Getting Started with Windows™ Speech Recognition (WSR)

A.   OVERVIEW

After reading Part One, the first time user will dictate an E-mail or document quickly with high accuracy. The instructions allow you to create, dictate, and send an E-mail without touching the keyboard. The second part discusses steps to attain highest accuracy. The final part has suggestions for increasing productivity when using Windows™ Speech Recognition.

II.   PART ONE

A.   WHY USE SPEECH RECOGNITION?

Most people will be able to dictate faster and more accurately than they type. My experience with Windows™ Speech Recognition is the ability to dictate over 80 words a minute with accuracy of about 99%. If you truly can type at 80 words a minute with accuracy approaching 99%, you do not need speech recognition. However, even a good keyboarder will benefit from reduced strain on the hands and arms by using Windows™ Speech Recognition.

It takes time to become comfortable with dictation into a computer. There will be moments of frustration as you go through the learning curve. If you are impatient or are a perfectionist, DO NOT read on and do not use Windows™ Speech Recognition. If you have reasonable patience, you will learn to dictate accurately and comfortably.

The best strategy is to keep things simple for first several days of using Windows™ Speech Recognition. When you are comfortable with the basics, move to part two of this document.

B.   THE MICROPHONE

A good microphone makes dictation into a computer a pleasure instead of a battle. Good microphones reproduce your voice accurately and block out background noise that distorts the audio signal. You also want to make sure to microphone is comfortable or else it will negatively affect dictation. You can see many types of microphones tested to show they work well with Windows™ Speech Recognition at **SpeechRecSolutions**.

I cannot emphasize enough the need to have a high quality microphone. If you start with a microphone poorly suited for speech recognition you are likely to abandon the practice. Some basic microphones work quite well at a cost under $50.00. More expensive microphones allow dictation in noisier environments and yield a 1-2% accuracy increase. Do not start speech recognition without a good microphone.

HOW MUCH COMPUTER POWER IS NECESSARY

A Pentium 4 with 2 GB RAM is a bare minimum for use with the Vista and Windows 7-10 operating systems. For best general computer performance, use a Core 2 Duo or faster with at least 2 GB RAM. Increasing RAM to 3 GB or 4 GB will allow Windows Speech Recognition to purr.

C.   ENABLING SPEECH RECOGNITION IN WINDOWS

Click **Start**, then **Control Panel**. Open **Speech Recognition Options**. Click **Start Speech Recognition**. **Follow the instructions** carefully. When you finish this process, Windows™ Speech Recognition is ready to accept your dictation. As the instructions describe how to use Windows™ Speech Recognition you will have a good idea how to issue basic commands.

D.   HOW TO READ DURING THE TRAINING SESSION

The training session is one way Windows™ Speech Recognition software to learns your style of speaking. In addition, think of it as a way of training yourself to dictate. As you go through the training process, you are asked to read text into your computer. Try to feel relaxed and speak clearly. By this we mean you should make sure to enunciate each word clearly. You will know you are enunciating properly if you feel your lips form each word as you speak. Speech recognition works not only by listening for sounds of words, it compares each word in the context of the surrounding words. These can be bi-grams, tri-grams or quad-grams, groups of 2, 3 or 4 words. For example, if I dictate, "They're going to park their car over there," the software needs to determine which "they're-their-there" is needed. It does this by looking at adjacent words.

Windows™ Speech Recognition is learning your voice quality and style of speaking as you perform the training sessions. You want it to get used to your normal speech. Train yourself to feel relaxed as you enunciate clearly. The point is, you do not want to trade strained arms, wrists and hands for voice strain.

E.   USE WINDOWS™ SPEECH RECOGNITION SIMPLY AT FIRST

It will take time to learn all of the commands, tricks and techniques available to speed your work. Our advice is to learn the basic commands in the beginning. Do not worry if you cannot remember all the commands. Very few people can. Most of what you need to do can be accomplished with fewer than 10 commands. These you memorize as you need them. You will figure out the others as you need them. You can always say, "What can I say" to see a list of available commands for any window.

