

Voice Dictation for Word Processing in Windows 7

with mention of other speech recognition programs

presented to the
Lexington Computer & Technology Group

by Gary Patrick
January 21, 2015

(with amendments since the presentation for completeness)

Table of Contents:

- Slide 3: What does Windows Speech Recognition Do?
- Slides 4-8 Examples of Dictation into Windows SR
- Slides 9-15 How to set up Windows Speech Recognition
- Slides 16-19 The Speech Reference Card (active help)
- Slides 20-21 Other Speech Recognition Programs
- Slides 22-25 More about Dragon NaturallySpeaking
- Slide 26 References (beyond footnotes already given)
- Slides 27-28 Appendix, Make Way for Ducklings

(Slide numbers are given in the bottom right corner)

Windows' Speech Recognition works in Windows 7 and Windows 8, in English or in several foreign languages:

- Dictating and Editing Text
 - In Wordpad or Notepad
 - In Microsoft Word
- Filling in cells or forms
 - In MS Office applications such as Excel or Access
 - In Internet Explorer Browser Web Pages
- As a substitute for mouse and keyboard actions
 - both within programs and for Windows commands in Windows Explorer and on the Desktop
- You can use Speech recognition interchangeably with your mouse and keyboard.

Here's an example of a magazine article I dictated at home:

"The Story of Dragon Systems," in the Tufts University Magazine, Fall 2012:

(an excerpt from an interview of Janet and Jim Baker by Michael Blanding.)

[This was the second speech-recognition dictation I did - after 1) completing the Windows SR Tutorial, plus 2) an additional session of training, and then 3) dictating the beginning of Robert McCloskey's Make Way for Ducklings, (included in the Appendix to these slides).]

[I have underlined the mistakes below, and indicated the correct word as [>word] where the correct substitution isn't obvious. Bolded words indicate places I had to spell in or keyboard-in to fix really garbled dictation.] All in all, the accuracy is pretty good! Now begins the dictation:

"In 1971 Janet Mciver [>MacIver], **J69** {Typed in comment: Jackson College didn't merge into Tufts College of Arts and Sciences until 1980}, a biophysicist, married Jim Baker, a mathematician. They were young and ambitious and wanted to have an impact on the world. So they decided to tackle the biggest technological challenge they can [>could] hope to solve in their lifetime: natural language speech recognition. They pioneered a radical new approach and started Dragon Systems in 1982 to commercialize it. The company took off. People everywhere started using Dragon products for dictation or for communicating with devices. But before the Bakers could reap the full profit of their technology (which today is said to power the eye phones it's eerie [iPhone's Siri]), they entered into a business deal so disastrous it left them speechless.

Janet Mciver [>MacIver] Baker simply calls it "_the debacle." It's as good a word as any to describe the cascade of events that led to the sudden demise of Dragon Systems, the company she and her husband Jim had grown for decades, then lost virtually overnight. Living up to its namesake's reputation for power and creativity, Dragon cracked one of the most intractable problems in computing - speech recognition technology. Then, just as their successful business had merged with a larger company and was poised to grow worldwide, the Bakers watched as the company went bust, taking all their technology, employees, and money with it.

Worse yet, they lost control of the patents to their own ideas, and we're now suddenly barred from developing the technology they had [>they'd] pioneered. Baker has a word for that, too: "_devastating," she says. Sitting at her dining table in an ornate Victorian home in the Boston suburb of Newton, Baker clearly didn't lose everything with the dissolution of her company. But even after more than a decade, and continuing lawsuits against those redeem [>they deem] responsible, she can [>can't] talk about the debacle without emotion.

She plucks artifacts of Dragon Systems from a large plastic BN filled with magazine clippings, articles from the New York Times and the Wall Street Journal, and other mementos up. [extraneous syllable from commanding "period."] Both of the Bakers has [>have] since landed on their feet. Jim is a distinguished career professor at Carnegie Mellon, and helped found the defense department funded Human Language technology center at Johns Hopkins University. Janet is a visiting scientist at the MIT Media lab and a lecturer at Harvard Medical School, where she is researching how the brain understands language.

