact; static elements are unaffected. While the user can use Dragon even when an HTML document does not adhere to these guidelines, he or she typically needs to perform one or more additional steps to achieve a particular task.

This document assumes you are already familiar with the commands that Dragon provides for use with Internet Explorer and Mozilla Firefox. These commands are documented in the section on Internet Explorer and Mozilla Firefox in the Dragon Help.

---

1. To make HTML documents accessible to disabled users with a variety of assistive technologies, follow the World Wide Web Consortium accessibility guidelines at http://www.w3.org/TR/WAI-WEBCONTENT/ or see their main page at http://www.w3.org/WAI/.

GENERAL REQUIREMENTS

There are two fundamental requirements. The user should be able to:

- Dictate into any input area, taking advantage of the dictation support in Dragon. This requirement is paramount when Dragon is being used as a productivity tool, for example to replace human transcription of documents dictated by medical or legal professionals.

- Use an intuitive, unambiguous, spoken phrase as a command to navigate to any element. This requirement is paramount when Dragon is used as an assistive technology by users with little or no use of their hands, but it is also important to productivity users who are not comfortable with the mouse, or whose hands are busy with other tasks.

DICTATION

Dragon allows dictation into any Windows application, into any control into which the user can type. Controls into which the user can dictate are classed as either "standard windows" or "nonstandard windows".

- In standard windows, Dragon sets and gets the text and the selection programmatically. Since Dragon has access to the text, it can allow the user to say commands that depend on the text contents, such as "select <words>" and "correct <words>". This capability is called Full Text Control (formerly known as Select-and-Say).

- In nonstandard windows, Dragon communicates with the control by sending it keystrokes. This capability is called Basic Text Control. Voice commands can be used for editing the text, but in a limited fashion. If the user intersperses keyboard or mouse input with dictation, or says a voice command that causes Dragon to lose track of the insertion point, commands such as "select <words>" no longer function. Dragon also loses the ability to provide the correct spacing and capitalization based on the existing text in the control.

HTML text entry fields (<INPUT type="text | password">) and text areas (<TEXTAREA>) are treated as standard windows. Nuance recommends that you use these HTML elements for all text entry in an HTML-based application.

If your application uses a custom or third-party control for text entry, that control may be treated as a nonstandard window. To understand the limitations of nonstandard windows, see the Dragon Help topic "Dictating in nonstandard windows". For a list of controls that are supported (that is, treated as standard windows), see the white paper "Guidelines for Developing Windows Applications Compatible with Dragon NaturallySpeaking and Dragon Medical".

One way for users to dictate into nonstandard windows is to use the "show dictation box" command documented in that Help topic; it displays a dialog box into which the user can dictate, and pastes the dictated text into the application when the user is done.

Many users find the limitations of nonstandard windows and the "show dictation box" command to be unsatisfactory. If your users require standard windows, you have several options:

- Replace the control with an HTML element (<INPUT> or <TEXTAREA>) or a supported control. This is the simplest and most highly recommended option.

- Incorporate the Dragon custom-dictation control into your application and use it to implement dictation support for the custom text control. See the Dragon SDK documentation for information on the custom-dictation control. (If you are using a third-party control, try asking the developer of the control to add dictation support to it using the Dragon SDK.)

- Develop a wrapper for the custom text control that responds to standard Windows messages such as EN_CHANGE, EN_SELCHANGE, EN_SETSEL, EN_GETSEL.

ELEMENTS PROBLEMATIC FOR DICTATION

Applications that Respond to Keyboard Events
If your application needs to take action as a result of changes to the text, do not assume that changes can come only from the keyboard. In other words, avoid writing an application in which important functions are triggered only by keyboard events, such as onkeypress, onkeydown, or onkeyup. Dragon dictation support changes the text by sending messages such as WM_SETTEXT and EM_REPLACESEL.

Dynamic Web Pages
Dragon enumerates the elements on a web page only in response to a BeforeNavigate event. Dragon is unable to detect when the elements on a page have changed by means of a script. In particular, it does not detect a text entry field that has been dynamically added to a page, or if it is in a DIV that has been hidden and redisplayed. The symptom is that dictation appears in the Results Box but does not get placed into the text entry field.

Therefore, if your application changes elements dynamically, it should fire the BeforeNavigate event after it has done so.

If it is not possible to fire BeforeNavigate, a workaround to allow dictation is to disable the HTML support in Dragon, on the Commands tab of the Options dialog; if you do this, Full Text Control and navigation commands are not available at all in the web browser.

NAVIGATION (VOICE COMMANDS)

The single most common action in typical HTML browsing is clicking hyperlinks to navigate among documents. HTML supports several ways of specifying links, some of which correspond more naturally to spoken commands than others. For example, text links naturally imply an analogous spoken command; this is less often true for image-based links, or other elements that are not text-based. However, you can make such links accessible by speech by associating appropriate text with each element. Typically, you do this with the HTML ALT attribute.