F.   DICTATE AN E-MAIL HANDS FREE

Once you have finished training the system and gone through the tutorial you are ready to start dictating E-mail. We use Microsoft Outlook as it uses the Microsoft Word as Outlook word processor. Microsoft Windows Mail works similarly.

Dictate an E-mail (say the words in quotes):

- **"Open Microsoft Outlook."** Outlook will open.

- **"New."** You will see the drop down menu under "New" open up with Mail Message highlighted.

- **"Mail Message."** A new mail message will open, with the cursor in the **"To"** box.

- Say your name. If you your name is in the contact list, it will automatically appear. Alternatively, you can dictate an e-mail address by saying, **"Martin at yahoo mail dot com"** and martin@yahoomail.com appears.

- **"Go to Subject."** The cursor jumps to the subject box.

- Fill in the Subject. "**Catching up**."

- "**New Paragraph**." The cursor appears in the body of the email.

- **"Dear Jim comma."** "**New Paragraph**. You will see, *Dear Jim,*

- At the flashing, **"I will meet you tomorrow under the clock tower at 11:00 AM period New Paragraph."**

- **"Fred"**

- **"Send."** and the E-mail is on its way.

It really is this easy. Now it is your turn. Dictate a real E-mail for practice.

G. COMMANDS AND TECHNIQUES TO BUILD A SPEECH PROFILE CUSTOMIZED TO YOUR STYLE OF SPEAKING.

ADDING WORDS OF PEOPLE WITH UNUSUAL NAMES:

1. **"Open the Speech Dictionary."**

2. If you are using the **WSRToolkit Version 3** from MyMSSpeech.com click on the Add to Dictionary tab or **"Add a new word,"** and type in the word or words.

3. **"Next."**

4. **"Record a pronunciation" or**, **"Press Spacebar"** to activate the checkbox.

5. **"Finish."** The numeral 3 appears on the word "Finish" at the bottom right hand corner. Say **"3."** The word **OK** shows. **"OK."** Alternatively you may use the mouse to click "Finish." The point being, you have the choice to work hands-free, use the mouse, or a combination of both.

6. **"Record."** Pronounce the word or words you added. Then say or click "Finish."

H. CORRECTING MISRECOGNIZED WORDS OR PHRASES

Speech Recognition software works best when you dictate phrases. The is software is not only listening for the sounds of each word, it is comparing the words in context of surrounding words. Therefore, when a word is misrecognized, it is best to correct the word in the context of at least one other word. This is how the system learns. Usually, one correction and the word or words appear correctly from then on. Here is an example:

You said:
Meet me at the clock tower.
The computer interpreted:
Beat me at the clock tower.
Say, **"Correct, Beat me."**

The correction box appears with a numbered list of words. Most often the correct word(s), in this case "Meet me," appear in the list. Just say the number next to the correct word(s). These become highlighted. Say, **"OK"** and the correct text *Meet me* replaces *Beat me* in the document. If you do not see the word(s) you dictated in the numbered list, say the correct word again and it should appear.

If for any reason the correct word or words does not show, say, **"Spell it."** A box appears. Spell out the words letter by letter. If you are correcting a phrase with 2 or more words, say, "Space," to separate the words. If a letter is misinterpreted, say the number and the cursor jumps back to that letter and you can say it again. You can also spell letters using sample words. For example if I want to spell "Meet" I could say, "Capital M as in Mike, e as in easy, e as in easy, t as in tango."

I. USING COMMANDS

Print out the basic list of commands. Say, **"Open speech reference card."** Click the topics you are interested in and print out the commands. You will see there are a lot of commands. You do not need to learn them all. You can, if you wish, use your mouse and click on things or use the keyboard.

When new to speech recognition, it is helpful to start with basic commands only. For example, if you want to make a word or words **bold**, you can double click the word or words and click on the **bold** button. With Windows™ Speech Recognition you can say select --------- ------ (the word or words you wish to format). The words are highlighted and you just say, "Bold."