And although their company disappeared, the technology that Baker helped invent lives on. In February 2010, the company that acquired Dragon's patents [**and**] software, Nuance Communications, collaborated with Siri Inc. a spinoff of SRI international (formerly Stanford Research Institute), to unveil a new **smartphone app**, an intelligent assistant called Siri. Two months later, Apple acquired a [extraneous syllable] Siri Incorporated for an undisclosed sum and made the app the most hyped aspect of its **iPhone** four S, released last fall (Siri is also a big part of the new iPhone five). Nuance itself produces the [>a] suite of dictation programs based on Dragon Naturally Speaking, licenses Dragon_brand software to control everything from car navigation systems to coffee makers to T V's, and just announced its own competitor to Siri called Nina. None of this would be possible if it weren't for a promise that Baker and her husband made to each other more than 40 years ago.

Now in her sixties, Janet Baker flashes a warm smile as she lays out muffins and coffee on the [>a] **sun-splashed** dining table. Her long gray hair is tied back in a no nonsense ponytail, and she wears dangling [>dangly] earrings, gold rimmed glasses, and a vaguely Asian looking purple blouse. Dragons are all around the house, with statuettes lining the mantel in the living room and a 200 [>two hundred] pound sculpture guarding the front hall. "in the Chinese world the dragon is the spirit of creativity, but every culture has some powerful, untamed force," she says. "we thought that speech and language was also a very powerful force."

The Bakers first vowed to tame that force soon after they married, during Columbus Day weekend of 1971. Of [extraneous syllable] the two were doctoral students at Rockefeller University, in New York City, later transferring to Carnegie Mellon University, he a mathematician and she a biophysicist. Young and ambitious, they sought an area in which they could pool their expertise to make a significant impact on the world. "we wanted to do something that would be practical and useful and more than a paper on a library shelf," Baker says. "and we decided the goal had to be satisfied in our lifetime."

Where did this dictation run into trouble?

1) some trouble mistaking the ending of a word with the beginning of the next, where the same consonant or sibilant is involved; it can be a tough distinction.

examples: "can't talk" was interpreted as "can talk."

"have since" was interpreted as "has since."

2) Saying "hyphen" was interpreted as a command several times that shrank the page width on screen; so I gave up on that, for "comma." There are hyphens in a great many words in the article text, that I have let go as spaces here.

I don't know how to insert a hyphen between words already dictated, other than putting it in by mouse-positioning and keyboard. Afterwards, I realized I should have tried "dash" or "endash."

3) It misses capitalizing the first word in a sentence beginning a quotation, and throws in an extraneous space first, where I had said "open double quote."

4) I had to capitalize words a number of times, even on later occurrences.

5) I kept getting "seri" instead of "Siri" even after correcting it three times.

During the January 21st meeting, I dictated this grocery list, unrehearsed, using the microphone inside my laptop pc:

Wheat germ

Paper towels

Egg beaters

Sauerkraut

Peanut butter

Two packages of chickpeas

The toothpaste

Spade and span powder

In the frozen peas

Two canned corn

Black beans

Raisin bread

Two canned salmon

Old Cape COD crackers

Sliced water chestnuts

Four diced tomatoes

Nonfat plain yogurt.

- Produced a very usable list.

- Windows Speech Recognition got it all right except:

← “The” is extraneous; background noise?

← Should be “Spic n Span” powder

← “In the” is extraneous; background noise?

← Capitalized “Cape” correctly, but don't know why “COD” is all caps.

