

# Speech Recognition (Windows XP)

## Microsoft © 2005

### Speech overview

Speech capabilities for a computer system refers to the ability to play back text in a spoken voice (referred to as text-to-speech or TTS), or to convert a spoken voice into electronic text (referred to as speech recognition or SR). The two capabilities are independent of each other. Many systems will have only TTS. SR may be installed later either by loading a speech package, or more commonly, through an application which has incorporated speech into it. For example, a new word processor or office tool suite may include speech and it will be loaded at that time.

In general, speech-enabled applications will use Speech properties in Control Panel to access and control features. In this way, speech may be customized for your personal preferences or a specific office location. Since TTS and SR may be loaded at different times, it is possible that not all the Help will be applicable. In cases where SR is not available, references to SR help and SR procedures may be ignored.

### Speech recognition overview

Speech recognition (SR) is the ability of the operating system to convert spoken words to written text. An internal driver, called an SR engine, recognizes words and converts them to text. The SR engine may be installed with the operating system or at a later time with other software. During the installation process, speech-enabled packages such as word processors and web browsers, may install their own engines or they can use existing ones. Additional engines are also available through third party manufacturers. These engines often use a certain jargon or vocabulary; for example, a vocabulary specializing in medical or legal terminology. They can also use different voices allowing for regional accents such as British English, or speak a different language altogether such as German, French or Russian.

You need a microphone or some other sound input device to receive the sound. In general, the microphone should be a high quality device with noise filters built in. The speech recognition rate is directly related to the quality of the input. The recognition rate will be significantly lower or perhaps even unacceptable with a poor microphone. The Microsoft Speech Recognition Training Wizard (Voice Training Wizard) guides you through the process and recommends the best position to place the microphone allowing you to test it for optimal results.

Once you have installed the system and it is working, it is important to train it for your environment and speaking style. On the **Speech Recognition** tab, click **Train Profile** and use the Voice Training Wizard to train the system to recognize background noises such as a fan, the hum of air conditioning, or other office sounds. It adapts to your speaking style including accents, pronunciations and even idiomatic phrases.

---

## Using the Microsoft speech recognition engine

The Microsoft speech recognition engine enables you to insert text into a document using specific programs. You can dictate text in any Microsoft Office XP program, in Internet Explorer 5 or later, and in Outlook Express 5.0 or later. Other software programs might eventually support the Microsoft speech recognition engine. You cannot dictate text in Notepad at this time.

### ◆ Important

- Speech recognition engines are language-specific. The first three Microsoft speech engines that are available are Simplified Chinese, U.S. English, and Japanese. Engines for other languages will become available.

Along with being language-specific, some speech engines may be region-specific. For example, the Microsoft English ASR Version 5 engine is intended for speakers of U.S. English. British, Australian, and other non-U.S. English speakers may have difficulty using it due to variations in accent.

To use speech recognition, you should have a high-quality close-talk (headset) microphone and a sound card or USB port. You should be able to position the microphone so it is close to your mouth. Use the Microphone Wizard to configure your microphone.

It is important to train the speech recognition engine to understand your voice. As you read the training text aloud, the engine looks for patterns in the way you speak to help it interpret the words you say. Training creates a speech profile for the individual speaker.

Speech recognition is not designed for completely hands-free operation; you'll get the best results if you use a combination of your voice and the mouse and keyboard, or your voice and a handwriting input device.

For more detailed information about using Microsoft speech recognition, click the **Help** button on the Language bar.

If you are using a tablet computer, you can use Tablet PC Input Panel for handwriting and speech tasks.

### Speech Recognition Tips

Speech recognition is not designed for completely hands-free operation; you'll get the best results if you use a combination of your voice and the mouse or keyboard. Also a consistent quality of speech results in the best results. When speaking to others, we usually understand from the context and environment even when whispered, shouted, or talking quickly or slowly. However, speech recognition understands words best when spoken to in a more predictable manner.

- Speak in a consistent, level tone. Speaking too loudly or too softly makes it difficult for the computer to recognize what you've said.
- Use a consistent rate, without speeding up and slowing down.

- Speak without pausing between words; a phrase is easier for the computer to interpret than just one word. For example, the computer has a hard time understanding phrases such as, "This (pause) is (pause) another (pause) example (pause) sentence."
  - Start by working in a quiet environment so that the computer hears you instead of the sounds around you, and use a good quality microphone. Keep the microphone in the same position; try not to move it around once it is adjusted.
  - Train your computer to recognize your voice by reading aloud the prepared training text in the Voice Training Wizard. Additional training increases speech recognition accuracy.
  - As you dictate, do not be concerned if you do not immediately see your words on the screen. Continue speaking and pause at the end of your thought. The computer will display the recognized text on the screen after it finishes processing your voice.
  - Pronounce words clearly, but do not separate each syllable in a word. For example, sounding out each syllable in "e-nun-ci-ate," will make it harder for the computer to recognize what you've said.
- 
- 

### **Speech setup overview**

For speech systems to work properly or for optimal results, the components need to be set up correctly. Speech capabilities have been designed to work with system defaults so that a minimal effort is required on your part. Apart from physically connecting speakers and microphones, all other aspects are intended to work automatically. Additionally, some systems are equipped with built in devices and it is possible that no configuration is required.

The first step for proper installation, or double-checking an existing installation, is following the set up guidelines presented here. If the system is not working properly, review the appropriate troubleshooting section.

---

---

### **Speech recognition requirements**

To use speech recognition, you need the following hardware and software components:

- A high-quality close-talk (headset) microphone. A Universal Serial Bus (USB) microphone is recommended.
  - A 400 megahertz (MHz) or faster computer with 128 megabytes (MB) of memory.
  - Microsoft Internet Explorer 5.0 or later.
  - The speech recognition engine must be installed. It is available with Microsoft Office XP, but might not have been installed.
  - Speech recognition engines are language-specific. The first three Microsoft speech engines that are available are Simplified Chinese, U.S. English, and Japanese. Engines for other languages will become available.
- 
-

## To add speech recognition

Speech recognition is normally added automatically as a text service when it is installed. Use this procedure only if you removed it as a text service and want to add it.

