

Teaching Tip

An On Demand Text-to-Speech Agent

George H Bodnar
School of Business
Duquesne University
Pittsburgh PA 15282 USA
bodnar@duq.edu

ABSTRACT

This Teaching Tip provides a Microsoft .NET 2.0 framework application that uses Microsoft Agent® for on demand text-to-speech (TTS) capability. One opens the program, then pastes selected text into a textbox to have Microsoft Agent® Merlin read, i.e., speak the text. One may also open text files directly into the application. One can use the application in a multimedia classroom environment to hold and enhance student interest with a variety of voices and also to provide some technological assistance when it is desired to have text read-aloud.

Keywords: Multimedia Classroom, Microsoft Agent®, .NET Framework, Text-to Speech Output

1. INTRODUCTION

Contemporary multimedia classrooms have sound capability. This permits the use of text-to speech (TTS) output technologies. A primary objective is to hold and enhance student interest with a variety of voices – the human lecturer/narrator and one or more TTS voices. This paper provides a Microsoft .NET framework 2.0 application that uses Microsoft Agent® for on demand TTS capability.

1.1 Microsoft Agent®

Microsoft Agent® is a set of programmable software services that support the presentation of interactive animated characters within the Microsoft Windows® interface. Microsoft Agent® enables software developers to incorporate a new form of user interaction, known as conversational interfaces, that leverages natural aspects of human social communication. Microsoft provides speech output (text-to-speech or TTS) engines that one can use with Microsoft Agent®.

1.2 Combining Human Speech and TTS

A number of applications and studies have addressed this topic. Ng S. T. Chong, Panrit Tosukhowong, and Masao Sakauchi discuss how the use of TTS (text-to-speech) synthesis is becoming increasingly attractive as a supplement or even as an alternative to human narration, and they present the results of a study that examines the combined use of TTS and human narration on Web lecture presentations (Chong, et al., 2002). They conclude that their study validates their intuitions about the benefits of TTS

technology in Web lectures. TTS demonstrates merits where the entire lecture is narrated using TTS or where the lecture contains both speech synthesis and human voice. While their project concerns Web based presentations, both it and the enclosed application possess the basic feature of augmenting human narration with TTS to capture audience attention and to hold their interest.

In a paper titled “Use of Interactive Multimedia to improve the lecturing experience”, Clive Chandler presents a review of the large body of work that addresses the topic of using multimedia in the field of education (Chandler, 2003). The traditional lecture is a basic tool in transferring information. The paper discusses the development of an agent assisted lecture process and an analysis of its effects. Chandler states “On the whole there would appear to be much mileage in the use of these characters in an interactive way to enhance the lecturing process.” A potential drawback is that the lecture may become too much of a performance.

1.3 The TTS Application AgentTalk

The application, titled AgentTalk, activates Merlin, a Microsoft Agent®, and provides basic TTS capability by having Merlin read selected text. The operation of the application is discussed in a subsequent section.

The design goal of the application is to provide basic TTS capability in direct fashion. The .NET Framework makes it possible to run an application such as AgentTalk directly from a personal storage device on demand. There is no need to have the program previously authorized or installed by an administrative function on the target computer. The .NET Framework handles the details of

locating and loading the components an application needs. This is known as xcopy deployment. One simply copies files to a target machine and applications can be deployed without registry entries or dependencies.

Microsoft Agent® is independent of the .NET framework and is by itself just a software technology rather than a complete, integral software application that automatically does something. Microsoft Agent® services are not included in nor installed by AgentTalk, but are only activated by AgentTalk. Thus, existence of these services on the target computer is a prerequisite.

2. PREREQUISITES

AgentTalk is a .NET 2.0 application that uses Microsoft Agent services. There are three categories of prerequisites for using AgentTalk. Windows Vista®, however, includes all but one of these prerequisites, as discussed below.

