



PureConnect®

2021 R4

Generated:

04-November-2021

Content last updated:

11-June-2019

See [Change Log](#) for summary of changes.



# CIC Text to Speech Engines

## Technical Reference

### Abstract

This document describes the Text-to-Speech engines supported in CIC and provides installation and configuration information.

For the latest version of this document, see the PureConnect Documentation Library at: <http://help.genesys.com/pureconnect>.

For copyright and trademark information, see [https://help.genesys.com/pureconnect/desktop/copyright\\_and\\_trademark\\_information.htm](https://help.genesys.com/pureconnect/desktop/copyright_and_trademark_information.htm).

# Table of Contents

Table of Contents	2
Introduction to CIC Text to Speech Engines	3
Supported TTS engines	3
Supported languages	3
TTS SAPI Engines	4
Microsoft SAPI engine	4
Other SAPI engines	4
SAPI architecture	4
Configure the SAPI TTS voice on the CIC server	4
TTS MRCP Engines	6
Interaction Text to Speech	7
Benefits of Interaction Text to Speech	7
Supported Languages for Interaction Text to Speech	8
Licensing for Interaction Text to Speech	8
Interaction Designer Tools for Interaction Text to Speech	8
Interaction Text to Speech with SAPI or MRCP TTS as default	9
Partially Supported SSML Objects	11
Supported Say-as Text Normalization	11
User-defined Dictionaries	15
Configure the TTS engine in Interaction Administrator	17
Add Voices and Languages for SAPI	19
Change Log	21

## Introduction to CIC Text to Speech Engines

The PureConnect platform uses a Text-to-Speech (TTS) engine to read text to callers over the telephone. For example, a user can take advantage of this system to retrieve an email message over the phone. The TTS engine then employs a speech synthesizer to read the sender, subject, and body of the message.

Genesys offers Interaction Text to Speech as a native TTS engine for Customer Interaction Center. Incorporated into Interaction Media Server, Interaction Text to Speech does not require a separate installation or separate hardware.

Apart from Interaction Text to Speech, CIC supports various TTS engines that comply with Speech Application Programming Interface (SAPI) and Media Resource Control Protocol (MRCP). The quality of the speech produced by these TTS engines varies from vendor to vendor.

You can use TTS through CIC handlers that you can create or modify through Interaction Designer, VoiceXML, and through Interaction Attendant nodes.

### Supported TTS engines

For a complete list of the third-party TTS engines that CIC supports, see <http://testlab.genesys.com>. You can purchase and use Interaction Text-to-Speech, which is integrated in Interaction Media Server, for basic TTS functionality.

### Supported languages

For Interaction Text to Speech, see [Supported Languages for Interaction Text to Speech](#). To view the list of languages supported by a specific third-party TTS engine, see the third-party TTS engine vendor's website.

*[Copyright and trademark](#)*

# TTS SAPI Engines

## Microsoft SAPI engine

The Microsoft SAPI-compliant TTS engine is available with the Windows Server 2008 R2 and 2012 R2 operating systems, along with one or more TTS voices. CIC supports the SAPI version 5 standard. Microsoft also offers software for your SAPI solution. For more information about Microsoft Speech Server SDK, Microsoft Speech Platform Runtime, and adding voices for SAPI, see [Speech Platforms](http://msdn.microsoft.com/en-us/library/hh361571(v=office.14).aspx) at [http://msdn.microsoft.com/en-us/library/hh361571\(v=office.14\).aspx](http://msdn.microsoft.com/en-us/library/hh361571(v=office.14).aspx).

**Note:**

For version compatibility information on Microsoft SAPI software, see <http://testlab.genesys.com>.

## Other SAPI engines

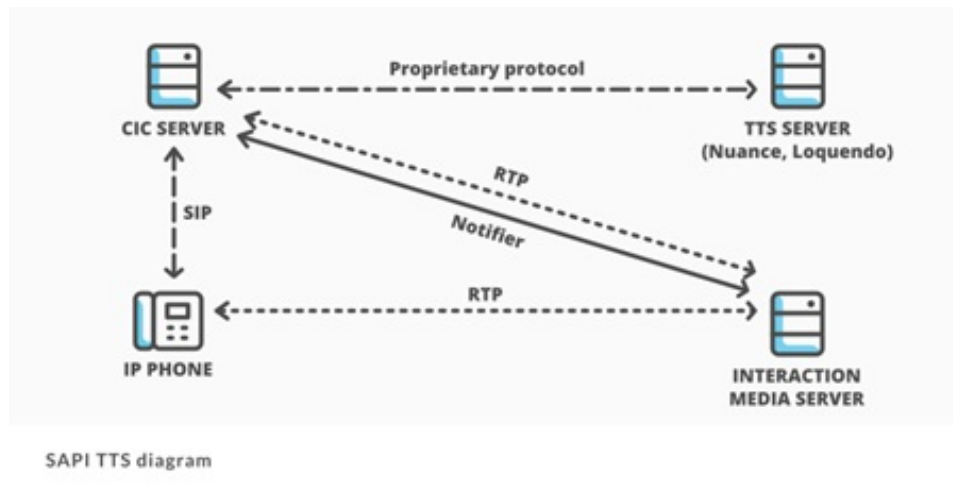
Any third-party TTS engine that supports these same standards should integrate with CIC. Nuance Vocalizer is the only SAPI TTS engine that you can purchase from Genesys. For TTS installation instructions for third-party products, see the vendor product installation documentation.