1. Open Regional and Language Options in Control Panel.
  2. On the **Languages** tab, under **Text services and input languages**, click **Details**.
  3. Under **Installed services**, click **Add**.
  4. In the **Input language** list, click an input language.
  5. Select the **Speech** check box, and then click an option in the list.
- Speech recognition engines are language-specific. The first three Microsoft speech engines that are available are Simplified Chinese, U.S. English, and Japanese. Engines for other languages will become available.
  - To open Regional and Language Options, click **Start**, point to **Settings**, click **Control Panel**, and then double-click **Regional and Language Options**.
  - Speech recognition requires computer memory and can affect your computer's performance, so you should remove it as a text service if you do not use it.
  - If speech recognition is not available, it may not be installed.
  - For more detailed information about using speech recognition, click the **Help** button on the Language bar.
  - If a language does not appear in the **Input language** list, the fonts for that language might not be installed.

## Install Speech Recognition

If you recently purchased a new computer, the Microsoft speech recognition engine might have been installed by the computer manufacturer. No further installation is necessary. If you installed Microsoft Office XP, or if you purchased a new computer with it installed, the speech recognition engine is included, but might not be installed. Follow one of the instructions below to install it.

☞ To determine if the speech recognition engine is installed

- Click **Start**, point to **Settings**, click **Control Panel**, and then double-click **Speech**.


If the **Speech Recognition** tab is available in Speech Properties, the speech recognition engine is installed. If it is not available, the engine is not installed.

To install speech recognition from Microsoft Word 2002

1. Open Microsoft Word.
2. On the **Tools** menu, click **Speech**.

Speech recognition is now available in all Office programs as well as other programs for which it is enabled, such as Microsoft Internet Explorer.

## To install speech recognition using Add or Remove Programs

1. Open  Add or Remove Programs in Control Panel.
2. Click **Change or Remove Programs**, click **Microsoft Office XP**, and then click **Change**.
3. Click **Add or Remove Features**, and then click **Next**.
4. Under **Features to install**, double-click **Office Shared Features**.
5. Double-click **Alternative User Input**, click **Speech**, click the down arrow, and then click **Run from My Computer**.
6. Click **Update**.

### Important

- Speech recognition engines are language-specific. The first three Microsoft speech engines that are available are Simplified Chinese, U.S. English, and Japanese. Engines for other languages will become available.

### Notes

- To open Add or Remove Programs, click **Start**, point to **Settings**, click **Control Panel**, and then double-click **Add or Remove Programs**.
- If you are using a tablet computer, you can use Tablet PC Input Panel for handwriting and speech tasks.
- For more detailed information about using speech recognition, click the **Help** button on the Language bar.

---

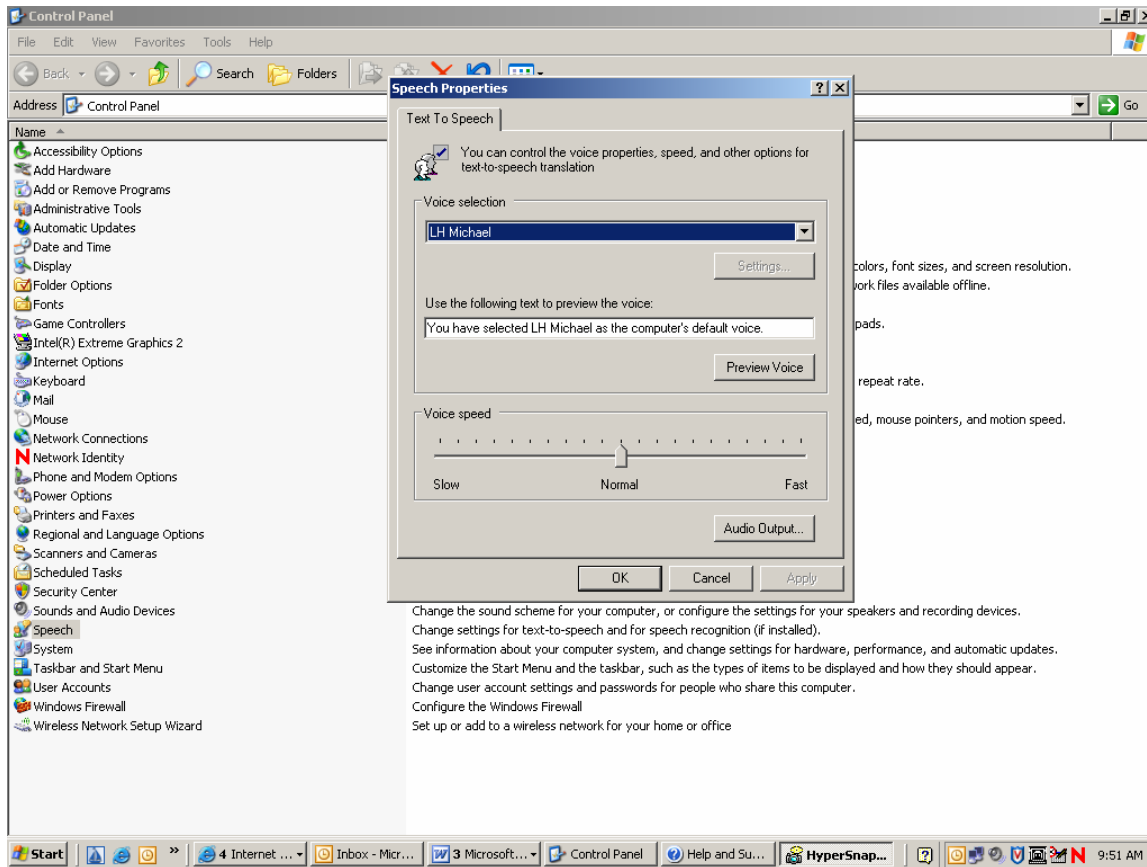
## To determine the selected speech recognition engine

1. Open **Speech** in Control Panel.
2. On the **Speech Recognition** tab, the name displayed in the Language drop-down list box is the active engine.

### Notes

- The language supported by an engine may not be obvious by the engine's displayed name. Refer to the specific user's guide for detailed information about the engine. This includes not only the language supported, but also the lexicon purpose. The lexicon purpose indicates whether it is a general grammar or jargon specific to a profession such as legal or medical.
- To open a Control Panel item, click **Start**, point to **Settings**, click **Control Panel**, and then double-click the appropriate icon.

## Start → Control Panel → Speech



Note: No speech recognition engine has yet to be installed.

Next steps:

- See requirements such as the install of a microphone
- Install Microsoft Windows speech recognition engine following the procedures below.

