# 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 have only TTS. SR may be installed later either through an explicit loading of 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 of the Control Panel to access and control features. In this way, speech may be customized for your personal preferences or 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.

# **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 set the profile to accommodate your speaking style. This includes training 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 a training wizard. Training for as little as ten minutes, you can achieve significant speech recognition improvement. The system also adapts to your speech on a ongoing basis and recognition will increase over time. Recognizer Profiles allow different users to share the same machine without interfering with each other's setups.

# Speech set up 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 <u>troubleshooting</u> section.

### 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.
- 3. Plug the microphone into that connection.

#### To test the connection:

- 1. In Control Panel, double-click the **Speech** icon.
- 2. Select the **Speech Recognition** tab.
- 3. Speak directly into the microphone. The sound level should register in the Microphone level graph.

- If no sound is received after connecting the microphone, See <u>Possible</u> <u>Speech Recognition Problems</u> 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. In Control Panel, double-click the **Speech** icon.
- 2. On the **Speech Recognition** tab, click **Audio Input**, 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.

#### **Notes**

## To set up speakers

Speakers vary greatly in design and purpose. They will continue to diverge and become more specialized. Consult the speaker manual for hardware and software specifics. However, most models install in a similar fashion.

Locate the sound connections and connect the speaker jack to the computer. Most computers use an internal sound card and often the connections will be in the back of the system. These will be a series of connections the same size and diameter as the speaker jack. In many cases there will be two sound out connections.

- 1. One will be labeled as a line-out connection. Most speakers requiring a separate power supply (such as an electrical adapter or batteries) should use this connection. It is also used to export amplified sound to recording devices including recordable CDs and tape cassette systems.
- 2. The other connection is for the non-powered speakers. Since the signal is boosted by the computer, powered speakers may be damaged if connected.
- 3. Plug the speaker into the proper connection.

### To test the connection:

- 1. In Control Panel, double-click the Speech icon.
- 2. On the **Text To Speech** tab, click **Preview Voice** to hear the currently selected voice; the text will be spoken, highlighting the words as they are spoken. If the speakers are working properly, you will hear the spoken words over the speakers.

#### **Notes**

• If no sound is heard after connecting the speakers, See <u>Possible Speech</u> <u>Recognition Problems</u> for troubleshooting procedures.

## To set up speaker options

The audio output line may be selected. By default, this option is disabled. However, other speech engines may include advanced properties for audio line out options. If so, the button will be active. Follow instructions on the screen or those documented separately for the specific engine.

- 1. In Control Panel, double-click the **Speech** icon.
- 2. On the **Text To Speech** tab, click **Audio Output**.
- 3. Follow instructions presented on the screen.

### Notes

## To change languages

The language used by either speech recognition (SR) or text to speech (TTS) is dependent on the respective engine. A system may have multiple engines loaded at the same time. In addition, multiple engines of the same language may also be present. However, only one engine of each type (SR and TTS) can be active at a given time. You may decide a particular engine better meets your requirements and use that engine more often. Once an engine is set, it is not usually necessary to change it.

## To change an SR language:

- 1. In Control Panel, double-click the **Speech** icon.
- 2. On the **Speech Recognition** tab, change the SR language by <u>changing the SR engine</u>.

### To change a TTS language:

- 1. In Control Panel, double-click the **Speech** icon.
- 2. On the **Text To Speech** tab, change the TTS language by <u>changing the TTS</u> <u>engine</u>.

- A TTS 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, click **Preview Voice** to test the voice and language.
- 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 by the engine, 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.

## Text to speech overview

Text to speech (TTS) is the ability of the operating system to play back printed text as spoken words. An internal driver, called a TTS engine, recognizes the text and using a synthesized voice chosen from several pre-generated voices, speaks the written text. A TTS engine is installed with the operating system. 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.