J. DICTATING IN A WORD DOCUMENT

With focus on the desktop window, Say **"Open Microsoft Word."** In the same way you dictate E-mail, you dictate in Word documents. Simply say what you want to say. Because Windows™ Speech Recognition listens for words in context of surrounding words, verbally adding punctuation marks and paragraph commands improve recognition accuracy.

For practice read a paragraph from a book, article or document that you have already created. Reading a news article from a newspaper is also good practice.

USING STYLES

If you want to create a heading, just say, **"heading 1"** or **"heading 2"** or "title", etc. The format you say will be used provided those styles are available in the document. Dictate your heading or title and say "N**ew paragraph**" to start dictating your text. Experienced speech recognition dictators use speech recognition to dictate words and punctuation. Final formatting is done the old fashioned way – with the mouse and keyboard although it is perfectly acceptable to do all formatting by voice.

K.   SUGGESTIONS

By now, you should see how easy it is to accurately dictate text.

1. For the next few days use speech recognition a few minutes every day, gradually increasing the amount of time you spend on it. If you find yourself getting frustrated, return to normal typing.

2. When you dictate something which is not correctly "heard," make sure to correct it.

3. Add new or unusual words to the speech dictionary if you plan to use them regularly. Record a pronunciation for each new word. I am always amazed whenever I add some unusual word, the next time I use it, the word or phrase usually shows up correctly. When you begin using Windows™ Speech Recognition you may find it frustrating to add words. Although it slows down the process of getting the document finished, in the long run your time investment will be worth the payback of high accuracy. If you are getting frustrated adding words, type them and add them to the speech dictionary another time.

The whole point of speech recognition is to make your work easier. Take your time at the beginning of the learning curve. Take it easy and enjoy it.

L.   SOLVING PROBLEMS

Here is a link to a webpage where you can ask questions, if you have problems, check out the MSSpeech-forum.

Here are a 3 simple solutions to solve some problems:

• I will start dictating into Word or Outlook and text will not show. However, the correction box shows. Just close that document and open a new document.

• Nothing happens when I dictate or weird stuff shows up. Just close speech recognition and open it up again.

• Right click the Speech Bar, select **Configuration** and **Set up microphone**. I usually run the Set up microphone utility before any dictation session. It sets the correct volume level for your voice and takes into account background noise levels to some degree. Running this Set up the microphone utility is also useful when you find accuracy decreasing over the course of the day. This is because your voice changes as the day wears on, background noise levels change. Many

people tend to get a little sloppy and start slurring words together. Running the Set up microphone utility reminds us to enunciate clearly and speak in phrases.

Part Two

ACHIEVE HIGH ACCURACY WITH WINDOWS™ SPEECH RECOGNITION

SPEAK THE WAY WINDOWS™ SPEECH RECOGNITION NEEDS TO HEAR YOU

Make your third grade teacher proud of the way you speak. If you have a good microphone and enunciate clearly, accuracy in the 99% range is easily attainable. Here are two links where the speakers are dictating properly.

https://www.mymsspeech.com/download/how_to.mp3

https://www.mymsspeech.com/download/marty.mp3

You will note a formality, a precision in the enunciation and pronunciation of each word. Speech recognition software has no idea what you are saying. It only "hears" electronic impulses (waveforms) matched to its memory bank of tens of thousands of words and their pronunciations. The software makes a best guess as to what it heard and then compares the word to tables of probability that asks, does this word work best in context of surrounding words.

Most of us have allowed our speech to be conversational or relaxed, as opposed to what you hear in the above links. This relaxed, conversational approach is perfectly normal and acceptable for talking with other people. The computer is not a person. It needs very clear enunciation to achieve high accuracy.

At first glance the list of 7 things to do might seem onerous. It involves a significant investment of time. However, if you prepare for and learn Speech Recognition in a systematic way, you will be successful. Following a methodical approach will increase the probabilities of achieving 99% accuracy. The difference between 96 percent accuracy and 99 percent accuracy is a great deal of time saved in editing and correcting misrecognitions. As the old auto mechanic saying goes, "You can pay a little now or a lot later. Take your pick."