GENERAL RECOMMENDATIONS FOR COMMANDS

- To be accessible by speech, an element must have some text clearly associated with it. For text links, this text is intrinsic. For non-text links, supply it in the element's ALT attribute. The user can then activate the element by saying the text, or any word or consecutive words within the text.

- The association of the text with the element should be obvious to a user. If the text is not displayed as part of the element, provide some additional indication to make the text and its relationship to the element clear to the user.

- Whenever possible, the text associated with a link should be unique within the page or frame set.[2] Although Dragon can deal with ambiguities by displaying numbers next to ambiguous items and letting the user choose a number, this requires an extra step on the part of the user. If the same phrase must be used for multiple elements on the same page or frame set, keep the number of duplicates to a minimum.

- Text used in a link should be pronounceable. While acronyms, names, and other terms that are not actual words, are generally recognized by Dragon, avoid any words that users might not be expected to know.

- Avoid using text in a link that the user might want to dictate as isolated text into an input area. For example, if there is a link whose text is "physical exam", the user would not be able to dictate the phrase "physical exam" by itself into an input area, nor the word "physical" or "exam" by itself; in each case Dragon would activate the link because commands take precedence over dictation. (However, the user could dictate a phrase or sentence containing the words "physical exam" because Dragon uses pauses to distinguish dictation from commands).

- Avoid altering text to affect its appearance. An example of such an alteration would be the addition of spaces between letters for emphasis, for example,

  I M P O R T A N T – R E A D T H I S N O W

  While the meaning of this is obvious to a human reader, it is very unlikely that speech recognition software would recognize this text as the phrase "Important – Read This Now".

Note that the format of text links as specified through HTML tags (such as <B></B>, <I></I>,etc.) does not affect recognition. There are no restrictions on formatting of text using HTML formatting codes.

COMMAND CONSIDERATIONS FOR SPECIFIC HTML ELEMENT TYPES

Anchor Elements (<A>…</A>)
Text links (anchor elements) naturally provide the text that should be spoken for speech access, and therefore require no specification beyond the guidelines above. If, for some reason, an anchor element has no text, Dragon NaturallySpeaking uses its ALT text.

Image (<IMG>) and Imagemap (<MAP>) Links
Images pose more of a challenge with respect to speech-enabled browsing than do text links, because there is no inherent requirement to have text associated with the image.

Two general approaches can be taken to associate text with an image link. First, you can assign ALT or TITLE attributes containing the text that corresponds to the image (or to each individual area of an imagemap). Second, you can place an equivalent text link adjacent to the image (or multiple text links adjacent to an imagemap), providing an alternate way to access the link.

---

2. A typical case that does not follow this suggestion is a page with numerous links that all use a common, generic phrase such as "Click Here".

In some cases, specifying ALT or TITLE text is sufficient. This is particularly appropriate for images, which actually represent the text, such as the following graphical representation of a button, shown at right.

In other cases, it is more difficult to infer the corresponding text, and some additional indication is called for. In such cases, a caption and/or corresponding text link is useful. Examples of these on a "Homework Help" web page look like those shown at right.
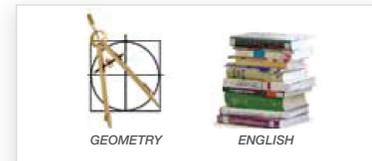
In all cases, the associated text should conform to the general recommendations previously discussed.

Buttons (<INPUT type="submit | button" | reset | image">)
Dragon uses the button's caption (the VALUE attribute) as the spoken text, as well as ALT or TITLE attributes if any. The caption should conform to the general recommendations previously discussed.

Edit Controls (Text Entry Fields) (<INPUT type="text | password">)
While Dragon provides commands such as "click text field" to navigate to text fields, it is preferable to be able to reach each field directly by speaking. These elements have NAME attributes that can be used as the spoken text. Although these elements do not typically have ALT or TITLE attributes, you can include those attributes as well. Using the ALT or TITLE attribute may be preferable in software-generated HTML, which often assigns unpronounceable, programmatic NAME values that cannot be easily changed without breaking the functionality of the document.

In either case, the associated text should conform to the general recommendations previously discussed, and you need to make the associated text apparent to the user.

Text Areas (<TEXTAREA>)
You can assign an ALT or TITLE attribute to a text area. As with edit controls, you need to find a way to make the associated text apparent to the user, and the text should conform to the general recommendations.

Menus ("List Boxes") (<SELECT>)
Assign an ALT or TITLE attribute in the <SELECT> tag to provide a spoken command for setting focus to the menu. You need to make the associated text apparent to the user, and the text should conform to the general recommendations.

Dragon uses the content of each option (<OPTION>) as the spoken text to select that option. However, it does not recognize the options unless the menu has focus.