## To set up a microphone

Microphones vary greatly in design and purpose. They will continue to develop and become more specialized. Consult the microphone manual for hardware and software specifics; however, most models install in a similar fashion.

1. Locate the sound connections and connect the microphone jack to the computer. Most computers use an internal sound card and often the connections will be in the back of the system.
2. These will be a series of connections the same size and diameter as the microphone jack. One will be labeled as the microphone connection, either with a small icon that looks like a microphone or explicitly labeled as such.
3. Plug the microphone into that connection.

To test the connection:

1. Open **Speech** in Control Panel.
2. Select the **Speech Recognition** tab.
3. Speak directly into the microphone. The sound level should register in the Microphone Level indicator.
  - If no sound is received after connecting the microphone, See [Possible Speech Recognition Problems](#) for troubleshooting procedures.
  - To open a Control Panel item, click **Start**, point to **Settings**, click **Control Panel**, and then double-click the appropriate icon.

## To set up microphone options

The audio input line may be selected.

1. Open **Speech** in Control Panel.
2. On the **Speech Recognition** tab, click **Audio Input**, and then click **Properties**.
3. Select either **Use automatically chosen line** or **Use this audio input line**.

**Use automatically chosen line** sets the input line to a default that is determined by the speech system. Because of differences in drivers, capabilities and languages used, there might be variances in the selected option. The selected default may not work with all options. If the device line does not work properly, you should manually select a new line using **Use this audio input line**.

**Use this audio input line** allows you to select another line for audio input. The options present all audio line possibilities for the system. Not all audio lines are supported for speech.

- To open a Control Panel item, click **Start**, point to **Settings**, click **Control Panel**, and then double-click the appropriate icon.

## MS Word → Tools → Speech

After clicking Speech, the following tool bar displays.



## Language bar overview

The Language bar is a floating toolbar that appears on your [desktop](#) automatically when you add handwriting recognition, speech recognition, or an [Input Method Editor \(IME\)](#) as a method of inserting text. In addition, if you add a second language or keyboard layout, you can display the Language bar from the [taskbar](#).

The Language bar enables you to easily switch between and perform tasks related to entering text. For example, if you are entering text with a [handwriting input device](#), use the buttons on the Language bar to open the Writing Pad window, from which you can insert handwritten text into your document.

The buttons and options that are displayed on the Language bar depend on the [text services](#) you have installed, and the software program that is currently [active](#). For example, Microsoft Word 2002 supports speech recognition, while Notepad does not. If both programs are running, the speech buttons are displayed when Microsoft Word is active, but disappear when you make Notepad the active program.

You can move the Language bar anywhere on your screen, minimize it to the taskbar, or make it nearly transparent. If you do not use it, you can close it. You should remove text services that you do not use, as they require computer memory and can affect performance.

If you are using a tablet computer, you can use Tablet PC Input Panel for handwriting and speech tasks.

---

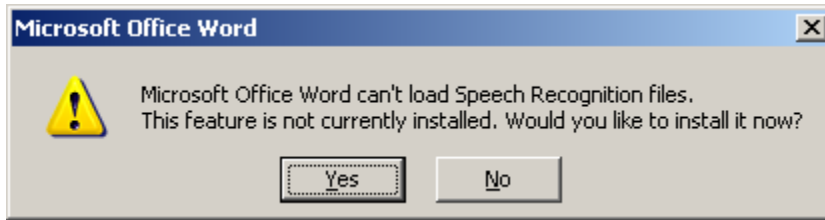
## To hide the Language bar



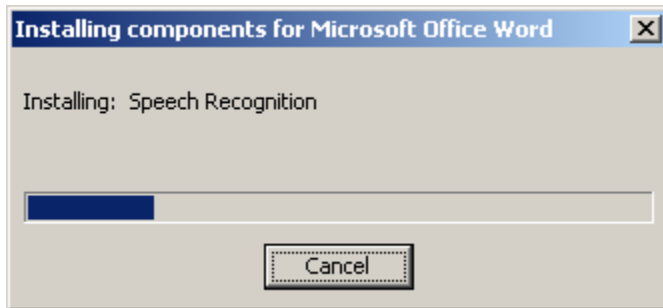
1. Right-click the Language bar.
2. Do one of the following:
  - Click **Minimize** to reduce the Language bar to an icon on the [taskbar](#).

This option removes the Language bar from your [desktop](#) but keeps it active and available.
  - Click **Close the Language bar** to turn off the Language bar and remove it from your desktop.
- To perform Language bar functions without having the Language bar on your desktop, right-click the Language bar and then click **Additional icons in taskbar**. You can then minimize the Language bar and click the text service icons on the taskbar to perform actions.
- Closing the Language bar does not remove any [text services](#). If you are not using a service, such as handwriting or speech recognition, you should remove it. Text services require computer memory and might affect performance.
- You can also minimize the Language bar by clicking **Minimize** on the Language bar.
- You cannot close the Language bar if you have installed an [Input Method Editor \(IME\)](#).
- If you are using a tablet computer, you can use Tablet PC Input Panel for handwriting and speech tasks.

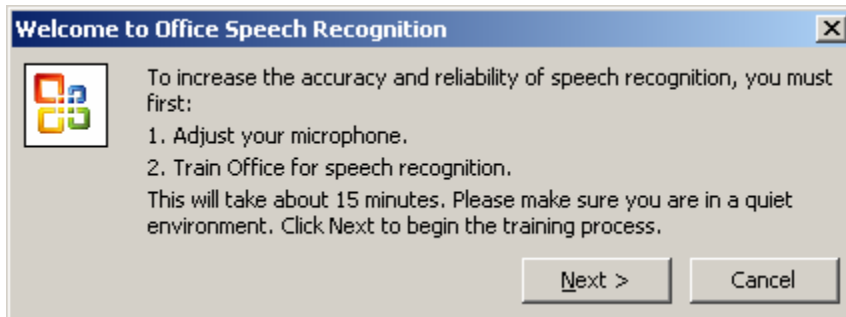
1. If the speech recognition engine has not been installed, then this dialog box displays.