2.1 .NET Framework 2.0

Windows Vista® includes the .NET framework 2.0 making xcopy deployment possible and eliminating any need to formally install AgentTalk. Windows XP® or earlier may have the framework installed. To check for the existence of the .Net framework on the target computer, open the Control panel and select Add or Remove programs. Check the list of installed programs for Microsoft .NET Framework 2.0. If not installed, the framework may be downloaded at no cost from <http://www.microsoft.com/downloads/details.aspx?FamilyID=0856EACB-4362-4B0D-8EDD-AAB15C5E04F5&displaylang=en>.

2.2 Microsoft Agent Services

Windows Vista® includes Microsoft Agent® Services and the Agent Merlin. Users of Windows XP® or earlier versions of Windows may install these services by going to the web page <http://www.microsoft.com/msagent/downloads/user.aspx>. The following four items should be downloaded and installed:

- the core components of Microsoft Agent;
- the Merlin Microsoft Agent character file;
- the Lernout & Hauspie TruVoice TTS engine - American English speech engine;
- the Microsoft SAPI 4.0a runtime binaries.

These four items appear as individual downloads in this order as one moves down the page.

2.3 Lernout & Hauspie TruVoice TTS engine - American English

Windows Vista® includes all necessary prerequisites for the operation of this application. However, the default speech engine is not likely to be acceptable as it will most likely be Microsoft Anna – a female voice – while Merlin the Wizard is a man. For best performance, one should download and install the Lernout & Hauspie TruVoice TTS engine - American English speech engine, which is available at no cost from <http://www.microsoft.com/msagent/downloads/user.aspx>. This is the first selection in the drop-down box that allows selection of a text-to-speech engine. It is noted that these text-to-speech engines are licensed only for use in

Microsoft Agent® enabled applications and Web pages with a visibly displayed Microsoft Agent® character.

It should not be necessary to install the SAPI 4.0 runtime support binaries that are offered immediately below the download of the TTS engines, although no harm will be done if one does so. This resource sets the Agent characters' TTSMODEID property to use the TTS3000 engines. However, this property is set directly in the AgentTalk application as installing the SAPI 4.0 runtime support did not override the use of the Microsoft Anna voice in Windows Vista® on my test machine.

3. USING THE APPLICATION AGENTTALK

This application strives to provide basic TTS capability with a minimum of effort. The user can simply download the application, unzip the files, and copy them to the target machine or a memory device and open (run) the application directly without installation.

3.1 Download of AgentTalk

I am the author of AgentTalk and have the authority to distribute it. Download the following ZIP file: <http://mysite.verizon.net/vzeu0s9j/AgentTalk.zip>. When unzipped (extracted), the download contains five files.

- AgentTalk (.exe)
- AgentTalk.vshost (.exe)
- AgentTalk (.xml)
- AxInterop.AgentObjects.dll
- Interop.AgentObjects.dll

AgentTalk.exe – the application file – is the program that one opens. The other four files are resources that are required by AgentTalk (.exe). They should be kept in the same directory as AgentTalk (.exe). Opening the file in Windows Vista® will prompt an “Open File -Security Warning”, stating that the publisher could not be verified as the file does not have a valid digital signature. Click run to open the application.

3.2 Using AgentTalk

Open the program AgentTalk. A textbox appears in the upper-left corner of the screen, and the Microsoft Agent® Merlin the Wizard appears and stands near the lower right corner of the textbox. Basic TTS capability is provided by copying the desired text into the Windows clipboard (i.e., a straightforward copy), then pasting it into the AgentTalk textbox, and then clicking Read on the AgentTalk menu. Merlin will read the text. The textbox will automatically clear for the next usage.

The textbox can be moved and resized as desired. It is not necessary to see the entire text in the textbox on the screen, nor is there any practical limitation to the amount of text that one can paste. Text from any source (Word® document, HTML, etc.) can be pasted as the textbox attempts to convert anything pasted to simple text. The application can be used to enhance a PowerPoint® presentation by having Merlin read selected points. It can be used to read announcements, portions of a web page, problems, assignments and other such material.