← “cans” implied in my saying “4 diced tomatoes”

Some Preliminary Steps to use Windows Speech Recognition

- If you want more than one person to be able to use Windows Speech Recognition, set up a separate Windows log-on account for each person.
- Plug in a microphone
(unless your computer has a built in microphone)
Windows seems to check before allowing SR to start up.
- A higher quality microphone will give better results.⁽¹⁾

(1) Christopher Null, PCWorld Article, July 8, 2014: (Comparing microphones)
www.pcworld.com/article/2458363/dragon-naturallyspeaking-13-review-better-than-ever-at-letting-you-speak-freely.html

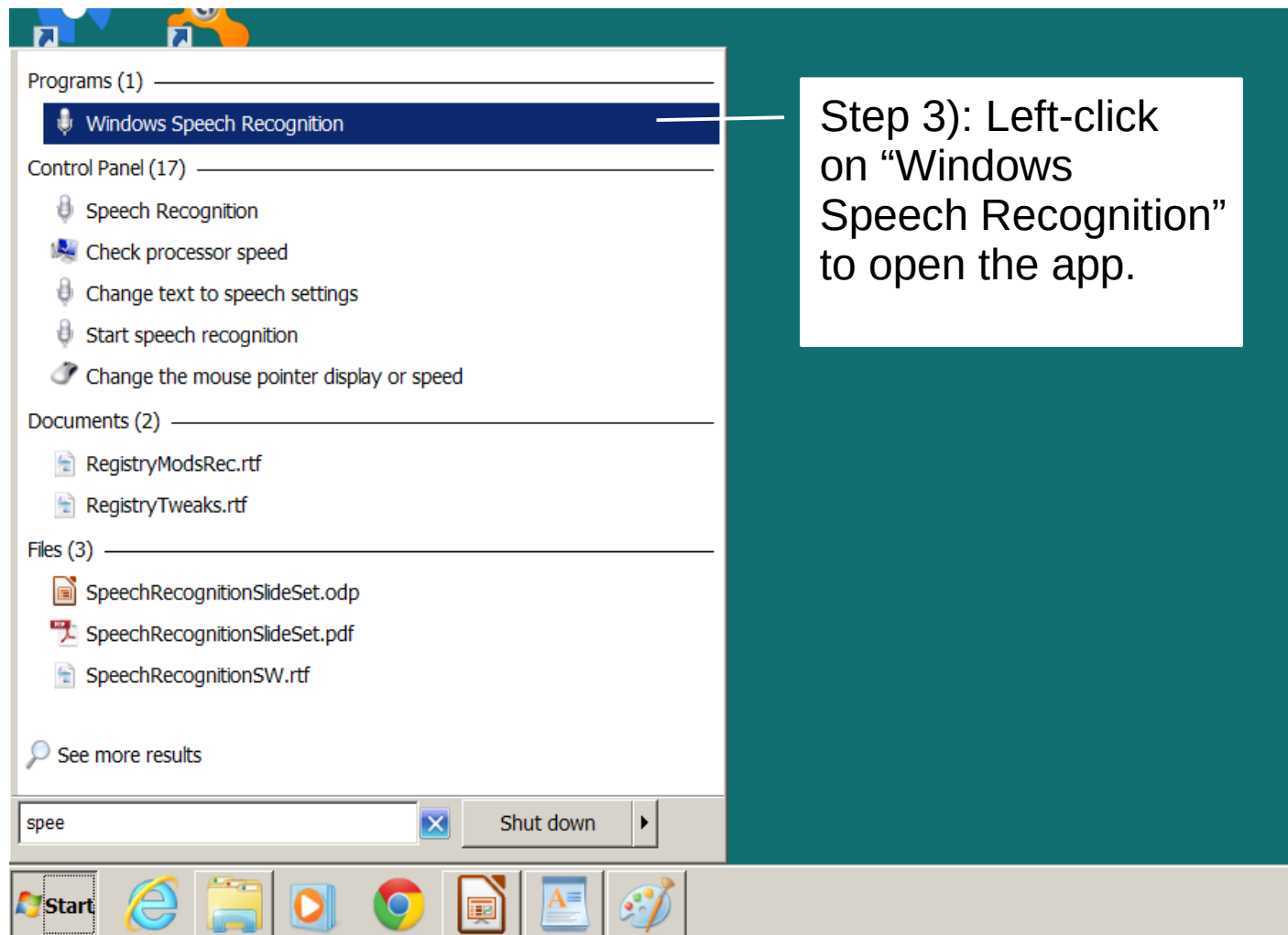
First you must set up Speech Recognition; A Windows Wizard guides you through:

- Indicating the type of microphone you are using
- Giving a sample of your voice for amplitude scaling
- Getting your permission (or not) for the speech dictionary to access your documents and email to add vocabulary
- Configuring a method of giving the “Start Listening” command
 - either by clicking the microphone icon in the SR Control Panel (that will be on your display),
 - or by saying “Start Listening.”
- Choosing whether (or not) to have Speech Recognition start up when Windows starts up.
- Finally it recommends you take the Tutorial, because SR learns your voice and pronunciation from the phrases you are asked to read (during the Tutorial).

The official way to start Speech Recognition is to click Start, All Programs, Accessories, Ease of Access, Windows Speech Recognition. But there's an easier way:

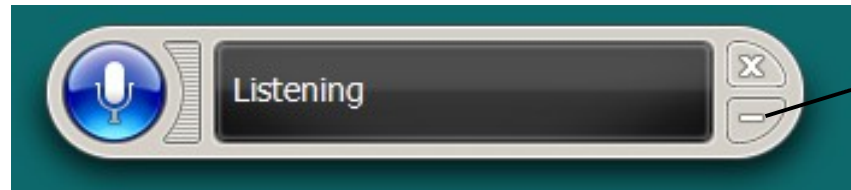
Step 1) Click Start, and

Step 2) type into the Start box “speech . .” You'll get:



The first time ever you start Windows Speech Recognition (from actions in the previous slide) you'll enter the Setup Wizard (not shown here).

Otherwise, the Speech Recognition Control Panel will appear:



Note “minimize”
button

- By default it appears first at the top center of your display.
- You can move this anywhere you like on your display, by clicking and dragging it.
- It will always reside on top of all other windows.

(unless you minimize it to the Notification Area – by button indicated above – and see illustration the slide after next)

- When you Shut Down your computer, the SR app will close automatically, just as other Windows programs do.

The Speech Recognition Control Panel provides

- A feedback status message window, and
- User controls (as called out below):

A Guide to the Features of the Speech Recognition Control Panel:

Click on this microphone symbol to turn listening mode "On" or "Off"

This Feedback Status Message Window tells listening status = "Off" or "Listening," or more.

"X" closes the SR app.

Speech amplitude indicator

Right-clicking on the frame of the Speech Recognition Control Panel opens this Controls Menu

Click on this for a Guide to spoken commands, punctuation, and more.

Recycle Bin Adobe Reader XI

Malwarebytes Anti-Malware Avast Free Antivirus

CCleaner Picasa 3

LibreOffice 4.2

On: Listen to everything I say
Sleep: Listen only for "start listening"
✓ Off: Do not listen to anything I say

Open Speech Reference Card
Start Speech Tutorial
Help

Options
Configuration
Open the Speech Dictionary
Dictation Topic

Help improve Speech Recognition...
About Windows Speech Recognition

Open Speech Recognition
Exit

How to access the Speech Recognition controls after the SR app has been “minimized” to the Taskbar Notification Area:

(read the pointers below in the order of their numbering)

(Illustration of a Windows 7 Desktop & Controls)

- 1) When Speech Recognition is running but minimized, it hides (with other minimized apps) under this up-arrow, in the Windows Taskbar Notification Area.
- 2) To un-minimize SR, left-click on the up-arrow to reveal the control icons for all the minimized apps;
- 3) The Speech Recognition icon is a microphone symbol.
- 4) Left- or right-click on the SR icon to open the Controls Menu, then click on "Open Speech Recognition" to un-minimize it.
- 5) (Optional) If you want to have Speech Recognition start up automatically as Windows starts up, then click "Options," and then "Run at Startup" in this pop-up panel.

