CHAPTER 1

■ ■ ■

Getting Started with Microsoft OCS 2007 Speech Server

Microsoft Office Communications Server (OCS) 2007 Speech Server is an optional component of OCS 2007, the successor of Live Communications Server. The Speech Server component is the successor of Microsoft Speech Server 2004, a stand-alone product. The Speech Server component can be installed separately from OCS 2007, from the OCS installation media. In this book, I will first cover how to create telephony applications using Speech Server separately, and then explain why the merging of these two products makes sense.

This chapter begins with an overview of Speech Server, including the new features of this version and how it works. Then you will learn how to install and configure Speech Server.

OCS 2007 Speech Server Overview

Speech Server is an interactive voice response (IVR) platform that integrates with Visual Studio 2005.

Speech Server provides tools for developing applications that run over a telephone, or *telephony applications*. For example, telephony applications let you check your bank balance via a telephone or get an automated call from your doctor's office reminding you of your next appointment. Speech Server is to a telephony application what a web server such as Internet Information Services (IIS) is to a web application.

Speech Server applications can have the following capabilities:

- Speech recognition allows users to respond to application prompts.

- Touch-Tone capabilities, called dual-tone multi-frequency (DTMF), let users respond to application prompts via the telephone keypad.

- Text-to-speech (TTS) capabilities allow applications to read and speak written text to users.

New Features of OCS 2007 Speech Server

As I noted at the beginning of this chapter, the Speech Server component of OCS 2007 is the successor to Microsoft Speech Server 2004. The new version offers a lot of additional features, including Voice over Internet Protocol (VoIP) support, new project types, and more tools.

VoIP Support

One of the more exciting new features of Speech Server is its native support for VoIP. VoIP basically allows users to place and receive calls over the Internet. Speech Server can accept VoIP calls without any additional software or hardware.

Project Types

Microsoft Speech Server 2004 supported only Speech Application Language Tags (SALT). The new version supports three types of projects:

- *SALT*: A World Wide Web Consortium (W3C) standard language for speech over the Web and telephone. It runs on top of ASP.NET pages.

- *VoiceXML*: VoiceXML follows the same web-based paradigm as its rival standard, SALT, also running on top of ASP.NET pages. You can now convert your current VoiceXML-based IVR application to Speech Server without a lot of coding changes.

- *Voice Response Workflow*: If you do not want to follow the web paradigm, you have a new alternative: Voice Response Workflow applications. This project type is based on Windows Workflow Foundation (WF). Unlike SALT and VoiceXML, Voice Response Workflow allows you to see your application's call flow.

If you are creating new IVR applications, Voice Response Workflow is highly recommended over SALT or VoiceXML, as all IVR applications generally follow a workflow. If you want to incorporate a SALT or VoiceXML page within a Voice Response Workflow application, you can do this with the SALT and VoiceXML interpreter controls, as you'll learn in Chapter 6.

New Tools

Several new tools are available with OCS 2007 Speech Server:

Conversational Grammar Builder and Grammar Design Advisor: The Conversational Grammar Builder allows you to easily and rapidly build grammar in a natural conversational flow. You still have the options of building your grammar either by writing your own Grammar XML (GRXML) or the Visual Studio Grammar Editor. Debugging support for your grammar has also been added, along with the Grammar Design Advisor, which provides warnings about possible incorrect grammar when you are compiling. Chapter 2 discusses grammar and grammar building.

Lexicon Editor: The Lexicon Editor allows you to add or change the pronunciation of words for the Conversational Grammar Builder, which affects how the Speech Recognizer Engine recognizes audio input from the user. This editor is discussed in Chapter 2.

Pronunciation Editor: Like the Lexicon Editor, the Pronunciation Editor allows you to add or change the pronunciation of words. However, this editor applies only to the Grammar Editor's grammar. The Pronunciation Editor is also covered in Chapter 2.

Analysis and tuning: The new analysis and tuning tools allow you to view details about your speech application, such as user responses that did not have a matching grammar. This will help you in making decisions for improving your speech application. Chapter 8 covers analyzing and tuning applications.

Business Intelligence Tools: If you have high call volume, you might want to consider installing the Business Intelligence Tools. These tools allow you to create SQL Server Integration Services (SSIS) packages in SQL Server 2005, for activities such as creating an online analytical processing (OLAP) cube. These tools are installed separately from Speech Server. You can install them from the Business Intelligence Tools directory of the OCS 2007 installation media.