2. Click "Yes" and the Microsoft Office Word speech recognition begins installing.

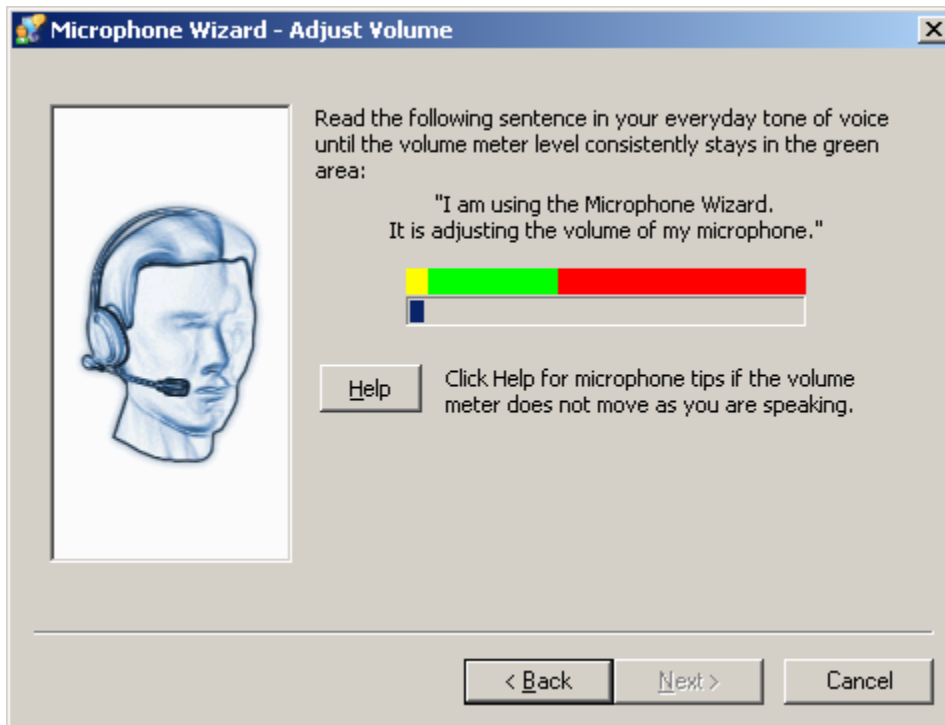


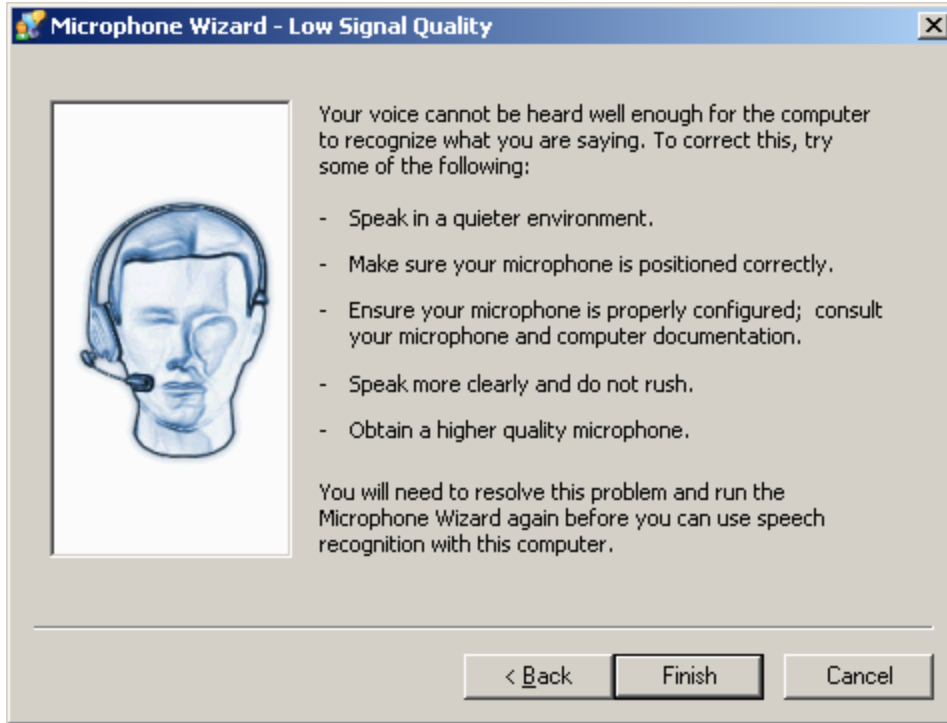
3. After the install routine, the Welcome to Office Speech Recognition box displays.



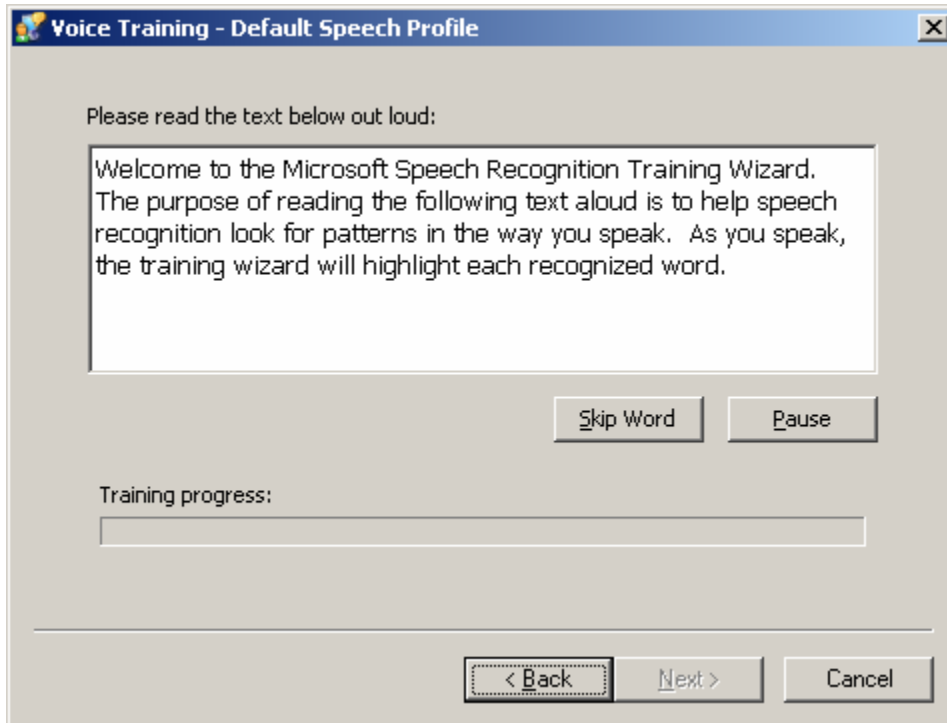
**3. The next two steps are to go through the wizard to adjust the microphone and then the next step is to train Microsoft Office for speech recognition.**