The screenshot shows a Windows 7 desktop with a teal background. The taskbar at the bottom contains several icons, including a microphone icon for Speech Recognition. A context menu is open over the microphone icon, showing options like 'Open Speech Recognition', 'Options', 'Configuration', and 'Exit'. A second context menu is open over the 'Options' menu item, showing options like 'Play audible feedback', 'Run at startup', 'Speak text in correction dialog', and 'Enable dictation scratchpad'. The taskbar notification area is visible on the right, showing a 'Taskbar Notification Area' label and a 'Customize...' button. The system tray at the bottom right shows the time '4:48 PM' and date '2/6/2015'. Red arrows point from the numbered instructions to the corresponding UI elements: 1) to the up-arrow in the taskbar, 2) to the up-arrow, 3) to the microphone icon, 4) to the 'Open Speech Recognition' menu item, and 5) to the 'Options' menu item and its sub-menu.

Speech Recognition deserves a more convenient way to start the app than Microsoft provided.

You can create a start-up icon in any of three places:

On the Start Menu, on the Desktop, or on the Taskbar.

The screenshot shows a Windows search window with the search term 'speech'. The search results are categorized into Programs (1), Control Panel (17), Documents (2), and Files (3). The 'Windows Speech Recognition' program is selected, and its context menu is open, showing options like 'Open', 'Run as administrator', 'Scan sapisvr.exe', 'Pin to Taskbar', and 'Pin to Start Menu'. The 'Desktop (create shortcut)' option is highlighted in the 'Send to' sub-menu. Text boxes with arrows point to these options, explaining their functions: 'Run as administrator', 'Pin to Taskbar', and 'Pin to Start Menu' are used to create a Taskbar icon; 'Send to' > 'Desktop (create shortcut)' is used to create a Desktop icon.

As before, click the "Start" button, and type in "speech."

On the search result, right click to open this menu, with the 3 choices

Left-click to create a Taskbar icon, or a Start Menu icon, or

a Desktop icon.

This is the Speech Reference Card, top level, that opens individual topics:



Here's a summary of a few most useful voice commands:

- editing text that you just spoke:

“delete that,” (“scratch that” also works) deletes the most recently dictated phrase, or deletes any **selected** text.

“correct” <re-say a word you dictated that texted wrong>

“capitalize” <say the word, recently dictated> or

“no caps” <say the word>, to change capitalization, or

“change next <n> words to uppercase” or “. . . lowercase.”

- managing dictation:

“new line,” “new paragraph,” “tab,”

“select that” to select the phrase most recently input, or

“select” <say the word, that's in the text somewhere>

“clear selection,” to de-select text.

More cursor management voice commands:

- Working relative to present cursor position:
 - “Go to <start, end> of <sentence, paragraph>”
 - “Go to <say word> goes to beginning of named word;
 - “Go after <say word> puts cursor after named word.
 - “Select <next, previous> <n> words.
 - “Delete <next, previous> <sentence, paragraph>
- Presenting numbers during dictation (or fixing later):
 - “Numeral <say number>” puts digits into text.
- mouse actions:
 - “Click,” “Double-click,” (then naming what)
 - Scroll up, scroll down, scroll left, scroll right

Punctuation Marks that are recognized while dictating:

- “comma,” ”semicolon,” ”colon,” ”period,” ”dot,”
“decimal point,” “exclamation point,” ”asterisk,”
“question mark,” “apostrophe,” “caret.”
- “Literal <any of the above>” means insert the literal name of the punctuation, instead. (Said to work for the name of any SR command).
- “<Open, Close> <double quote, quote>”
- “hyphen,” “minus sign,” “dash,” “en dash,” “em dash,,”
en dash is longer than a dash, “em dash” is longer still;
“underscore.”
- “Plus sign,” “equal sign,” etc.