The **Text To Speech** tab on Speech Properties in Control Panel presents the options for each TTS engine. See the individual Help topics for specific help. In addition to the general options, each engine can have a different set of specific features. For that reason, not all the resulting dialogs will look the same. It is possible that no special features have been included and some of the properties buttons will not have an associated dialog.

To determine the selected text-to-speech voice

- 1. In Control Panel, double-click the **Speech** icon.
- 2. On the **Text To Speech** tab, the displayed name in the voices drop-down box is the current active voice.
- 3. Click **Preview Voice** to hear the active voice; the text will be spoken, highlighting the words as they are spoken.

- A TTS 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.

## To preview the text-to-speech voice

- 1. In Control Panel, double-click the **Speech** icon.
- 2. On the **Text To Speech** tab, the displayed name in the voices drop-down box is the current active voice.
- 3. Click **Preview Voice** to hear the currently selected voice; the text will be spoken, highlighting the words as they are spoken. During playback, **Preview Voice** will change to **Stop**. Click **Stop** to interrupt the voice playback.

- You can change the text to be read by the **Preview Voice** by highlighting the text and typing in new text. These changes are not permanent and when you reopen Speech Properties or select a different voice, the text will reset to the default.
- 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. In Control Panel, double-click the **Speech** icon.
- 2. On the **Text To Speech** tab, the name displayed in the drop-down list box is the active voice.
- 3. Click the active voice inside the drop-down list, 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.

- A TTS 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.

## To change the text-to-speech voice rate

- 1. In Control Panel, double-click the **Speech** icon.
- 2. Select the **Text To Speech** tab.
- 3. Move the **Speed** slider to change the rate of the text-to-speech voice. By default, it is set to normal.
- 4. Click **Preview Voice** to hear the currently selected voice at the new rate; the text will be spoken, highlighting the words as they are spoken.

## Notes

## To select an audio output device

- 1. In Control Panel, double-click the **Speech** icon.
- 2. On the **Text To Speech** tab, click **Audio Output**.
- 3. Select either **Use preferred audio output device** or **Use this audio output device.**

Use preferred audio device sets the output device as the default for the system. Select this option if you want to use the same output device for speech as all other sound for the system. It is also the default option for Speech Properties. Often, computers will have only one output device, such as a pair of speakers. The default device is designated by Sounds and Multimedia in Control Panel as Sound Playback on the Audio tab.

**Use this audio output device** allows you to use another device for speech applications only. The drop-down box is active if other devices are available. From this drop-down list, select the device you want. This does not change the default device for other audio applications. For example, you may wish all speech output to go through your headset rather than the speakers.

#### **Notes**

To change text-to-speech format options

Often regional or national preferences dictate different formats for calendar and monetary designations. Because text to speech performs text normalization (changing text into numeric or monetary units automatically), these preferences are set using this option.

- 1. In Control Panel, double-click the **Speech** icon.
- 2. On the **Text To Speech** tab, click **Settings.**
- 3. Select either **Decimal Point and Number Separator Preference** or **The Short Form Date Order Preference.**
- 4. **Decimal Point and Number Separator Preference** allows you to select the placement of the comma and decimal point in numbers. This preference includes two options:

**Comma separator, period decimal point** sets the comma as a thousands designator. For example, the text "1,234.56" will be spoken as "one thousand two hundred thirty four point five six."

**Period separator, comma decimal point** sets the decimal point as a thousands designator. The text "1.234,56" will be spoken "one thousand two hundred thirty four point five six." Other combinations of commas and decimal points will be interpreted strictly as individual numbers, not as a numeric value. For example, if you use the **Period separator, comma decimal point** option, and the number "1,234.56", the number will be spoken "one comma two hundred thirty four dot fifty six."

- 5. **The Short Form Date Order Preference** allows you to select the format for the abbreviated date. For the date to be read in order of month, day and year, select the option **mm/dd/yy**. For example, if you select this option, the date 05/01/00 would be spoken "May first 2000." However, if you select the option **dd/mm/yy**, the same date would be spoken "January fifth 2000."
- 6. To accept an 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.

