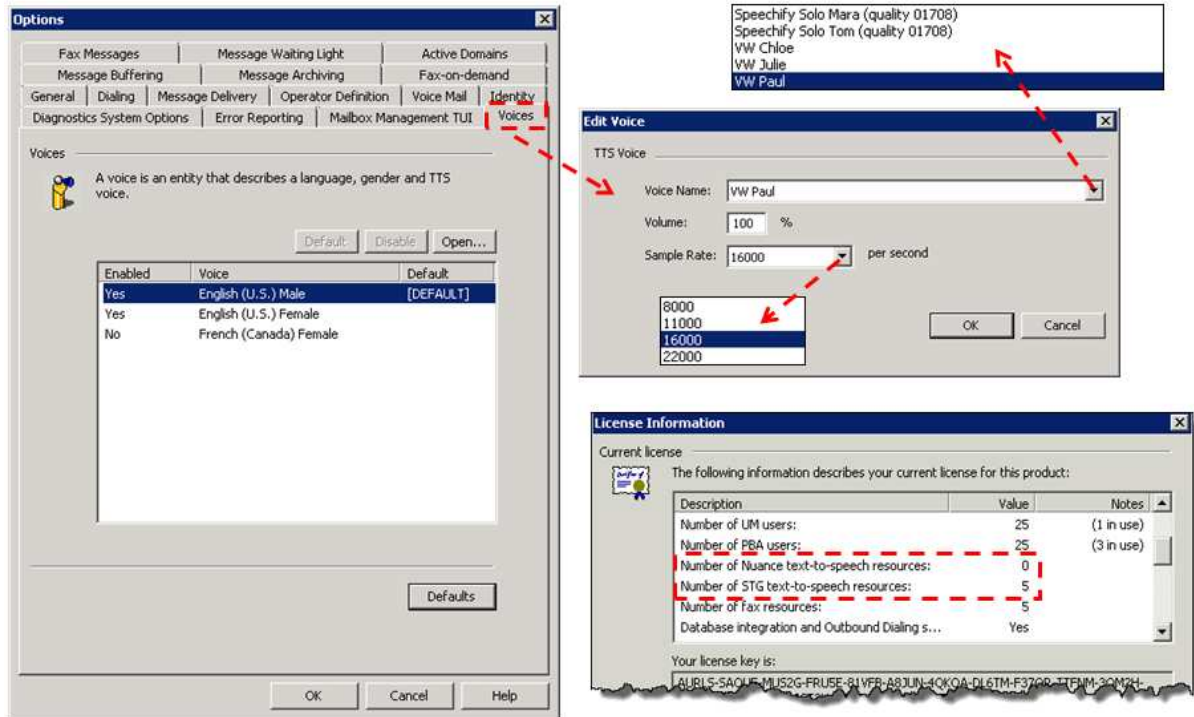




## Configuring the Text to Speech (TTS) Voice Properties



The default TTS voice properties can be changed for available voices using the **Edit Voice** menu in the **Tools > Options > Voices** tab. This menu allows administrators to configure the voice to use for TTS, as well as the TTS volume and the sample rate. The available TTS voices are dependent upon the TTS vendor that was licensed.

**NOTE:** Only one TTS vendor, either Nuance Communications or Speech Technology Group (STG), can be licensed.

The table below shows the TTS voices available from each vendor and the voices with which they are associated by default.

Voice	Default TTS Voice Name	
	Nuance Licensed	STG Licensed
English (U.S.) Male	Speechify Solo Tom (quality 01708)	VW Paul
English (U.S.) Female	Speechify Solo Mara (quality 01708)	VW Julie
French (Canada) Female		VW Chloe

### To configure the TTS voice properties

1. In the menu bar of the Administrator application, select **Tools > Options**.

2. Select the **Voices** tab.
3. In the list of available voices, select the voice for which you would like to configure the TTS properties. The voice will become highlighted.
4. Select the **Open** button. The **Edit Voice** menu will appear.
5. In the **Edit Voice** menu, use the **Voice Name** drop-down menu to select the desired TTS voice.
6. Enter the desired volume for the TTS voice in the **Volume** field. This value is a percentage of the maximum playback volume for the TTS voice. The default value is **80** percent.
7. Use the **Sample Rate** drop-down menu to select the sample rate in samples per second (Hz). Higher values correspond to higher TTS voice quality. The default value is **16000**.
8. Select **OK**.