How Speech Server Works

Speech Server has two main components: Speech Engine Services and ASP.NET. While IIS isn't technically a part of Speech Server, it plays a vital role in its architecture. Figure 1-1 shows a conceptual overview of how Speech Server works with other servers and with Visual Studio.
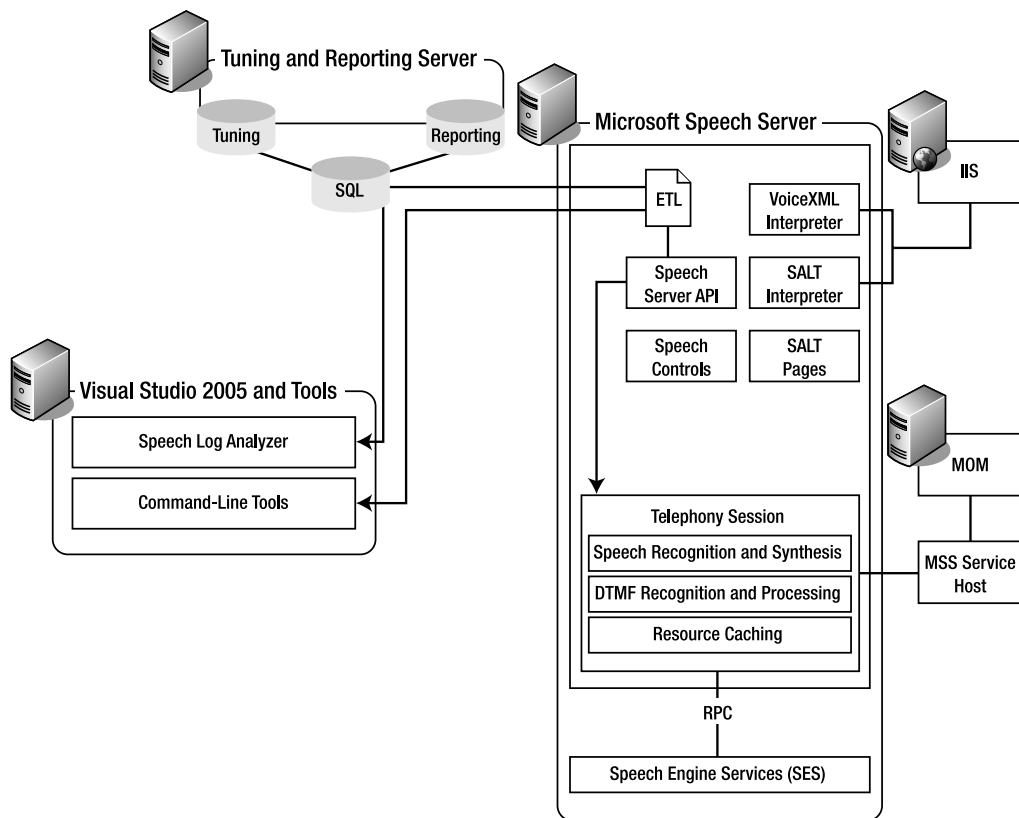
Figure 1-1. *How Speech Server works with servers and Visual Studio*

Speech Recognition and Speech Synthesis

Speech Engine Services (SES) has two main components: the Speech Recognizer Engine and the Speech Synthesis Engine.

The Speech Recognizer Engine produces an XML document in Semantic Markup Language (SML). This SML document contains the words or phrases the Speech Recognizer Engine recognizes and a numerical value for how confident it is that the user has said the word or phrase, based on the application's defined grammar. SML is discussed in Chapter 2.