**Note:**

Requires a third-party TTS license key.

## SAPI architecture

The following diagram depicts the protocol flow between servers when using SAPI for TTS plays. All audio streams from the TTS server to the CIC server using the vendor's proprietary method. The CIC server streams the audio using Real-time Transport Protocol (RTP) to Interaction Media Server, which then streams that audio using RTP to the IP device. For more information, see the *Interaction Administrator Help* at [https://help.genesys.com/cic/mergedProjects/wh\\_ia/desktop/interaction\\_administrator\\_help.htm](https://help.genesys.com/cic/mergedProjects/wh_ia/desktop/interaction_administrator_help.htm).



## Configure the SAPI TTS voice on the CIC server

On Windows, SAPI uses a selected voice. As such, the CIC server uses this voice by default for all TTS operations, unless you configure other voices in Interaction Administrator.

### To configure the SAPI TTS voice on the CIC server

1. Log on to the Windows Server hosting CIC with the user account that the **Interaction Center** service runs under.

**Note:**

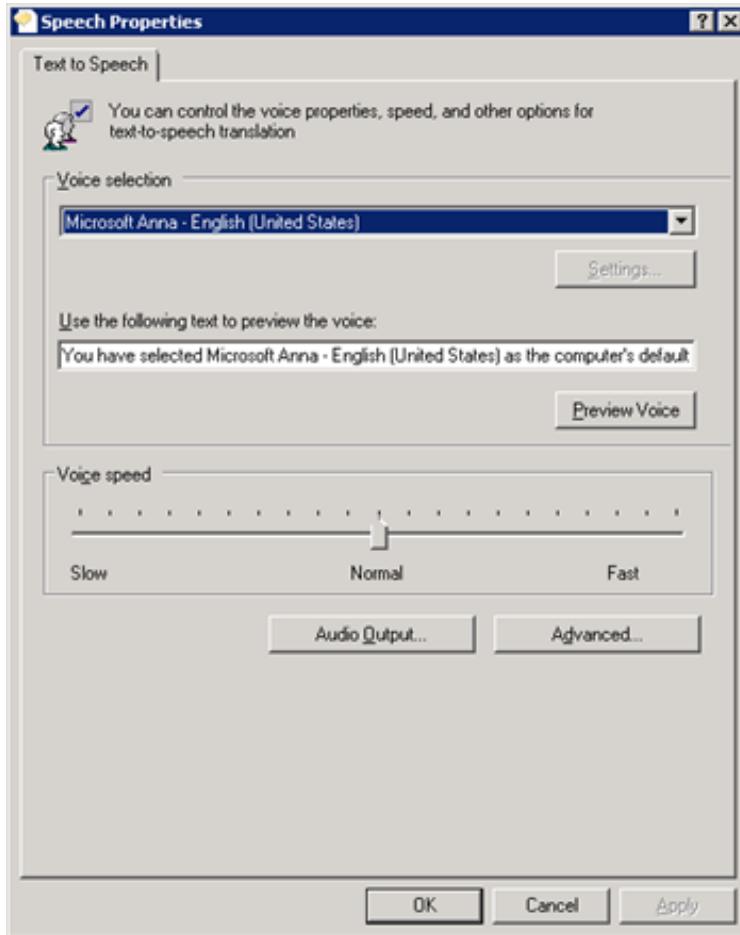
If you log on to the Windows Server with a different user account than the one under which the **Interaction Center** service runs, the selected TTS voice applies only to that account and doesn't affect the voice that the CIC server uses for TTS operations.

2. Run the speech applet, `sapi.cpl` from the following folder: `C:\Windows\SysWOW64\Speech\SpeechUX\`.

**Important!**

You must use the `sapi.cpl` program file in the specified directory path as it is the 32-bit version. Using the 64-bit version of `sapi.cpl` from other directory paths or from the **Speech** applet in the **Control Panel** does not configure SAPI TTS operations for the CIC server.