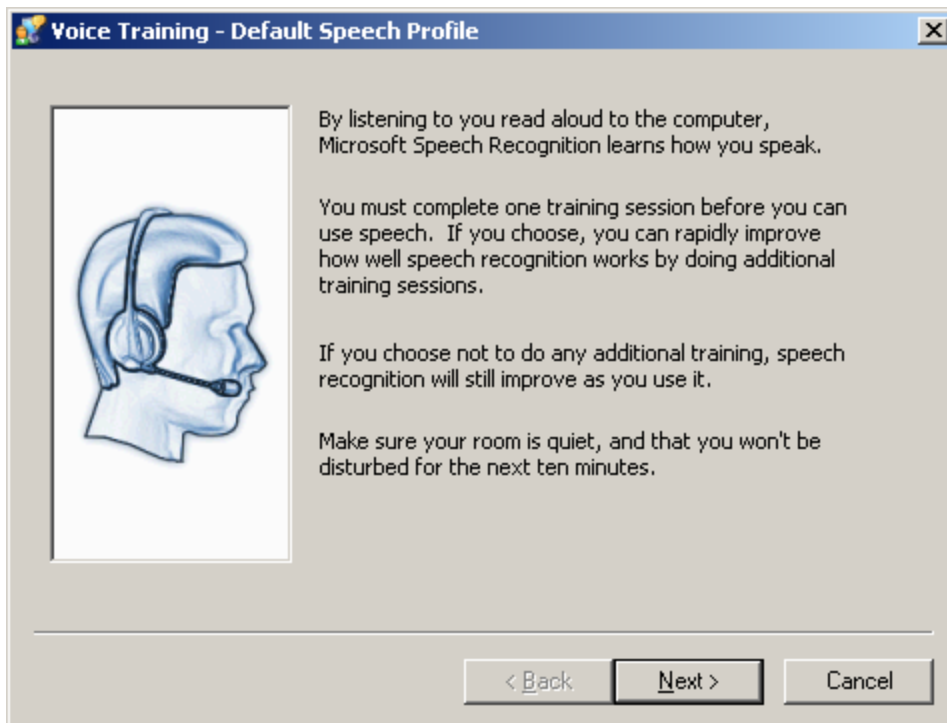
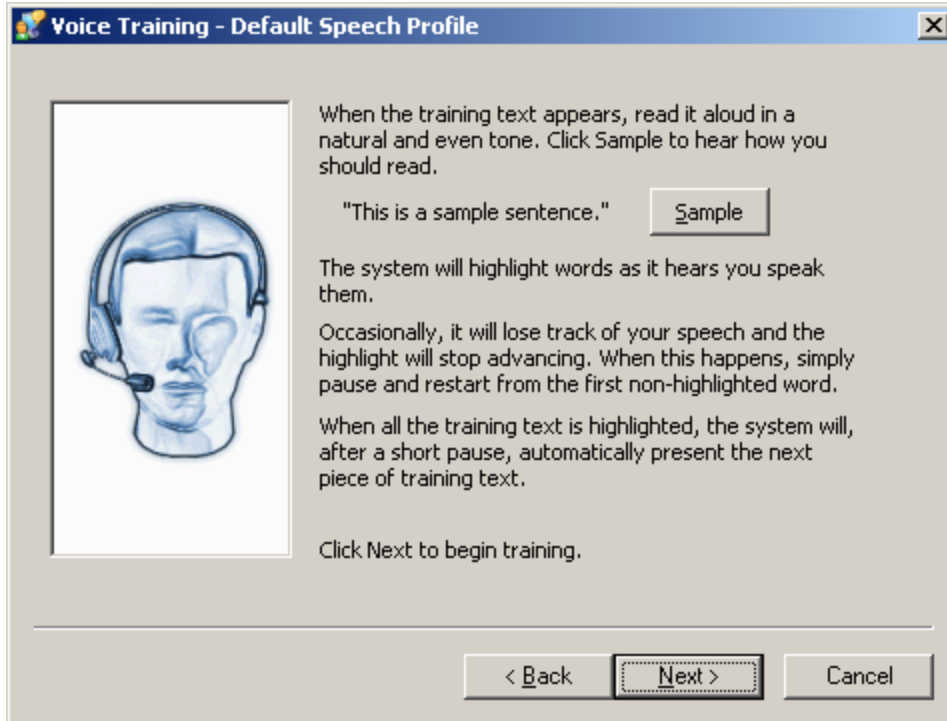
**First, complete the Microphone Wizard steps.**





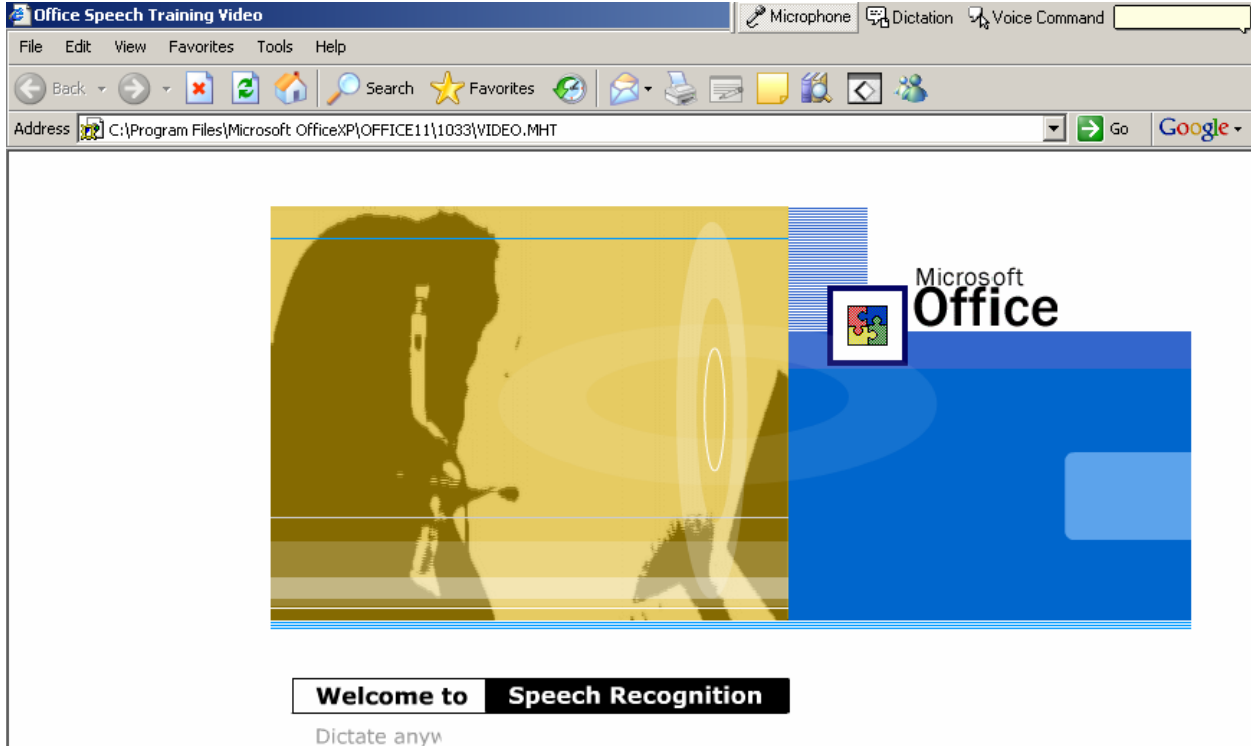
#### 4. Microsoft Voice Training Wizard: Default Speech Profile