The Speech Synthesis Engine, also known as the TTS engine, produces audio based on an XML document in Speech Synthesis Markup Language (SSML). The SSML document contains the text that the Speech Synthesis Engine will say. Before the Speech Synthesis Engine speaks, it checks the Prompt Engine. If the Prompt Engine contains the phrase or word that the Speech Synthesis Engine should say, it will play the associated prerecorded prompt instead of using the TTS engine. SSML is discussed in Chapter 3.
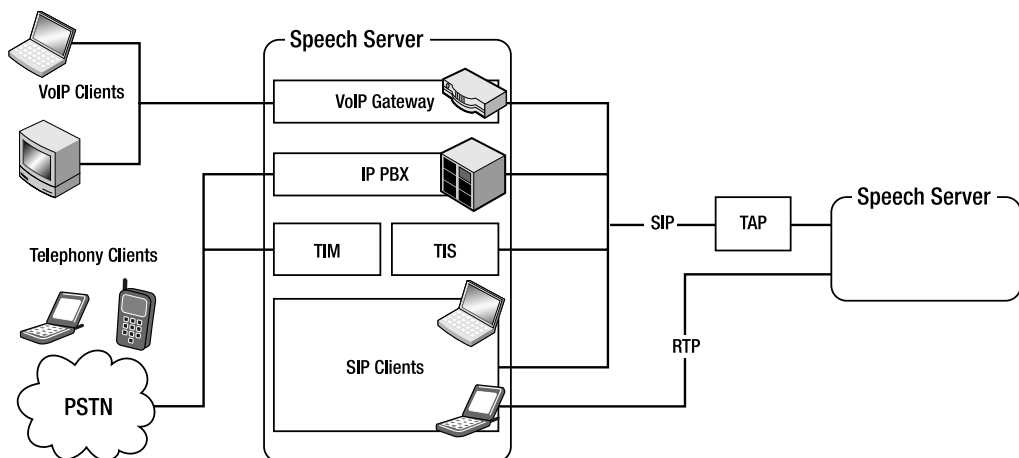
Client Connections

As noted earlier, Speech Server supports VoIP for telephone calls over the Internet. The underlying protocol that makes native VoIP support for Speech Server possible is Session Initiation Protocol (SIP).

For communications, first a client, such as Office Communicator, registers with an SIP server—in this case, OCS 2007—giving the SIP server certain data, such as its location, IP address, and the protocol it supports. When a user wants to place a call to another registered user, an INVITE message is sent to the registered user, based on the data that client gave the SIP server when it registered. Once the client accepts the INVITE, the two clients can communicate. The reason for the INVITE is that a client can be registered at multiple locations, such as a mobile device and a computer. The first location to accept the INVITE will receive subsequent messages.

If your network does not currently support VoIP, you will need to use the more traditional setup (as was the case with Speech Server 2004): a telephony card and the associated Telephone Interface Manager (TIM) software. Speech Server provides backward compatibility via the Telephony Interface Service (TIS), which serves as an interface between your TIM and the telephony application proxy (TAP). The TAP acts very much like a SIP redirect server, as it interprets requests to and from the TIS and other SIP peers, and routes them to the appropriate server.

Figure 1-2 shows a conceptual overview of how Speech Server works with client connections.

Figure 1-2. *How Speech Server works with client connections*

TELEPHONY APPLICATION ACRONYMS

The following are common acronyms you'll come across when working with telephony applications:

- *DTMF*: Dual-tone multi-frequency. Refers to Touch-Tone on a standard telephone.

- *IVR*: Interactive voice response. An automated telephone system with speech recognition capabilities.

- *PBX*: Private branch exchange. Allows for private companies to control their telephone lines inside their buildings, instead of having private telephone numbers for each extension.

- *PSTN*: Public switched telephone network. Provides access to outside telephone lines.

- *RTP*: Real-time Transport Protocol. A common protocol that allows real-time video and audio over User Datagram Protocol (UDP) transport.

- *SALT*: Speech Application Language Tags. A World Wide Web Consortium (W3C) standard language for generating speech over the Web and telephone.

- *SIP*: Session Initiation Protocol. A protocol to initiate VoIP and instant messaging (IM) sessions.

- *TIM*: Telephony Interface Manager. Provides an interface between a telephony board and Microsoft Speech Server.

- *TTS*: Text-to-speech. An automated component that turns text into speech, essentially reading text.

- *VoIP*: Voice over Internet Protocol. A protocol to carry telephone calls over an IP network, such as the Internet.

- *VUI*: Voice user interface. Describes an IVR application's presentation layer, similar to a graphic user interface (GUI) for a visual application.

Speech Server Installation

Speech Server requires certain hardware and software to run properly. Before proceeding with the installation, make sure that your system meets these requirements. If this is your first time using Speech Server, you should do an installation in a development environment before installing it in a production environment.

■**Caution** OCS 2007 Speech Server cannot run on the same machine as Microsoft Speech Server 2004. You must first uninstall Microsoft Speech Server 2004.

Installation Requirements

Speech Server can run under the following Microsoft operating systems:

- Windows XP Professional Service Pack (SP) 2

- Windows Vista Business, Enterprise, and Ultimate

- Microsoft Windows Server 2003 Standard Edition and Enterprise Edition

---

■**Caution** The Telephony Interface Manager Connector component is supported only on the server versions of Windows. It will not run on the client versions, such as Windows XP.