Merlin can be moved as desired and may fall asleep if not used for awhile (his snoring usually amuses a class). You

can hide Merlin by right-clicking on him, then selecting Hide from the short menu that appears. This does not close the application. When you click Read, Merlin will reappear. Hide Merlin if his snoring (or yawning) becomes bothersome, or if you wish to terminate the reading of text. The short menu also displays Advanced Character Options which exposes an Agent Properties menu. You can use this to change Merlin's speaking speed if so desired. Move the selector and click Apply. Note that clicking Apply will turn on the Agent's balloon feature if this box is selected by a checkmark in the display just above the speaking speed indicator. The balloon feature displays spoken text in a cartoon-like balloon. This feature is turned-off by AgentTalk when it starts, but the user has the option of turning it on in the Agent properties page.

3.3 Using Text Files

The application can also open and read text files directly into the textbox. The user might prepare a text file for narration in class, then open it and click Read on the application menu to have TTS enabled. The Microsoft Agent® services support modifying speech output through special tags inserted in the speech text string. These tags allow one to change the characteristics of the output expression of the character. Documentation is at <http://msdn2.microsoft.com/en-us/library/ms699259.aspx>. The most useful tag is used to insert a pause in the speech. The syntax is \Pau=number which pauses speech for the specified number of milliseconds. For example, the text string "Hello /Pau=2000/ I am Merlin the Wizard" contains a two second pause between Hello and I. (Please note that the text file does not have to have the text enclosed in quotes ("") as I have done here in this example.) Emphasis is possible with the Emp tag with syntax \Emp. This tag emphasizes the next word spoken. This tag must immediately precede the word. The other tags are less likely to be useful and are somewhat complicated. Please note that it is not necessary to edit a text file with tags. The Agent will read a text file using its rules of grammar. Interested users may consult the documentation for a thorough description of using tags to edit speech.

4. SUMMARY

Contemporary multimedia classrooms have sound capability. This permits the use of text-to speech (TTS) output

technologies. A primary objective is to hold and enhance student interest with a variety of voices – the human lecturer/narrator and one or more TTS voices. This paper provides a Microsoft .NET framework application that uses Microsoft Agent® for on demand text-to speech (TTS) capability. Windows Vista® includes all prerequisites excepting the best speech engine, which may easily be downloaded at no cost from Microsoft.

5. REFERENCES

- Chong, Ng S. T., Tosukhowong, Panrit, and SakauchiRex, Masao (2002), " WhiteboardVCR: a Web Lecture Production Tool for Combining Human Narration and Text-to-Speech Synthesis." Educational Technology & Society Vol. 5, No. 4. http://www.ifets.info/journals/5_4/ng.html
- Chandler, Clive (2003), "Use of Interactive Multimedia to Improve the Lecturing experience." <http://www.interdisciplinary.net/ci/mm/mm1/chandler%20paper.pdf>
- Dumdum, U. Rex and Tastle, William I. (1998), "Towards a Broader Competency-Based IS Education: A proposed Improvement Package for Analysis of Case Studies." Proceedings of ISECON'98, October 15-18, pp. 28-33.

AUTHOR BIOGRAPHY

George H. Bodnar is a Associate Professor in the School of Business Administration of Duquesne University, Pittsburgh, PA. He earned his Ph.D. in 1975 from the Wharton School, University of Pennsylvania with an Accounting major and a minor in Operations Research. He authored a textbook Accounting Information Systems in 1980. It has been in continuous publication since then, and is now in a 9th Edition with Prentice Hall. He has integrated computer technology in his teaching efforts for several decades, and has been programming in Microsoft Visual Basic for more than 10 years.





STATEMENT OF PEER REVIEW INTEGRITY

All papers published in the Journal of Information Systems Education have undergone rigorous peer review. This includes an initial editor screening and double-blind refereeing by three or more expert referees.

Copyright ©2008 by the Information Systems & Computing Academic Professionals, Inc. (ISCAP). Permission to make digital or hard copies of all or part of this journal for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial use. All copies must bear this notice and full citation. Permission from the Editor is required to post to servers, redistribute to lists, or utilize in a for-profit or commercial use. Permission requests should be sent to the Editor-in-Chief, Journal of Information Systems Education, editor@jise.org.

ISSN 1055-3096