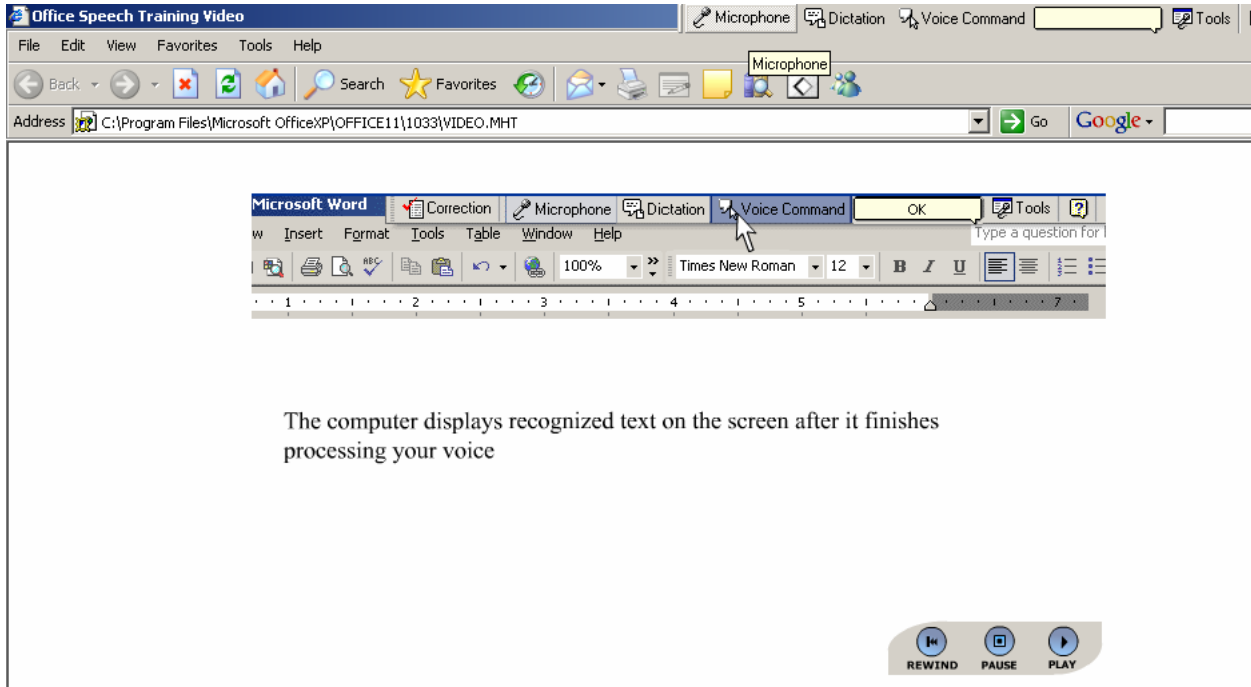
**Microsoft Office online video streaming demo of Speech Recognition software**

- 4. Then go to the Office Speech Training Video to observe the features and functionality of speech recognition.
- 5.



**Observe the Speech Recognition video and follow the steps as described.**

The speech training video takes you through each of the features and functionalities in a very clear way.



## To change speech recognition options

These options control the methods the speech recognition engine uses to process voice inputs.

1. Open **Speech** in Control Panel.
2. On the **Speech Recognition** tab, click **Settings** in the Recognition Profiles area.
3. Select **Pronunciation Sensitivity**, **Accuracy vs. Recognition Response Time**, or **Background Adaptation**.

**Pronunciation Sensitivity** controls the level of confidence needed for the system to respond to your command. High sensitivity indicates that you want the system to reject any command you utter when it is not confident of what you said. This means the system will make fewer recognition errors, but it will also increase the frequency of rejecting what you said. As a result, you may need to enunciate more slowly and clearly. Low sensitivity indicates that you want the system to respond to your command when it has little confidence that it has correctly recognized what you said. This will result in the system making more mistakes in recognizing your commands, but it will very rarely reject your command. This setting affects only command and control applications.

**Accuracy vs. Recognition Response Time** controls the trade off between accuracy in recognizing speech and computer processing time needed to generate the recognized speech. Low/Fast indicates that the system will perform limited processing. This results in the recognized text appearing on the screen quickly, but with a low accuracy rate. High/Slow indicates that the system will perform a much larger amount of computation to obtain the best accuracy possible, but at the expense of producing the dictated text much more slowly. This setting affects both command and control applications and dictation applications.

**Background Adaptation** controls whether the system adapts to your voice and speaking manner. When this box is checked, the system will learn what your voice sounds like (in a particular environment). This information is stored in your profile and contributes to an improvement in recognition accuracy. It is highly recommended that you select this setting.

4. To accept a new option, click **Apply**. To ignore any changes made in the dialog box and retain the current options, click **Cancel**. **Restore Defaults** overrides all changes and sets the options to a default state.