The system on which you install Speech Server determines which edition you are installing. For example, if you install it on a client operating system such as Windows XP, you are essentially installing a developer version, which is limited to two simultaneous connections.

Hardware Requirements

Speech Server has different hardware requirements depending on which edition you have installed, as shown in Table 1-1.

Table 1-1. *Microsoft Speech Server Hardware Requirements*

Hardware	Developer Edition	Server Edition
Processor	One 2.5GHz or faster	Two 2.5GHz or faster
Memory	1GB	4GB
Hard drive	3.5GB available for installation; 5GB minimum	3.5GB available for installation; 20GB minimum
Network interface card	Required	Required
Telephony board	Not required (only one per server supported)	Required for systems without VoIP (one per server supported)
Microphone	High-quality universal serial bus (USB) microphone (recommended)	High-quality universal serial bus (USB) (recommended)
Sound card	Recommended	Recommended

For a list of supported TDM and VoIP hardware, visit the Microsoft Speech Server web site, www.microsoft.com/speech/partners/default.mspx.

Software Requirements

The following are the software prerequisites for Speech Server, listed in the recommended order in which they should be installed:

- IIS

- Microsoft Message Queuing (MSMQ)

- Microsoft .NET Framework 3.0

- Visual Studio 2005 with SP1

- Visual Studio Extensions for Windows Workflow Foundation

■**Note** If you are installing Speech Server in your production environment, you do not need to install Visual Studio 2005 or the Visual Studio Extensions for Windows Workflow Foundation.

Installing Speech Server Components

The first step to installing Speech Server is running the setup.exe file from the installation media. After choosing Install, you must enter your user information and product key, and then accept the terms of the End User License Agreement (EULA) to continue the installation.


■**Tip** You should read the product release notes before installing any application. This file may contain directions or warnings pertaining to your particular setup.

Next, select the components you wish to install, as shown in Figure 1-3. If you are installing the 32-bit version, you can also select to install the Telephony Interface Manager Connector component, if a telephony board and the TIM are already installed on your system. Note that if you are using a VoIP-based setup, you do not need to install the Telephony Interface Manager Connector component.

Production machines must have the Server Components installed, and the Administrative Tools are recommended as well. You do not need to install the Development Tools or the Documentation on production machines. This book focuses on the development areas, including the Data Processing Utilities.

After you select and confirm the components, the Installation Wizard will install Speech Server.
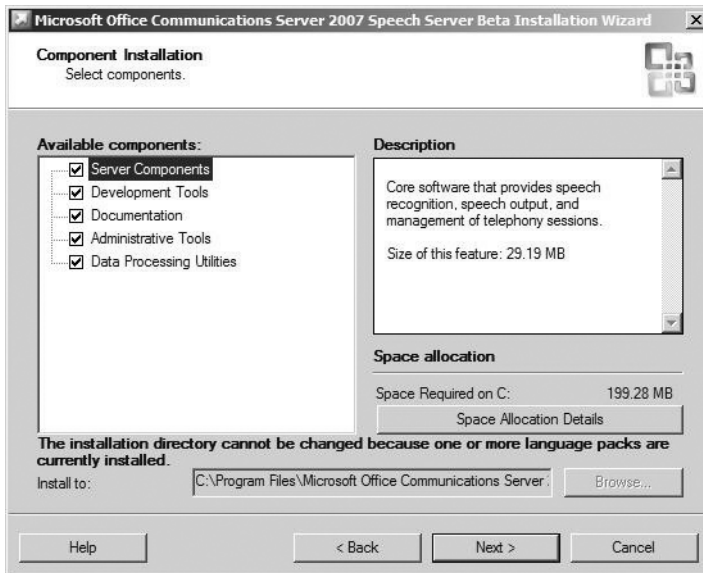
Figure 1-3. *Choosing Speech Server components to install*

When the component installation is complete, you will need to install a language pack. The language packs can be found on your installation media, in the language pack directory. Currently, the following language packs are available:

- English (United States)

- English (United Kingdom)

- German

- French

- Spanish

These language packs provide the necessary support for the language(s) your applications will be speaking. You must install the English (United States) language pack. If your application will support multiple languages, you will need to install all applicable language packs.