MICROPHONE AND SOUNDCARD OR USB SOUNDPOD

A good microphone has several important attributes. First, it passes the sound of your voice to the computer as pure audio. Next, it eliminates most background noise that can distort the sound of your voice and make it difficult for words to be understood by the software. Finally, the microphone should be ergonomic in the sense that it meets the needs of your style of working. If you are constantly getting up and walking away from your desk, you either want a headset microphone that has a quick disconnect or you should use a desktop mounted microphone so you don't need to put on a headset and take it off constantly or even worse muss up your hair. Some people prefer a microphone with an On/mute switch so they can mute the microphone from a switch on the microphone or microphone cable.

The microphone is a simple device that uses a diaphragm to capture the analog sound waves of your voice and transmit them to the computer as electrical impulses. No computer can read analog data. The

data must come in as digital information. A Soundcard or USB Sound Pod is an analog to digital converter. Many Soundcards do an admirable job today. However, many of the sound chips built into notebook and desktop computers are poorly shielded and introduce electronic noise into the audio stream. A USB microphone, or a regular microphone connected to a USB Sound Pod, bypasses electronic noise from within the computer enclosure and injects the audio directly into the computer.

A good microphone with a USB audio input guarantees you the best audio for speech recognition software. If you're not getting the highest recognition possible when using a good microphone/soundcard, you must look at your dictation style and denunciation as the culprits.

INSURE PROPER PLACEMENT OF THE MICROPHONE ELEMENT

It is important to find the proper place for the microphone element in relation to your mouth. For a headset microphone, the microphone element at the end of the headset boom should be approximately one half to 1 inch off the corner of your mouth. If the microphone is in front of your mouth, it will pick up your breath inhales and exhales as unwarranted words. For a handheld or desktop mounted microphone you should talk right down the head of the microphone, not across the top. The head of most handheld/desktop microphones should be approximately one half to 1 inch from your mouth. An exception to this is the Buddy Desktop, Buddy Gooseneck and Buddy Flamingo microphones which have a very hot audio signal which requires they be used at 3-6 inches. Be advised, the further in the desktop microphone element is from your mouth, a stronger signal must be and the more likely it is to pick up unwanted background noise. Also, you must train yourself when using a desktop/handheld microphone to not turn her head away from the microphone while dictating as you might when looking for a report or a physician looking at a film.

CONSISTENT PLACEMENT

Use a mirror the first few times to visually see the correct position. You can use one or two finger circumferences as a spacer for positioning the microphone off the corner of your mouth with a headset microphone. Once you have placed the microphone properly, get a sense of where it is relative to your mouth and to establish that position each time. A simple way to tell if the microphone is in the path of the breath flow is to place your finger over the microphone element and see if you can feel any breath.

PERFORM A FULL TRAINING

One way to access the training text as well as the tutorial is through Control Panel. Open Speech Recognition Options. Select, Train your computer to better understand you.

TESTING THE MICROPHONE AND SYSTEM

Get some simple, easy reading, approximately 300 words. A good selection is **The Rainbow Passage**. This selection was developed by linguists as it has an even distribution by usage of the 46 phonemes (individual sounds within words) in the English language.

Dictate the passage. Count the number of errors. Correct the errors using the correction window. Read this same passage again. You should have few errors. If you do have fewer errors, then the microphone

and software are working properly. This will be the case almost every time. However, if there is no improvement you may have some hardware, software or microphone problem.

The importance of this step is to avoid spending time trying to train a system that has problems. Find out right now if it's not working well and get it fixed. Then move on to the next step.

TRAINING AND MORE TRAINING

At this point you should be above 95 percent accuracy. What happens now is you simply dictate what you need to do in the normal course of your activities. Always use the correction window. This is the only way you and the computer and software will move to 99% accuracy. Depending on which program you use, or how much you dictate, or how technical your work, it can take a week or two to achieve 99% percent accuracy. In this 1 to 2 week time period, you will learn the commands for dictation and navigation.