Label Elements (<LABEL>...</LABEL>)
Label elements naturally provide the text that should be spoken for speech access, and therefore require no specification beyond the guidelines above.

Frames
Use frames with discretion, as they make navigation more complicated. Scrolling commands apply only to the active frame that has the focus. In addition, since Dragon processes all frames that are displayed, speakable text should be unique across all frames to avoid ambiguity.

ELEMENTS PROBLEMATIC FOR SPEECH NAVIGATION

Active Content

Certain HTML elements do not lend themselves to speech navigation, such as active content (for example, embedded applets or ActiveX controls). Because of the open-ended nature of active content, it is not technically feasible for the developers of Dragon, or any other speech-enabled browsing software, to implement a general solution for speech-enabling active content.[3]

In addition, since speech-recognition users do not generally move the mouse pointer around the document, avoid content that can be seen only when the mouse is placed over a specific region (such as a menu that appears only when the mouse is "rolled" over a certain location). Restrict use of such effects to providing visual enhancements, rather than functionality essential to navigation of the document.

Scrollable <DIV> Elements

As previously mentioned, if the user speaks ambiguous text, Dragon displays a number next to each matching element. If a <DIV> element is used to organize elements on a page, and if it is allowed to scroll (using one of the attributes overflow-x:scroll, overflow-y:scroll or overflow:scroll), the numbers do not scroll if the <DIV> element is scrolled. This is confusing to the user because the numbers do not remain next to their corresponding HTML elements. To avoid this problem, use frames instead of scrollable <DIV> elements.

<DIV> Elements with Large Z-INDEX Values

Commands such as "click text link" display numbered flags whose Z-INDEX property is 100. If a page contains a nontransparent positioned element whose Z-INDEX is 100 or greater, it can obscure the numbered flags, therefore Nuance recommends using a value of 99 or less.

Dynamic Web Pages

See the previous discussion above under "Elements Problematic for Dictation".

ADDITIONAL INFORMATION ON HTML SUPPORT

Enabling HTML Support

The HTML support in Dragon is enabled by default; you can make sure it is enabled by looking in the Dragon Options dialog for "Enable HTML Support" on the Commands tab, (or in the Nuance Management Console of Dragon Medical 360 | Network Edition, look on the DM360 Network Edition Settings tab for a group). This setting is stored on a per-user basis and a default setting can be specified at installation time. In addition to enabling Dragon to generate navigation commands, it also affects dictation support into text areas; in other words, if this option is disabled, then Dragon treats text areas as nonstandard windows.

Enabling HTML Commands

The Dragon option "Enable Commands in HTML Windows," also on the Commands tab of the Options dialog, (or on the DM360 Network Edition Settings tab for a group in the Nuance Management Console), enables commands specific to a web browser. If "Enable HTML Support" is checked and "Enable Commands in HTML Windows" is unchecked, Full Text Control is available in the fields that support it.

---

3. As an alternative, however, active content can be explicitly speech-enabled by its author through the Dragon NaturallySpeaking SDK.

Requiring 'Click' Before Commands

Users sometimes complain that hyperlink commands are recognized accidentally. This is more likely to happen if there are many commands, or commands containing short words, or words that users often need to dictate in isolation. If this happens, you can require the user to say the word "click" before a hyperlink command. This setting can be made in Dragon 10 and higher by enabling the option "Require 'Click' to select hyperlinks in HTML windows" on the Commands tab of the Options dialog, (or on the DM360 Network Edition Settings tab for a group in the Nuance Management Console). With this option enabled, users can still navigate to links, but must say the word "Click" before the link name. This is a user-specific option and can be set as a default option at installation time.

In Dragon 9, this setting is made by editing options.ini, which is a user-specific file. Add the following line to the [Options] section of options.ini:

Click Command Required in IE=1

Supporting Documents with Many HTML Elements

By default, Dragon processes only the first 200 HTML elements on a page. Links beyond the first 200 elements cannot be reached by voice, and text areas beyond the first 200 elements cannot be dictated into. This restriction exists to limit the amount of time that Dragon uses to process HTML elements, which it does every time the user begins speaking. You can increase this limit by editing the file options.ini, a user-specific file that resides in the "current" folder of the user's Dragon profile.

You can make this the default option by editing nsdefaults.ini, which resides by default in C:\Documents and Settings\All Users\Application Data\Nuance\NaturallySpeaking nn (Windows XP), or C:\Users\AppData\Nuance\NaturallySpeakingnn (Windows Vista or 7).

At the bottom of options.ini, add this line:

Number Of Enumerated Html Elements=n

where "n" is the maximum number of elements you are likely to present on one page.

For example: Number Of Enumerated Html Elements=500

If this setting causes a perceptible slowdown in Dragon, your only option is to redesign the HTML document to reduce the number of HTML elements. ∎

NUANCE