Configuring Speech Server

After installation, you may need to configure some components depending on your setup:

- If you have multiple servers, you may want to set up deployment groups. This will allow you to manage settings for every server in one spot.

- If you are using VoIP, you will need to set up the SIP peers before Speech Server will accept any calls from your VoIP gateway.

You perform most Speech Server configuration and management tasks through the Administrator Console. To open the Administrator Console, select Start ➤ Programs ➤ Microsoft Office Communications Server 2007 Speech Server. The Administrator Console is divided into two panes: the left pane lists nodes for Servers, Applications, and SIP Peers, and the right pane shows a details view of the selected node.
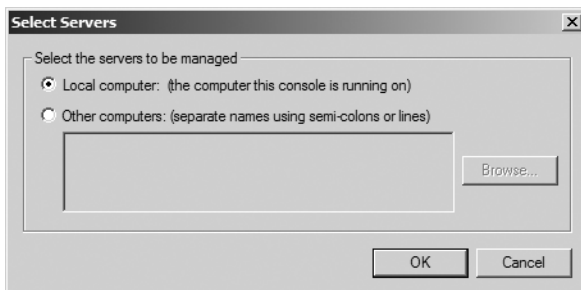
Setting Up Deployment Groups

Deployment groups allow you to easily manage multiple servers at the same time, so you don't need to adjust settings for each individual server. You can also copy settings from a server to the other servers in the same group.

Adding a Deployment Group

You can set up deployment groups through the Administrator Console, as follows:

1. Select Start ➤ Programs ➤ Microsoft Office Communications Server 2007 Speech Server to open the Administrator Console.

2. To create a new group, right-click the topmost node and choose New Group. You will be prompted to enter the name of the group you wish to add.

3. To add servers to the deployment group, right-click the Servers node in your deployment group and choose Add Server. You will be prompted to enter the servers' names, as shown in Figure 1-4.

**Figure 1-4.** *Adding servers to a deployment group*

Copying Server Settings

After you have added multiple servers to the same deployment group, you can copy server settings to all of these servers. This is useful when you have a distributed deployment model for your production environment, as you do not need to manually change the settings for each group.

To copy the server settings to the group, select the Servers node in the Administrator Console, and then click the Copy Server Settings to Group link, as shown in Figure 1-5. You will be prompted for which settings you want to copy. As shown in Figure 1-6, you can copy the properties, application settings, and the trusted SIP peers.

Figure 1-5. *Viewing your servers*
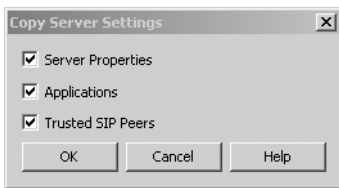
**Figure 1-6.** *Copying server settings*

Setting Up SIP Peers

For Speech Server to accept or place phone calls, you need to first set up a SIP peer. SIP peer configuration is done through the Administrator Console.