### Notes

- The features described here are unique to Microsoft speech recognition engines. Some engines provide additional and specialized features. However, special features are not a requirement and not all engines support them. If the currently selected engine does not support custom features, **Settings** in the Recognition Profiles group will be unavailable.
- To open a Control Panel item, click **Start**, point to **Settings**, click **Control Panel**, and then double-click the appropriate icon.

## To determine the selected speech recognition engine

1. Open **Speech** in Control Panel.
2. On the **Speech Recognition** tab, the name displayed in the Language drop-down list box is the active engine.

### Notes

- The language supported by an engine may not be obvious by the engine's displayed name. Refer to the specific user's guide for detailed information about the engine. This includes not only the language supported, but also the lexicon purpose. The lexicon purpose indicates whether it is a general grammar or jargon specific to a profession such as legal or medical.
  - To open a Control Panel item, click **Start**, point to **Settings**, click **Control Panel**, and then double-click the appropriate icon.
- 

## To change speech recognition engines

1. Open **Speech** in Control Panel.
2. On the **Speech Recognition** tab, the name displayed in the Language drop-down list box is the active engine. Click the active engine inside the Language drop-down list, or use the arrow to display a list of available engines.
3. Click a new engine to select it.
4. Click **OK** or **Apply** to accept the new engine.

### Notes

- If you select a different engine, it will not start until you stop all applications using speech.
  - The language supported by an engine may not be obvious by the engine's displayed name. Refer to the specific user's guide for detailed information about the engine. This includes not only the language supported, but also the lexicon purpose. The lexicon purpose indicates whether it is a general grammar or jargon specific to a profession such as legal or medical.
  - To open a Control Panel item, click **Start**, point to **Settings**, click **Control Panel**, and then double-click the appropriate icon.
-

## To train the speech recognition engine

1. Open **Speech** in Control Panel.
2. On the **Speech Recognition** tab, select the speech recognition engine you want from the Language drop-down menu.
3. Select the profile you want from the Recognition Profile group. Training is specific to an engine and profile so that training one engine/profile set has no effect on any other engine/profile set.
4. Click **Train Profile**. The Voice Training Wizard will appear; follow the directions in the wizard. Not all engines support training. If your engine does not, **Train Profile** will be unavailable.

### Notes

- To open a Control Panel item, click **Start**, point to **Settings**, click **Control Panel**, and then double-click the appropriate icon.
  - It is recommended that you spend at least 15 minutes training the computer. The more training you do, the higher your recognition accuracy will be.
- 
- 

## Troubleshooting Speech Systems

To determine if either speech system (text-to-speech or speech recognition) is working properly, use the following tests.

### For text-to-speech (TTS)

1. Open **Speech** in Control Panel.
2. On the **Text-to-Speech** tab, click **Preview Voice**.
3. The text in Preview Voice should be spoken audibly with each word highlighted in turn. If so, TTS and the speakers are working.
4. If you cannot hear the Preview Voice and see words highlighted as they are spoken, see [Possible Text-to-Speech Problems](#) for troubleshooting solutions.

## For speech recognition (SR)

1. Open **Speech** in Control Panel, and then click the **Speech Recognition** tab.
2. Speak into the microphone. A properly working microphone will display the input volume in the display area. The level will increase and decrease as the speaking volume changes.

### Notes

- To open a Control Panel item, click **Start**, point to **Settings**, click **Control Panel**, and then double-click the appropriate icon.
- 

## Possible text-to-speech problems

If no speech is heard after testing the system consider the following:

1. The speaker volume is not turned up or is muted. Some speakers have external controls for volume and muting. Make sure the volume is turned up sufficiently or that muting is off.
2. The speakers may not be selected as the current output device. On the **Text-to-Speech** tab, click **Audio Output** to make sure that the speakers are selected.
3. The speakers might not be connected properly. Consult the speaker hardware documentation for additional information. Make certain that the sound card for the computer is also properly seated, installed and that the correct drivers are available. The text-to-speech engine may be corrupted. To test for corruption, switch to another engine. If another engine is working properly, reinstall the specific engine from the original source. If no engine is working properly, reinstall the SAPI component from the original source or CD.

### Notes

- To open a Control Panel item, click **Start**, point to **Settings**, click **Control Panel**, and then double-click the appropriate icon.
- 

