



**CHANT**<sup>®</sup>

# **Fine-tune Speech Synthesis Using Text-to-Speech Markup**



Copyright © 2011 Chant Inc. All rights reserved.

Chant, SpeechKit, Getting the World Talking with Technology, talking man, and headset are trademarks or registered trademarks of Chant Inc. Other marks are trademarks or registered trademarks of their respective holders.

# Fine-tune Speech Synthesis Using Text-to-Speech Markup

**A**void putting your end users to sleep with boring synthesized speech. You will be amazed at how adjusting the speed, adding pauses, injecting emphasis, and switching voices can break up the monotony of synthesized speech.

Text-to-speech (TTS) markup is text with imbedded indicators that control speech synthesis from the text. Speaking qualities such as the speed, pitch, emphasis, and word pronunciation may be tailored in reproducing speech from text.

A TTS grammar is a collection TTS markup. A text-to-speech engine (i.e., synthesizer) uses TTS markup to enhance its ability to synthesize speech from text and generate the audio for playback.

## WHAT IS VOICE MARKUP?

Synthesizers (i.e., speech APIs) support unique markup syntax. By generating TTS markup at runtime, your application can maximize the quality of TTS playback and offer your end users the flexibility of using various synthesizers with your application.

## WHAT IS VOICEMARKUPKIT?

Chant VoiceMarkupKit is comprised of application ready software components that handle the complexities of text-to-speech (TTS) markup language management that enables you to generate TTS markup to enhance the playback quality when synthesizing.

The VoiceMarkupKit component library includes a voice markup management class that provides you a

simple way to generate text-to-speech markup. Your application can markup text as part of its runtime operation to enable real-time customization and tailoring of your text-to-speech environment.

It simplifies the process of generating Nuance L&H Native Control Sequence, Microsoft SAPI 4, Microsoft SAPI 5, and W3C SSML markup language to use with your favorite speech synthesizers.

VoiceMarkupKit includes ActiveX, C++, C++Builder, Delphi, Java, .NET Framework, Silverlight, and Web component library formats to support all your programming languages and provides sample projects for popular IDEs—such as the latest Visual Studio 2010 from Microsoft.

The component libraries can be integrated with 32-bit, 64-bit, and mobile applications.

## **VOICEMARKUPKIT FEATURES**

The goal of text-to-speech markup is to enhance the quality of the text-to-speech playback. With Chant VoiceMarkupKit, you can:

- Generate markup language in Nuance L&H Native Control Sequence, Microsoft SAPI 4 control tags, Microsoft SAPI 5 XML, and W3C SSML (Cepstral) syntax.
- Generate pronunciation phonemes for Cepstral, Microsoft SAPI 4, Microsoft SAPI 5, Nuance Real-Speak Solo, and Nuance Vocalizer synthesizers.
- Dynamically switch among speech APIs and syntax formats.

Chant VoiceMarkupKit is comprised of application ready software components that handle the complexities of generating text-to-speech markup for various markup syntax. This enables you to tailor speech synthesis to produce sounds in familiar dialects, speaking patterns, and accents of your end users. You can adjust TTS markup as needed for the synthesizer to enhance the playback quality when synthesizing.

Synthesizers (i.e., speech APIs) support unique markup syntax. VoiceMarkupKit supports the following synthesizers and their markup syntax:

| Synthesizer                                 | Speech API                  | Markup Syntax  |
|---|-----------------------------|--|
| Cepstral (all languages)                    | Cepstral Swift              | W3C SSML   |
| Microsoft SAPI 4 (all languages)            | SAPI 4                      | SAPI 4 Control Tags                                      |
| Microsoft SAPI 5 (all languages)            | SAPI 5                      | SAPI 5 XML Markup  |
| Nuance RealSpeak Solo (all languages)       | Nuance RealSpeak Solo       | L&H Native Control Sequence, SAPI 5 XML Markup           |
| Nuance Vocalizer Auotmotive (all languages) | Nuance Vocalizer Auotmotive | L&H Native Control Sequence, SAPI 5 XML Markup           |
| Nuance Vocalizer Mobile (all languages)     | Nuance Vocalizer Mobile     | L&H Native Control Sequence                              |
| Nuance Vocalizer Network (all languages)    | Nuance Vocalizer Network    | L&H Native Control Sequence, SAPI 5 XML Markup, W3C SSML |

By generating TTS markup at runtime, your application can maximize the quality of TTS playback and offer your end users the flexibility of using various synthesizers with your application.

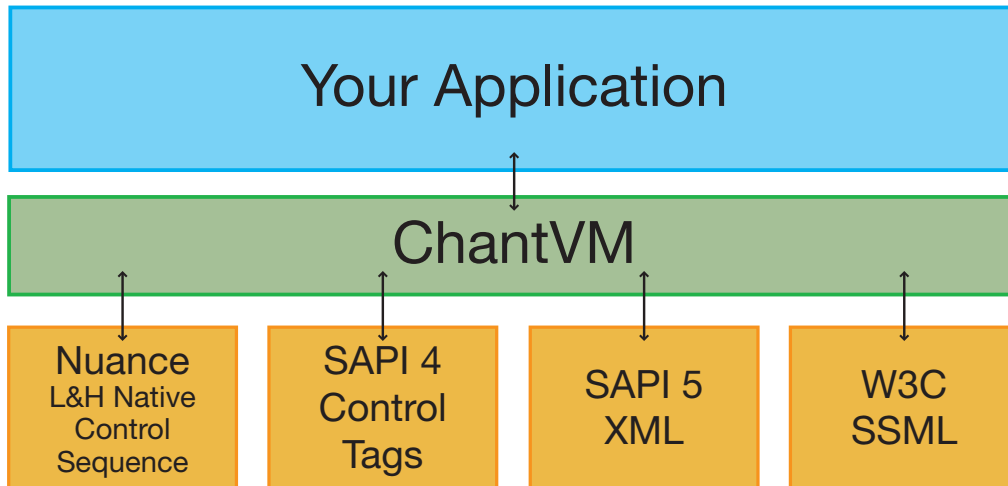
## VOICE MARKUP COMPONENT ARCHITECTURE

The VoiceMarkupKit component library includes a voice markup management class that provides you a simple way to generate text-to-speech markup. Your application can markup text as part of its runtime operation to enable real-time customization and tailoring of your text-to-speech environment.

The voice markup management class, ChantVM, enables you to specify a markup language, the markup options, and generate markup prior to synthesis. Your application uses the ChantVM class to manage the activities for generating the markup in the format you need. The ChantVM class supports the following markup syntax:

- Nuance L&H Native Control Sequence,
- Microsoft SAPI 4 control tags,
- Microsoft SAPI 5 XML markup, and
- W3C Speech Synthesis Markup Language (SSML).

The ChantVM class encapsulates all of the technologies necessary to make the process of generating markup simple and efficient for your application.



You instantiate a ChantVM class object before you want to generate markup within your application. You destroy the ChantVM class object and release its resources when you no longer want to generate markup within your application.

## MORE INFORMATION

To learn more about developing software that speaks and listens, explore how easily you can manage grammars, profiles, lexicons, recognizers, synthesizers, and text-to-speech markup directly within application software you develop in the following documents:

- Develop Software That Speaks and Listens,
- Integrate Speech Technology for Hands-free Operation,
- Design Grammars for High-Performance Speech Recognition,
- Administer Speaker Profiles for Accurate Speech Recognition, and
- Fine-tune Speech Synthesis Using Text-to-Speech Markup.