LEARNING TO SPEAK JUST CLEARLY ENOUGH

When *training the system* to your voice, *train yourself* to enunciate each word distinctly and clearly. I know I'm dictating correctly when I feel my lips forming each word as it is spoken. The trick is to learn to enunciate clearly, speak in phrases but to also relax when doing so. You do not want to trade strained hands from keyboarding for a strained voice. Do not strain your voice, or over enunciate. You are not only training the software to listen to you, but training yourself to speak clearly but without strain.

Along with a good microphone and correct microphone placement, proper enunciation is the most significant factor in achieving high accuracy. You must not and need not strain your voice. The trick is, as you speak, feel your lips form every word and imagine someone facing you is reading your lips.

When I get tired, or lazy, accuracy immediately drops. I need to speak a bit slower and need to put a more distinct break between words.
Dictate the way and at the speed newscasters read the news. Be very clear and precise. Speak so that your third grade teacher would be proud of you. You are likely to see a marked improvement in accuracy. You need not sound like a newscaster, try to speak like one. Be clear and precise in your enunciation.

It is more effective to dictate from notes when you first start to develop good dictation habits. Dictation to a computer is not natural for most people. You must work to develop the skill to dictate with good enunciation.

Even after months of practice, I still have lower accuracy when I dictate without having thought through what I want to say and put some of the points down on paper.
Tips

- • Have an outline of what you wish to dictate.
- • The first document you dictate should be simple like a letter to a friend.

- Correct phrases rather than single words.
- Always use the correction window to correct misrecognitions.
- Edit text by deleting and dictating the new text.
- It is not helpful to stare at the screen when dictating. Focus on your notes or thoughts.
- Learn to get a flow of words going.

DETERMINING ACCURACY

It is worth doing a real count on accuracy, rather than just eyeballing the mistakes. There's a story about a fellow who was complaining about all the errors his speech program was making, and when an actual accuracy assessment was made, he found he was getting 98 percent accuracy. Meanwhile, others are thrilled to get 95 percent. This story points to the importance of expectations and perceptions. An accuracy count can serve as a baseline and when performed later can show progress that will motivate you to keep plugging along.

Read or dictate five paragraphs. Do not correct the errors. Select the errors and boldface them so that you can count them easily (print this out and save it for future reference). In Microsoft Word a word count is found by clicking the Review menu at the top and then click Word Count. Make note of this number. Now count the errors. Subtract the errors from the total words and divide this number by the total words. Example: six errors, 300 words. 300-6 = 294. 294/300 = 98%. There are of course other ways to do the math. Future accuracy measurements should follow the same methodology.

Count as a single error instances in which a short phrase has several errors; allow some forgiveness given the program's use of context. Don't count strange words or proper names that are missed as errors if the program has not learned them yet.

Do this every few days. You should gain encouragement from your visible progress and improvement.

Additional Resources to gain highest accuracy

1. *The* *WSRToolkit*
   The best $29.99 software investment you can make. Features include:

   a) **Text Macros** – issue a simple command to insert a block of text

   b) **Command Macros** - Step by step keystroke commands

   c) **Macro Editor Window** for editing or creating script macros

   d) **Train from Text** allows you to enter text of your choosing and read it to train the 'Acoustic Model' of your speech profile.

e)   Add to Dictionary provides an easy way to add words or phrases to your personal speech dictionary

f)   **Add From File** increases accuracy by parsing your reports and documents typical of your dictation style. The way you use words, their context to surrounding words, is used to improve your personal speech profile.

g)   **Transcription** reads your .wav or .mp3 file and transcribes it to text. This is only for a single user who speaks into a high quality recorder as if they were sitting in front of the computer

Get the WSRToolkit at:
https://mymsspeech.com/

2.   _WSRMacros: The User's Guide_ by Brad Trott
Uses easy to follow examples to create sophisticated command macros in WSR.
https://mymsspeech.com/

3.   _Forums_

The MSSpeech-Forum **is an excellent resource for discussions, questions and useful files and macros for download. There are sections for New Users, various Professions, Macros, Microphones, etc.**
http://www.msspeech-forum.com/