## Text-To-Speech Element

The Text to Speech element enables you to convert text to speech. The supplied text is "spoken" to callers by the Text-to-Speech engine. Text can be spoken in either English (U.S.) Female or Male voice or French (CA) Female voice. The default is English (U.S.) Male voice. The voice gender can be changed in the UC Client administration mode by navigating to **Tools > Options > Voices** tab. You can use this element for a variety of purposes:

- If you prefer not to make audio recordings you can manually type the text to be spoken
- To speak out digits that were collected from a caller
- To speak out contextual information like the current time, date, caller Id, etc...
- To speak out information fetched from a database

The UC Server supports escape codes and VTML (VoiceText Markup Language) Tags to read out numbers/letters differently from the defaults depending on which TTS vendor is licensed.

Nuance Communications TTS escape codes:

<http://docs.voxeo.com/voicexml/n2.0/frame.jsp?page=appendixm.htm>

Speech Techonology Group (STG) TTS usable VTML Tags:

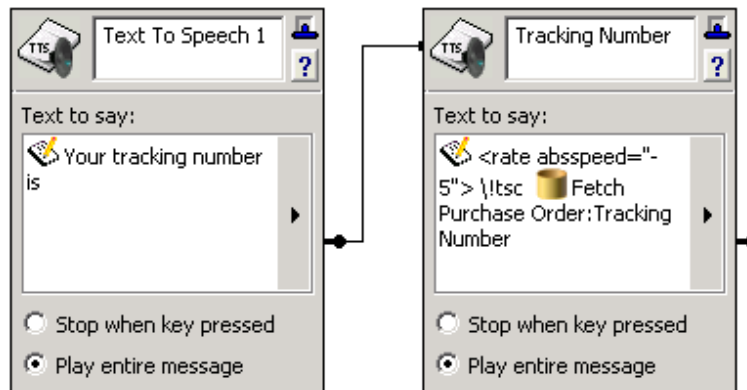
*Posted soon in Support Community at <http://supportforums.adtran.com>*

The Text To Speech element also supports embedded tags to fine-tune the way that the server speaks out the text-based information. Embedded tags are special codes that can be used in the element to customize text-to-speech (TTS) behavior in a variety of ways. Below is list of common embedded tags that can be used in the Text To Speech element using the **Nuance Communications** TTS vendor:

Nuance Markup	Explanation	Example
\!pN	Create a pause N milliseconds long.	\!p300 Create a pause for 300 ms

\!tsc	All-character spellout mode: pronounce all characters individually by name	My account number is \!tsc 487-B12. My account number is four eight seven dash bee one two.
\!tsa	Alphanumeric spellout mode: pronounce only alphanumeric characters by name.	My account number is \!tsa 487-B12. My account number is four eight seven bee one two.
\!tsr	Radio spellout mode: like alphanumeric mode, but alphabetic characters are spelled out according to the International Radio Alphabet. Supported in English only.	The last name is spelled \!tsr Dvorak \!ts0 and is pronounced Dvorak The last name is spelled delta victor oscar romeo alpha kilo and is pronounced Dvorak.
\!ts0	Default mode - resets any previous \!ts modications.	
\!ny0	Quantity interpretation	In May \!ny0 1945 people emigrated. In May one thousand nine hundred forty five people emigrated.
\!ny1	Year interpretation (default)	\!ny0 1945 \!ny1 people emigrated in 1945 One thousand nine hundred forty five people emigrated in nineteen forty five.