Adding a New SIP Peer

To create a new SIP peer, open the Administrator Console, right-click the SIP Peers node, and choose New ➤ SIP Peer, as shown in Figure 1-7. You will be prompted to enter information about the SIP peer, as shown in Figure 1-8.
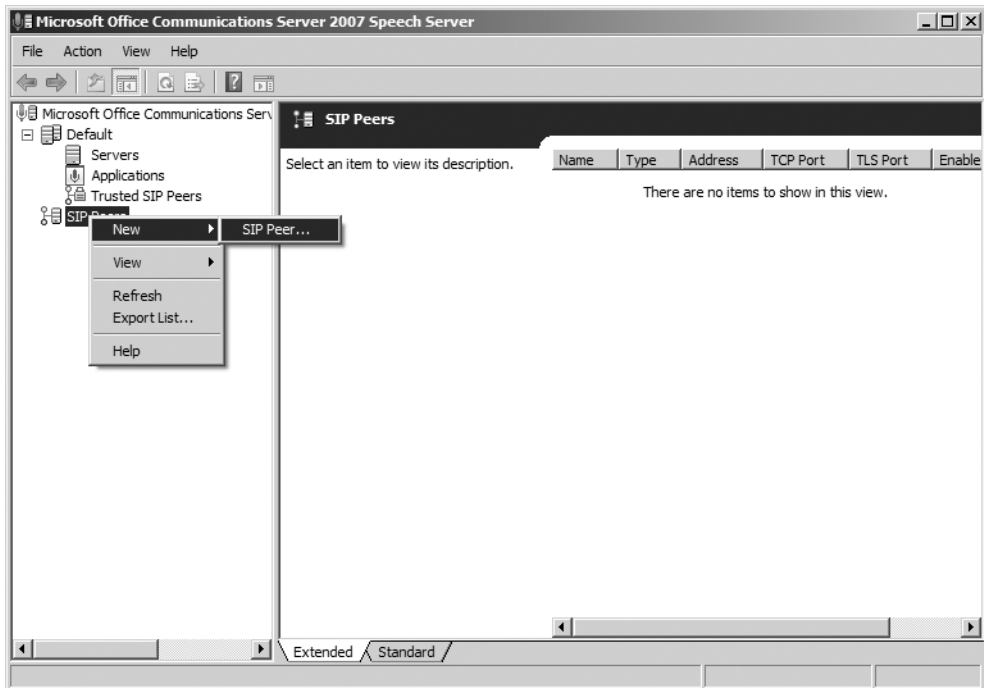
Figure 1-7. *Choosing to add a new SIP peer*
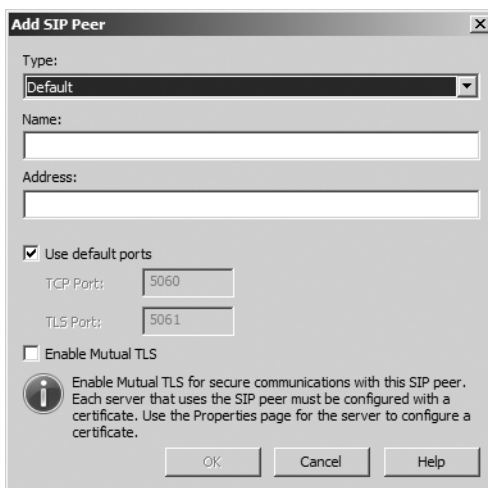
Figure 1-8. *The Add SIP Peer dialog box*

The Add SIP dialog box has three fields:

- *Type*: The drop-down list has the choices Default and TIS. Default describes VoIP gateways, IP PBXs, and SIP clients. TIS describes your traditional telephony card and TIM.

- *Name*: Enter a name that is descriptive enough for you to recognize it from a list of other SIP peers.

- *Address*: Enter either the host name or IP address of the SIP peer.

You also have the option of changing the default ports and enabling Mutual Transport Layer Security (TLS) for secure communications.

After you click OK, the SIP peer will be added to the details view (the right pane) of the SIP Peers node.

Adding a Trusted SIP Peer

If your application will be making outbound calls or transfers through this SIP peer, you will need to identify it as a trusted SIP peer. The SIP peer must be added, as described in the previous section, before it can be set up as a trusted SIP peer.

To add a trusted SIP peer, right-click the Trusted SIP Peers node in the Administrator Console and choose New ➤ Trusted SIP Peer. You will be prompted to choose which SIP peer to trust. You can enable and disable outbound and transfer calls after selecting the SIP peer, as shown in Figure 1-9.

■**Note** You can change only the Allow Transfers and Allow Outbound Calls options in the Add Trusted SIP Peer dialog box. Any other properties must be changed from the SIP Peers node, rather than from the Trusted SIP Peers node.



Figure 1-9. *Adding a trusted SIP peer*

You can optionally enable Secure RTP (SRTP) for standard SIP peers. SRTP is not available for TIS type peers. After you click OK, the SIP peer will be listed in the details view of the Trusted SIP Peers node.

Installing Applications in a Production Environment

If you are migrating applications from Microsoft Speech Server 2004, you will need to install your applications on the Speech Server Administrator Console before they will run in a production environment. And, of course, once you have developed and deployed your application, you will need to install it in your production environment via the Administrator Console. Development and deployment are discussed in the following chapters. This section covers only the installation of your application on the server.