<ul> <li>To open a Control Panel item, click <b>Start</b>, point to <b>Settings</b>, click <b>Control Panel</b>, and then double-click the appropriate icon.</li> </ul>

## To change the text-to-speech volume

To adjust the volume output levels, follow the procedure below. However, not all devices support this option in the same way. Some devices will not support volume control and the **Volume** button will be unavailable. Other devices may use their own display. In those cases, follow the instructions presented on the screen or documented separately with the engine.

- 1. In Control Panel, double-click the **Speech** icon.
- 2. On the **Text To Speech** tab, click **Audio Output**, then click **Volume**.
- 3. A volume control mixer will be displayed. Adjust the appropriate device to the required level.

### Notes

# **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 OS 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 installation 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. Using the same training wizard, you can 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.

## **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 prepared training text in the Microsoft Speech Recognition 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.

To determine the selected speech recognition engine

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

- 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. In Control Panel, double-click the **Speech** icon.
- 2. On the **Speech Recognition** tab, the name displayed in the Language dropdown 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.

- 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 change engine settings

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 Language drop-down list box will be inactive. If it is active, you can use these procedures:

- 1. In Control Panel, double-click the **Speech** icon.
- 2. On the **Speech Recognition** tab, click **Settings** in the Language box and follow the instructions in the dialog box or the instructions provided by the engine vendor. Special features are dependent on the engine's manufacturer and will be documented separately.

### Notes

## To select an audio input device

- 1. In Control Panel, double-click the **Speech** icon.
- 2. On the **Speech Recognition** tab, click **Audio Input.**
- 3. Select either **Use preferred audio input device** or **Use this audio input device**.

Use preferred audio input device sets the input device to the default for the system. Select this option if you want your speech applications to use the same input device as all other sound for the system. Often, computers will have only one input device, such as the headset microphone. This is the default in Speech Properties. The default device is designated by Sounds and Multimedia in Control Panel as Sound Recording on the Audio tab.

Use this audio input device allows you to select another device for speech applications only. Select this option to change to another input device. The drop-down list box becomes active only if other devices are available. From this drop-down list, select the device you want. Making a selection here does not change the default device for other audio applications. For example, you may wish all speech input to use the headset microphone rather than a omnidirectional microphone.

#### **Notes**

## To train the speech recognition engine

- 1. In Control Panel, double-click the **Speech** icon.
- 2. On the **Speech Recognition** tab, select the speech recognition engine you want.
- 3. Select the profile you want. Training is specific to an engine and profile and training one engine/profile set has no effect on any other engine/profile set.
- 4. Click **Train Profile**. A training wizard will appear; follow the directions in the wizard. Not all engines support training. If your engine does not, **Train Profile** will be inactive.

#### Notes

## To configure the microphone

Getting the best results from your microphone depends on a variety of factors including the distance from mouth to microphone and correct positioning of the microphone. Each manufacturer addresses specific requirements.

- 1. In Control Panel, double-click th Speech icon.
- 2. On the **Speech Recognition** tab, click **Configure Microphone**; a training wizard appears. The features supported are dependent on the manufacturer of the engine. Follow the instructions in the wizard. **Configure Microphone** is active only if the current engine supports microphone training.

### Notes

## To check the input sound levels

- 1. In Control Panel, double-click the **Speech** icon.
- 2. On the **Speech Recognition** tab, the Microphone box displays a linear register of the input sounds. If the microphone is accepting input, a bar is displayed. This indicates the sound level. Normal speech should register slightly less than halfway on the graph.
- 3. If no sound registers, make sure the microphone is turned on; a few models have a mute switch. Also, double check the connections to the computer. Some systems have several locations where a microphone can be plugged in.

#### Notes

## To change the speech recognition volume

To adjust the volume input levels, follow the procedure below. However, not all devices support this option in the same way. Some devices will not support volume control and the **Volume** button will be unavailable. Other engines may use their own display. In those cases, follow the instructions presented on the screen or documented separately with the engine.