## To change the text-to-speech voice or engine

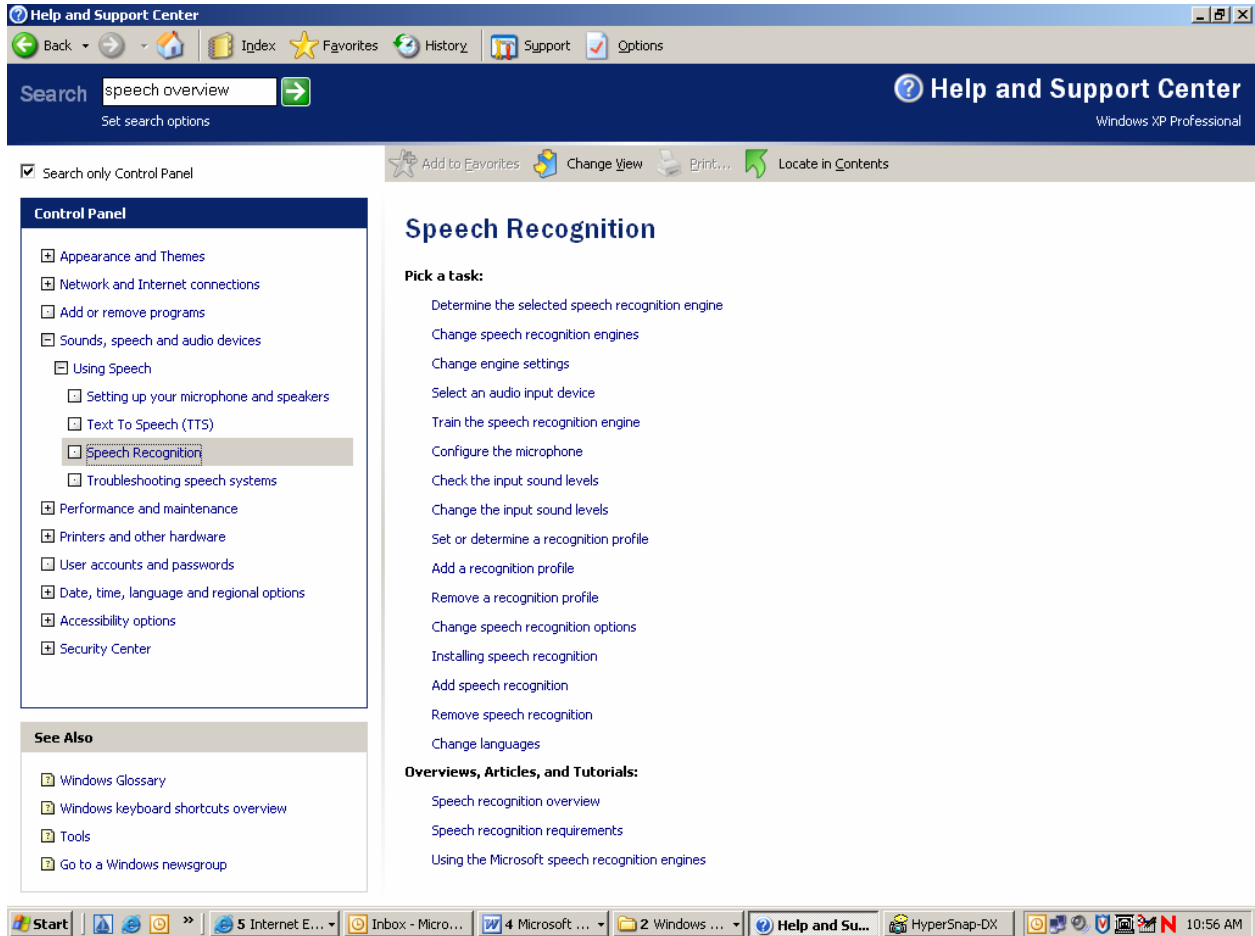
1. Open **Speech** in Control Panel.
2. On the **Text-to-Speech** tab, the name displayed in the Voice selection drop-down list box is the active voice.
3. Click the active voice inside the drop-down list box, or use the arrow to display a list of available voices.
4. Click a new voice to select it.
5. The newly selected voice will speak using the text in **Preview Voice** box.

6. Click **OK** or **Apply** to accept the new voice.

#### Notes

- A text-to-speech voice is closely associated with a particular engine. It may not be clear from the displayed name which language a voice is using. After selecting an engine or voice, test the voice and language by clicking **Preview Voice**.
- The language or voices supported by an engine may not be obvious by the engine's displayed name. Refer to the specific user's guide for detailed information about the engine. This includes not only the language supported, but also the lexicon purpose. The lexicon purpose indicates whether it is a general grammar or jargon specific to a profession such as legal or medical.
- To open a Control Panel item, click **Start**, point to **Settings**, click **Control Panel**, and then double-click the appropriate icon.

## If you experience any difficulties with Speech Recognition, then use Windows Help and Support Online.



## **Using Speech Properties**

You can use either the Speech Recognition tab or the Text-to-Speech tab in Speech properties to initialize and customize speech-enabled applications. These settings control general attributes such as input and output devices, the language used, as well as the playback voice and the accuracy of word recognition. Individual applications may implement speech differently and you need to check the respective user's manuals for particulars. However, some attributes are shared by speech engines or systems across different software applications. These shared features are set in Speech properties.

You can choose between several preferential option settings and customize the speech profile according to your needs. However, many options are proprietary and may vary between computer systems. Therefore, some of the buttons and screens may not act the same under all circumstances. Please check instructions for your computer, speech software or device hardware.

In addition to choosing pre-determined options, you may create a new profile or set a profile to accommodate your speaking style. A new Recognizer Profile should be created if you move offices, the noise level changes on a permanent basis or additional people are often present. This is done by clicking New on the Speech Recognition panel. Recognizer Profiles allow different users to share the same computer without interfering with other speech configurations.

Training allows the speech recognizer to adapt to the sound of your voice, word pronunciation, accent, speaking manner, and even new or idiomatic words. This is done through the Voice Training Wizard. Training for as little as ten minutes can achieve significant speech recognition improvement. The system also adapts to your speech on an ongoing basis and recognition will increase over time.

## Troubleshooting Speech Systems

To determine if either speech system (text-to-speech or speech recognition) is working properly, use the following tests.

### For text-to-speech (TTS)

1. Open **Speech** in Control Panel.
2. On the **Text-to-Speech** tab, click **Preview Voice**.
3. The text in Preview Voice should be spoken audibly with each word highlighted in turn. If so, TTS and the speakers are working.
4. If you cannot hear the Preview Voice and see words highlighted as they are spoken, see [Possible Text-to-Speech Problems](#) for troubleshooting solutions.

### For speech recognition (SR)

1. Open **Speech** in Control Panel, and then click the **Speech Recognition** tab.
2. Speak into the microphone. A properly working microphone will display the input volume in the display area. The level will increase and decrease as the speaking volume changes.
3. If no level is indicated, see [Possible Speech Recognition Problems](#) for troubleshooting solutions.

### Notes

- To open a Control Panel item, click **Start**, point to **Settings**, click **Control Panel**, and then double-click the appropriate icon.

---

## Possible speech recognition problems

If no sound is detected after testing the system, consider the following:

1. The microphone is muted. Some microphones have an external control for muting. Make sure the mute is turned off.
2. The microphone input level is set too low. On the **Speech Recognition** tab, select **Configure Microphone**. Follow the directions in the subsequent Microphone Wizard to set the input level.
3. The microphone may not be selected as the current input device. On the Speech Recognition tab, click **Audio Input** to make sure that the microphone is selected.
4. It is possible that the microphone connections are not properly secured. Make sure the microphone wires are intact and unbroken. The connections to the computer may be loose;

unplug the wires and reconnect them making sure they fit snugly. See [Set Up a Microphone](#) for additional connection information.

5. Check the documentation for the microphone's hardware for information specific to the manufacturer. This may include making certain the sound card for the computer is also properly seated, installed and that the correct drivers are available.
6. The SR engine may be corrupted. To test for corruption, switch to another engine if you have one. See [Changing the SR engine](#). If the other engine is working properly, reinstall the specific engine from the original source. If no engine is working properly, reinstall the SAPI component from the original source or CD.

#### Notes

- To open a Control Panel item, click **Start**, point to **Settings**, click **Control Panel**, and then double-click the appropriate icon.