Other Software that Provides Speech Recognition for Windows Computers:

- There are three types of functionality:
 - a) Can read a text file to you as audio output (some can make a DVD);
 - b) Can take your speech as commands to the computer;
 - c) Can take your speech as dictation, to write text.
- Various programs do one or more of the above:⁽²⁾

Windows 7 & 8 Speech Recognition can do all three;

Dragon NaturallySpeaking Home or Professional Edition can do all three.

Tazti, VoiceAttack, and MIA PC Friend can do only (b); all are shareware with a free trial, then a purchase fee, \$8 - \$40.
- TTS Reader, TextAloud, and DocumentSpeaker can do only (a). TTS Reader is free; the other two are \$15 - \$30 with a free trial period.

Speech Recognition Offered in the Cloud:

- TalkTyper is an online speech to text converter:
 - uses the Google voice algorithm; works inside the Chrome browser.
 - Its output options are:
 - copy text to clipboard
 - email the text
 - print the text
 - translate it to another language, using Google Translate.
 - A PC World editor⁽²⁾ thought its accuracy was not as good as Windows or Dragon NaturallySpeaking, and recommends its use for shorter spoken content only.
- Google offers speech recognition for its search function – will not be discussed here – see Google online documentation.

Dragon NaturallySpeaking is updated by Nuance Communications about once a year, now at Version 13.

Features common to both Home and Professional Editions:

- Nuance-approved headset microphone included, but can switch among different microphones on demand.
- Microsoft Office 2010 and 2013, and Wordperfect are supported for dictation, editing, and commands by voice.
- Internet Explorer, Chrome, and Firefox browsers supported for doing searches and dictating into text fields.
- Most web-based email, and social networks supported.
 - e.g. Gmail, Outlook.com, and Yahoo mail, and
 - Twitter and Facebook
- Text to Speech output for proofreading.
- Interactive Tutorial has been expanded over V.12, so it now includes orientation to the “Dragon Bar” and a new Dragon Vocabulary Editor.

Dragon NaturallySpeaking Premium Edition adds a few features vs. the Home Edition:

- Give commands and dictate entries in Excel spreadsheets and Powerpoint presentations.
- Import or export custom word lists, to synchronize with another user of DNS.
- Create custom voice commands to insert frequently-used text and graphics, such as logos, boilerplate text, or a signature.
- Digital Voice Recorder included, for later transcription.
Mobile version providing a DVR for smartphone to p.c. is \$299.99.
- Audio playback with associated text highlighted on screen (supplementing the text-to-speech capability in the Home Edition)
- Premium edition list price is \$199.99, currently on sale for \$149.99. (Also it pays to check Amazon.com or NewEgg.com)
- Premium Wireless version includes a Bluetooth wireless headset, for \$299.95.

User Feedback about Dragon Naturally Speaking:

- V.13 is thought to be even better than V.12. or V.12.5:
 - improved transcription accuracy, but incremental, not revolutionary.
 - better performance in web page forms; vs. previous tendency for cursor to jump unexpectedly
 - Improved learning
 - Microphones built into laptops now officially supported.
- What types of users are most appreciative:
 - Professional writers, authors, lawyers, students
 - Those with dyslexia.
- But there have been some problems installing V.13 if upgrading a p.c. from V.12. or V.12.5. both from CD-ROM and from downloaded installation. [amazon.com reviews] Stay with V.12 or 12.5 if you have it; almost as good as v.13.
- Complaints about unhelpful customer support, and refusal to allow return for refund if it won't install. [amazon.com reviews]
- One complaint V.13 runs very slowly on Windows 8.1, where V.12.5 did not have this problem. [amazon.com review]

More User Feedback about Dragon NaturallySpeaking:

- It installs two auxiliary programs that have caused compatibility problems with Chrome and Internet Explorer for some users.⁽²⁾⁽³⁾
 - Rich Internet Application Support is supposed to make Dragon work more fully in a browser, but some users experience browser freezes or crashes.
 - Common Software Manager checks for updates to Dragon, but intrudes, feeling like malware
- Both of these programs can be uninstalled like any other Windows program, and won't likely be missed.
 - Check for program updates manually, of course.