Adding a New Application

To add a new application, open the Administrator Console (select Start ➤ Programs ➤ Microsoft Office Communications Server 2007 Speech Server), right-click the Applications node, and choose New Application. You will be prompted to enter information about your application, as shown in Figure 1-10.
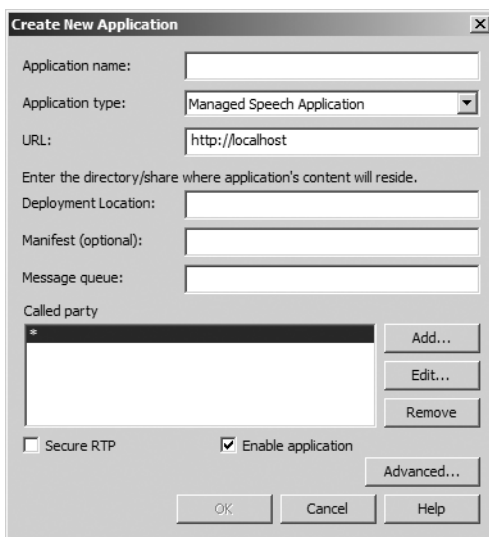
Figure 1-10. *Creating a new application*

The Create New Application dialog box has the following fields:

- *Application Name*: Enter a name that is descriptive enough to be picked out of a list, so that you can find your application easily.

- *Application Type*: The drop-down list contains three choices: Voice Response Workflow Application, SALT, and VoiceXML. This represents your application's programming model.

■**Note** If your application is a Voice Response Workflow type, the application must be deployed to the IIS instance on which Speech Server is installed. SALT and VoiceXML applications may be deployed on remote machines.

- *URL*: Enter the URL where your application is hosted.

- *Deployment Location*: Enter the directory where the source of your application is located.

- *Manifest*: Optionally, enter the relative path to the manifest XML file for your application that contains the grammars and prompts to preload.

- *Message Queue*: For outbound calling applications, this should reference the queue in MSMQ that is to be used for outbound calling. The next section explains how to configure your application for outbound calling.

After you click OK, your application will be added to the Applications node in the Administrator Console.

Setting Call Answering Precedence

Speech Server allows you to set call answering precedence, which lets applications with the same phone number determine which application the user is trying to call.

When you select the Applications node in the Administrator Console, you will see the list of installed applications in the details view, along with Set Call Answering Precedence link. Click the link to open the Call Answering Precedence dialog box.

Figure 1-11 shows an example of setting call answering precedence. In the example, the Employee Application has precedence over the Customer Application, simply because it is at the top of the list. The Customer Application uses the wildcard, *, in its phone number. This means all calls to any number will go to the Customer Application. However, since the Employee Application has precedence over the Customer Application, it is evaluated first. Therefore, if the number called is 555-1212, it will go to the Employee Application and calls to any other number will be forwarded to the Customer Application.
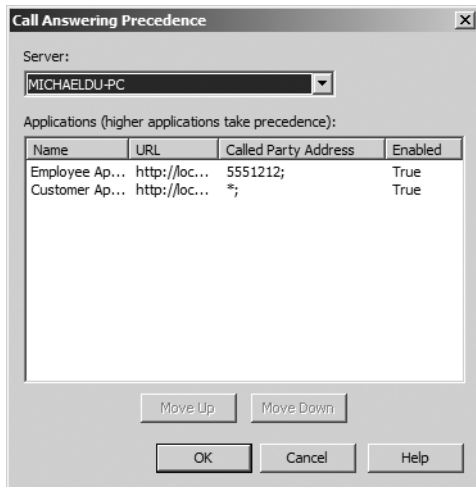
Figure 1-11. *Setting call answering precedence*

Setting Up Outbound Calling for an Application

To allow your application to place outbound calls, you need to first enable outbound calling on the Speech Server application by entering the name of the queue in the Message Queue field of the Create New Application dialog box, as described in the previous section. Then you need to do the following:

- Create a message queue in MSMQ.

- Change security to allow Speech Server access to the queue.

Creating a New Message Queue

To create the message queue, from the Windows Control Panel, open Administrative Tools, and then Computer Management. Under the Services and Applications node, you will find the Message Queuing node. Expand that node, as shown in Figure 1-12.