The **Speech Properties** dialog box appears.



3. In the **Voice selection** list box, click the voice you want to use as the default voice.

**Note:**

Some Windows Server versions offer only one voice, by default.

**Tips:**

- If you want to preview the selected voice, click **Preview Voice**.
- If you want to adjust the rate of speech for the voice playback, move the **Voice speed** slider to the right to increase the speed or to the left to decrease the speed.

4. Click **OK**. These changes take effect immediately on the CIC server for any SAPI TTS operations.

# TTS MRCP Engines

Media Resource Control Protocol (MRCP) enables speech servers to provide various speech services to clients. PureConnect supports the MRCP v2.0 protocol for connecting to speech servers that provide text-to-speech (speech synthesis) services. Third-party TTS engines that support MRCP v2.0 can integrate with PureConnect but Genesys only resells the Nuance TTS product line.

For more information about these engines, see the *MRCP Technical Reference* at [https://help.genesys.com/cic/mergedprojects/wh\\_tr/mergedProjects/wh\\_tr\\_mrmp/desktop/mrmp\\_technical\\_reference.htm](https://help.genesys.com/cic/mergedprojects/wh_tr/mergedProjects/wh_tr_mrmp/desktop/mrmp_technical_reference.htm). Also, see your vendor's product documentation.

PureConnect is compliant with the Media Resource Control Protocol Version 2 (MRCPv2), RFC 6787: <http://tools.ietf.org/html/rfc6787>.

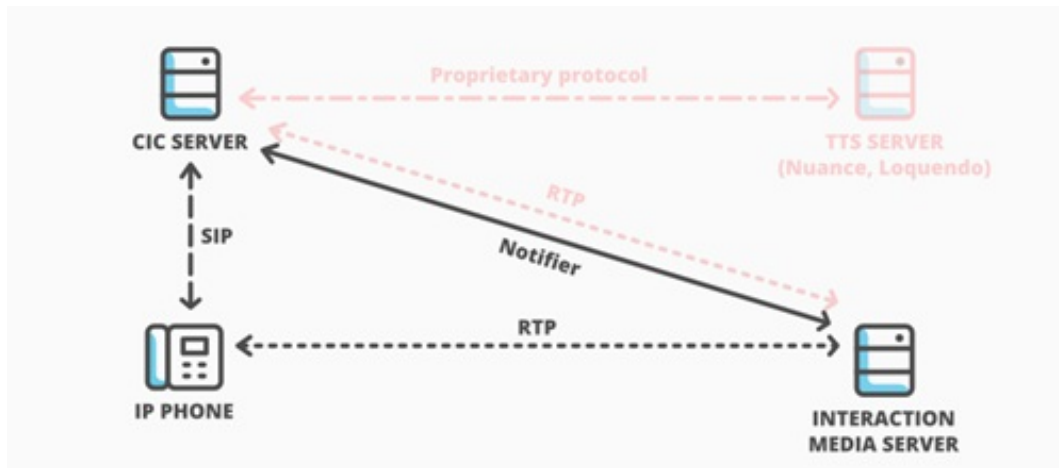
# Interaction Text to Speech

Interaction Text to Speech is a native TTS engine within Interaction Media Server. Interaction Text to Speech is in continuous development to comply with the following standards:

- Speech Synthesizer Markup Language (SSML) v1.1 at <https://www.w3.org/TR/speech-synthesis/>
- Pronunciation Lexicon Specification (PLS) Version 1.0 at <https://www.w3.org/TR/pronunciation-lexicon/>

## Benefits of Interaction Text to Speech

- Simpler deployment than a third-party TTS solution through SAPI or MRCP



Interaction Text to Speech diagram

- No extra hardware or software requirements for the CIC server or Interaction Media Server
- Simplification of selection rules – Interaction Text to Speech uses only *Media Server selection rules*. TTS solutions based on MRCP require both Media Server selection rules and MRCP selection rules in Interaction Administrator.