(2) Mark O'Neill, PCWorld, November 6, 2013, Comparing Five Speech Recognition Programs: editor's review of Windows Speech Recognition, Google Voice Search, Dragon NaturallySpeaking Premium Edition, TalkTyper, and Tazti.

www.pcworld.com/article/2055599/control-your-pc-with-these-5-speech-recognition-programs.html

(3) www.knowbrainer.com/forums/forum/messageview.cfm?catid=4&threadid=16276

References in addition to those given in footnotes

Top Ten Reviews, online, with a bar-graph comparison of various versions of Dragon Naturally Speaking, Tazti, and others

<http://voice-recognition-software-review.toptenreviews.com>

- A Google Web link that explains Google Voice use on Android, iPhone, and in the Chrome browser (on any platform):

<https://support.google.com/websearch/answer/2940021?hl=en>

- Windows 7 Speech Recognition review from 2010:

www.arstechnica.com/information-technology/2010/05/win-7s-built-in-speech-recognition-a-review

- Reviews for Nuance Dragon Naturally Speaking, versions 12 and 13, can be found on Amazon.com and Newegg.com

- Version & Levels (editions) history:

[en/wikipedia.org/wiki/Dragon_NaturallySpeaking](http://en.wikipedia.org/wiki/Dragon_NaturallySpeaking). The version for Apple Mac OS is called Dragon Dictate now at Version 4, (or Dragon for Mac in some older versions).

Appendix: Dictation of a little of Make Way for Ducklings

This was the first document I dictated after completing the tutorial and one session of the extra training for Windows 7 Speech Recognition. I have underlined the mistakes (and provided the correction as [> xxxx] if it's not clear what the word should be). I did have to fix capitalization, and have not shown everywhere that I did that. There are still more missed capitalizations.

I am surprised at how accurate this dictation is, even recognizing "McCloskey," (although not capitalizing correctly). It's possible, of course, that Microsoft Software Engineers used this text in developing the program.

"Make Way For Ducklings, by Robert Mccloskey. <--(underlining done manually)

Mr. and Mrs. Mallard or [> were] looking for a place to live. But every time mr. Mallard saw what looked like a [nice] place, Mrs. Mallard said it was no good. There were sure to be fox's in the woods or turtles in the water, and she was not going to raise a family where there might be fox's for turtles. So they flew on and on.

When they got to Boston, they felt too tired to fly any further. There was a nice fond in the public garden, with a little island on it. " the very place to spend the night," [elect > quacked] Mr. Mallard. So down they flapped. Next morning they fished for their breakfast [him one > in the mud] at the bottom of the pond. But they didn't find much.

Just as [the A or > they were] getting ready to start on their way, [the > a] strange enormous word came by. It was pushing a boat full of people, and there was a man sitting on its back. " good morning," quacked Mr. Mallard, being polite. The big bird was too proud to answer. But the people on the [vault through > boat threw] peanuts into the water, so the mallards followed them all rounds the pond and got another breakfast, better than the first period " I like this place," said Mrs. Mallard as they climbed out on the bank and [wobbled > waddled] along. " why don't we build a nest and raise our ducklings right in this pond? There are no fox's and no turtles, and the people [VS > feed us] peanuts. What could be better?"

" good," said Mr. Mallard, delighted that of last Mrs. Mallards had found a place [best > that] suited her. But - " lookout!" Squawks Mrs. Mallard [all of the other > all of a dither]. " you'll get run over!" And when she got her breath she added: " this is no place for babies, with all those [hard > horrid] things rushing about. We'll have to look somewhere else." "