Right-click Private Queues and click New ➤ Private Queue. You will be prompted for the name of the new queue, as shown in Figure 1-13. Enter the same name as you supplied in the Message Queue field of the Create New Application dialog box, as described in the previous section. (Speech Server does not currently support transactional queuing, so leave the Transactional option unchecked.)
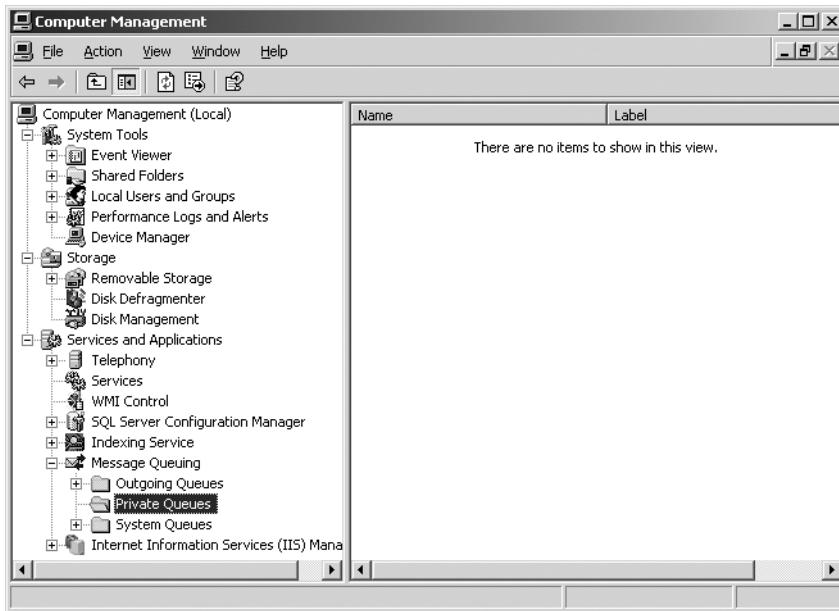
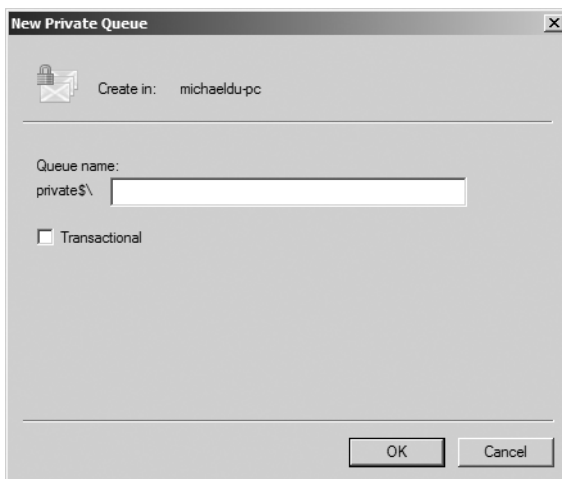**Figure 1-12.** *Navigating to Message Queuing*


Figure 1-13. *Creating a new queue*

■**Note** To avoid any problems, double-check to make sure the new queue has the same name as the queue for the application it will be serving. This is the name you entered in the Message Queue field when you set up your new application (see Figure 1-10).

Setting Up Queue Security

After creating the queue, you need to set up security for the queue so that Speech Server can access it. Right-click the newly created queue and choose Properties. In the Properties dialog box, choose the Security tab, as shown in Figure 1-14.
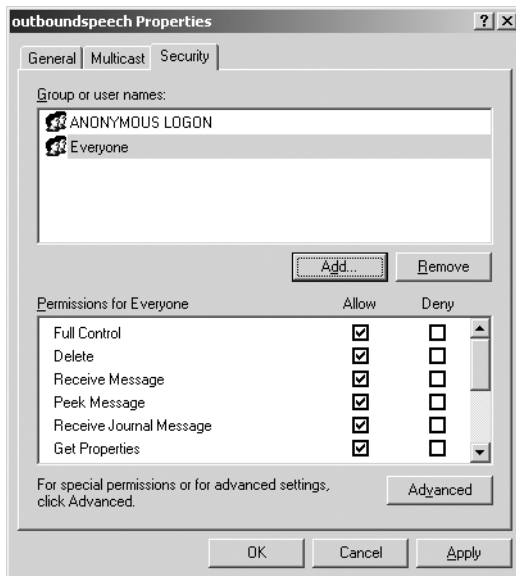
Figure 1-14. *Setting security for the message queue*

Click the Add button, and type **Network Service** in the Names to Select field. After you click OK, Network Service will be listed in the Groups or User Names list. Select Network Service and check the Allow boxes for Peek Message and Receive Message in the Permissions list. Clicking OK completes your queue setup.

Conclusion

This chapter provided an overview of OCS 2007 Speech Server, including the new features available in this version and a summary of how it works. Then you learned the basics of installing and configuring Speech Server. Installation is very important to get correct the first time, as it can be time-consuming to go back and fix a Speech Server installation.

The following chapters will walk you through each step of creating an IVR application. One of the first steps in your IVR application development should be building grammar, and that is the topic of the next chapter.