- 1. In Control Panel, double-click the **Speech** icon.
- 2. On the **Speech Recognition** tab, click **Audio Input**, then click **Volume**.
- 3. A volume control mixer will be displayed. Adjust the appropriate device to the required level.

### Notes

To set a profile or determine which profile is selected

- 1. In Control Panel, double-click the **Speech** icon.
- 2. On the **Speech Recognition** tab, the Recognition Profiles list box displays the currently active profile; that is, the one with the check mark next to it.
- 3. To select a different profile, click the one you want. The new choice will become highlighted and have a check mark next to it. Only one profile may be active at a time.

### Notes

## To add a profile

- 1. In Control Panel, double-click the **Speech** icon.
- 2. On the **Speech Recognition** tab, click **New** in Recognition Profiles. An Add Profile wizard will appear; follow the directions it presents. Profiles are specific to engines and not all engines support similar functions.

- Profiles accommodate your speaking style. This lets the speech system know about the sound of your voice, word pronunciation, accent, speaking manner, and background noises. A new profile should be created if you move offices, the noise level changes on a permanent basis or additional people are often present. Recognition Profiles allow multiple users to share the same machine without interfering with each other's setups.
- To open a Control Panel item, click **Start**, point to **Settings**, click **Control Panel**, and then double-click the appropriate icon.

## To remove a profile

- 1. In Control Panel, double-click the **Speech** icon.
- 2. On the **Speech Recognition** tab, select the profile you want to delete from the Recognition Profiles list box. This highlights the profile and selects it; **Delete** becomes active.
- 3. Click **Delete**.
- 4. Click **OK** or **Apply** to remove, otherwise click **Cancel**.

- An engine must have at least one profile associated with it. Unless otherwise changed, the default profile will be used automatically.
- To open a Control Panel item, click **Start**, point to **Settings**, click **Control Panel**, and then double-click the appropriate icon.

## 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. In Control Panel, double-click the **Speech** icon.
- 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 <u>Possible Text To Speech Problems</u> for troubleshooting solutions.

## For Speech Recognition (SR)

- 1. In Control Panel, double-click the **Speech** icon.
- 2. On the **Speech Recognition** tab, speak into the microphone.
- 3. A properly working microphone will display the input volume in the display area. The level will grow and shrink as the speaking volume changes.
- 4. If no level is indicated, see <u>Possible Speech Recognition Problems</u> for troubleshooting solutions.

### Notes

## Possible Text To Speech Problems

If no speech is heard after testing the system (see <u>Troubleshooting Speech Systems</u>), consider the following suggestions:

- 1. The TTS engine may be bad. Try switching to another engine. See <u>Changing the TTS engine</u>. 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.
- 2. 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.
- 3. Speakers are also controlled through software. Look in either **Sounds and Multimedia** or **Multimedia** properties item in the Control Panel, click the Audio tab and check the preferred device in the Sound Playback section.
- 4. The speakers might not be connected properly. Consult the speaker hardware documentation for additional information. Make certain the sound card for the computer is also properly seated, installed and the correct drivers are available.

#### **Notes**

## Possible Speech Recognition Problems

If no sound is entered after testing the system (see <u>Troubleshooting Speech Systems</u>), try the following suggestions.

- 1. The SR engine may be bad. Try switching to another engine. See <u>Changing the SR engine</u>. 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.
- 2. The microphone level is 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 is muted. Some microphone have an external control for muting. Make sure the mute is turned off.
- 4. Microphones are also controlled through software. Look in either **Sounds** and **Multimedia** or **Multimedia** properties item in the Control Panel, under the Audio tab and Sound Recording section. Additional information is available through the associated Help files.
- 5. It is possible the microphone connections are not properly made. Check the documentation for the Microphones hardware for additional information. This may also include making certain the sound card for the computer is also properly seated, installed and the correct drivers are available.

#### **Notes**