The Nuance TTS license also supports SAPI XML tags to control the rate of the voice. The **rate absspeed** command controls the absolute rate of the voice, so a value of ten always corresponds to the value of ten; the value of five always corresponds to a value of five. The SAPI XML tags must be entered at the beginning of the Text To Speech element. The command should be entered as follows: `<rate absspeed="value">` The value of the attribute should be an integer between -10 and 10. The example below is going to slow down the rate in which the tracking number is going to be read from a data source. The `\!tsc` tag is used in conjunction to spell out the numbers by name.



Below is list of common embedded tags that can be used in the Text To Speech element using the **Speech Technology Group (STG)** TTS vendor. Punctuation marks like “?” and “!” can also be used to add inflections when the text is being read. For example, “I am very mad!” will add inflections compared to “I am very mad.”

**NOTE:** *If you start a TTS Element with a VTML tag, you MUST put a character before it or it will not be recognized. For example always put a period “.” Before the VTML tag when starting it in an Element.*

STG Tag	Explanation	Example
<vtml_pause time="msec"/>	Sets a pause to be inserted in the synthesized voice.	The arrest warrant issued in Alabama<vtml_pause time="1000"/> links the attorney to a government undercover case.
<vtml_sayas interpret-as="ssml:characters"> Text </vtml_sayas>	All-character spellout mode: pronounce all characters individually by name	My account number is <vtml_sayas interpret-as="ssml:characters"> 487-B12 </vtml_sayas>. My account number is four eight seven dash bee one two.
<vtml_sayas interpret-as="sapi:number"> Text </vtml_sayas>	Quantity interpretation	In May <vtml_sayas interpret-as="sapi:number"> 1945 </vtml_sayas> people emigrated in 1945. In May one thousand nine hundred forty five people emigrated in nineteen forty five.
<vtml_speed value="speed"> Text </vtml_speed>	Sets the speed at which the text surrounded with these tags is read. Value of speed with 50-400(%)	Your have <vtml_speed value="75"> 123568 </vtml_speed> people enrolled in class.
<vtml_volume value="volume"> Text </vtml_volume>	Sets the volume at which the text surrounded with tags is read. Value of Volume from 0 to 500(%)	<vtml_volume value="150"> The arrest warrant issued in Alabama links the attorney to a government undercover case. </vtml_volume>
<vtml_pitch value="pitch"> Text </vtml_pitch>	Sets the pitch at which the text surrounded by this tags is read. Level of pitch ranging between 50-200(%)	. <vtml_pitch value="150"> The arrest warrant issued in Alabama </vtml_pitch> links the attorney to a government undercover case.
<vtml_partofsp part="unknown"   "noun"   "verb"   "modifier"   "function"   "interjection"> text </vtml_partofsp>	Designates a word class of the word surrounded with tags. (Only available in English).	Did you <vtml_partofsp part="verb">record</vtml_partofsp> that <vtml_partofsp part="noun">record</vtml_partofsp>?

When using the “vtml\_speed” tag, the variables don’t have to continually be opened/closed if you want to have a global speed change, but a specific sub-change as well.

If changing the speeds multiple times in a single TTS element, it can be tagged as the following:

```
<vtml_speed value="85"> You have <vtml_speed value="75"> 123568 </vtml_speed>
people enrolled in class. </vtml_speed>
```

**Important Notes:**

- Text To Speech is a licensed feature of UC server and is never included with any base bundles; if Text to Speech is a customer requirement it must always be explicitly added as an expansion license in the bill of materials.
- Text To Speech is licensed on a per-port basis and is allocated and freed on a dynamic basis. For example, if two callers are using the same service and the system has one Text To Speech license, as long as the callers are both not using a Text To Speech element at the same time they will both hear text spoken aloud when the call enters the element. However, if both callers are using a Text To Speech element at the same time then only the first caller to enter the Text To Speech element will hear text spoken aloud; the other will not hear any text