For more information about Media Server selection rules, see the following:

- *Interaction Media Server Technical Reference* at [https://help.genesys.com/cic/mergedProjects/wh\\_tr/desktop/pdfs/media\\_server\\_tr.pdf](https://help.genesys.com/cic/mergedProjects/wh_tr/desktop/pdfs/media_server_tr.pdf)
- *Interaction Administrator Help* at [https://help.genesys.com/cic/mergedProjects/wh\\_ia/desktop/interaction\\_administrator\\_help.htm](https://help.genesys.com/cic/mergedProjects/wh_ia/desktop/interaction_administrator_help.htm).
- Less audio and signaling traffic on the network than a third-party TTS solution

## Supported Languages for Interaction Text to Speech

Interaction Text-to-Speech supports the following languages:

- Dutch, Netherlands (nl-NL)
- English, United States (en-US)
- English, Australia (en-AU)
- English, Great Britain (en-GB)
- French, Canada (fr-CA)
- French, France (fr-FR)
- German, Germany (de-DE)
- Italian, Italy (it-IT)
- Japanese, Japan (ja-JP)
- Mandarin Chinese, China (zh-CN)
  
- Portuguese, Brazil (pt-BR)
- Spanish, Spain (es-ES)
- Spanish, United States (es-US)

### Note:

To ensure that Interaction Media Server does not exceed memory resources, customers should test the performance and memory usage of Media Servers when using more than 4 TTS languages. It's possible that older Media Servers cannot handle more than 4 languages. Overuse of Interaction Media Server resources can result in defects or failures in audio processing.

## Licensing for Interaction Text to Speech

Interaction Text to Speech requires a license for the feature, a license for the number of sessions to allow, and a license for each language that you want to support. The following table provides the license names for Interaction Text to Speech:

License	Name
Interaction Text to Speech (ITTS)	I3_FEATURE_MEDIA_SERVER_TTS
ITTS Sessions (total across all languages)	I3_SESSION_MEDIA_SERVER_TTS
ITTS Language Feature – Dutch (NL)	I3_FEATURE_MEDIA_SERVER_TTS_LANGUAGE_NL
ITTS Language Feature - English (US)	I3_FEATURE_MEDIA_SERVER_TTS_LANGUAGE_EN
ITTS Language Feature - English (AU)	I3_FEATURE_MEDIA_SERVER_TTS_LANGUAGE_EN_AU
ITTS Language Feature - English (GB)	I3_FEATURE_MEDIA_SERVER_TTS_LANGUAGE_EN_GB
ITTS Language Feature - French (CA)	I3_FEATURE_MEDIA_SERVER_TTS_LANGUAGE_FR_CA
ITTS Language Feature – French (FR)	I3_FEATURE_MEDIA_SERVER_TTS_LANGUAGE_FR
ITTS Language Feature - German (DE)	I3_FEATURE_MEDIA_SERVER_TTS_LANGUAGE_DE
ITTS Language Feature - Italian (IT)	I3_FEATURE_MEDIA_SERVER_TTS_LANGUAGE_IT
ITTS Language Feature – Japanese (JP)	I3_FEATURE_MEDIA_SERVER_TTS_LANGUAGE_JA
ITTS Language Feature – Mandarin (CN)	I3_FEATURE_MEDIA_SERVER_TTS_LANGUAGE_ZH_CN
ITTS Language Feature – Portuguese (BR)	I3_FEATURE_MEDIA_SERVER_TTS_LANGUAGE_PT_BR
ITTS Language Feature - Spanish (ES)	I3_FEATURE_MEDIA_SERVER_TTS_LANGUAGE_ES_ES
ITTS Language Feature - Spanish (US)	I3_FEATURE_MEDIA_SERVER_TTS_LANGUAGE_ES_US

## Interaction Designer Tools for Interaction Text to Speech

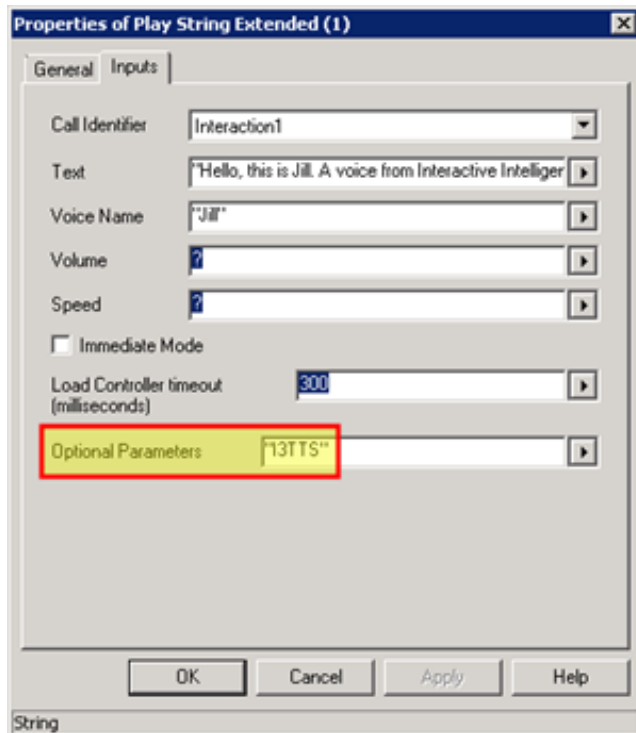


You can use the following Interaction Designer tools with Interaction Text to Speech when Interaction Media Server is set as the default Text-to-Speech provider:

Play tools	Record tools	Play prompt tools
<ul style="list-style-type: none"> <li>• Play String</li> <li>• Play String Extended</li> <li>• Play Text File</li> <li>• Play Text File Extended</li> </ul>	<ul style="list-style-type: none"> <li>• Record String</li> <li>• Record String Extended</li> <li>• Record Text File</li> <li>• Record Text File Extended</li> </ul>	<ul style="list-style-type: none"> <li>• Play Prompt Phrase</li> </ul>

## Interaction Text to Speech with SAPI or MRCP TTS as default

If you use SAPI or MRCP, you can still configure these tools to use Interaction Text to Speech by specifying an optional parameter, "I3TTS", in the properties for that tool step:



You can also specify other parameters to control various characteristics of the synthesized speech of Interaction Text to Speech.

**Note:**

To use multiple parameters, use a space between each parameter in the **Optional Parameters** box. Separate parameters from values with a colon (:). You must use double quotation marks around the entire string of characters. For example, "I3TTS i3tts.content.language:text/plain".

Parameter	Values	Description
i3TTS	N/A	If Interaction Text to Speech is not selected as the default TTS provider, use this parameter to set Interaction Text to Speech as the TTS provider.  <b>Note:</b> Specify this parameter before any other parameters.
i3tts.content.language	Language values as specified in <a href="#">RFC 3066</a>	Use this parameter to specify the language for ITTS to use to synthesize the text.
i3tts.content.type	<ul style="list-style-type: none"> <li>text/plain (default)</li> <li>application/ssml+xml</li> </ul>	Use this parameter to specify the content type of the text for ITTS to synthesize.
i3tts.voice.name	ITTS voice to use for the specified language.	Use this parameter to specify the i3tts voice for ITTS to use to synthesize the text. ITTS uses the following voices for the supported languages: <ul style="list-style-type: none"> <li>Dutch (NL) - Marina</li> <li>English (US) - Jill</li> <li>English (AU) - Kandyce</li> <li>English (GB) - Ellene</li> <li>French (CA) - Hilorie</li> <li>French (FR) - Manon</li> <li>German (DE) - Arabella</li> <li>Italian (IT) - Luisa</li> <li>Japanese (JP) - Miki</li> <li>Mandarin (CN) - Mei-Ling</li> <li>Portuguese (BR) - Viviane</li> <li>Spanish (US) - Isabel</li> </ul>
i3tts.voice.rate	<ul style="list-style-type: none"> <li>A non-negative percentage</li> <li>default</li> <li>x-fast</li> <li>fast</li> <li>medium</li> <li>slow</li> <li>x-slow</li> </ul>	Use this parameter to specify the SSML prosody rate of the selected ITTS voice.
i3tts.voice.volume	<ul style="list-style-type: none"> <li>A positive or negative value in decibels (dB)</li> <li>default</li> <li>x-loud</li> <li>loud</li> <li>medium</li> <li>soft</li> <li>x-soft</li> <li>silent</li> </ul>	Use this parameter to specify the SSML prosody volume of the selected ITTS voice.
i3tts.voice.pitch	<ul style="list-style-type: none"> <li>A value in hertz (Hz)</li> <li>default</li> <li>x-high</li> <li>high</li> <li>medium</li> <li>low</li> <li>x-low</li> </ul>	Use this parameter to specify the SSML prosody pitch of the selected ITTS voice.

## Partially Supported SSML Objects

The following table lists the SSML objects that Interaction Text to Speech partially supports.

Element	Attribute	Notes
token	N/A	ITTS does not support the token element.
emphasis	level	ITTS does not support a value of none for the level attribute.
phoneme	alphabet	ITTS supports only the ipa alphabet and PureConnect's own Arpabet-style phoneme alphabet.
prosody	duration range contour pitch semitones (st)	ITTS supports the following attributes for the prosody element: <ul style="list-style-type: none"> <li>pitch</li> <li>rate</li> <li>volume</li> </ul>
sub	alias	ITTS does not support the sub element.
lookup	N/A	ITTS does not support the lookup element.

## Supported Say-as Text Normalization

In verbal conversations, certain categories of speech, such as currency and time, use a specific method to convey information. For example, when people read \$12,345, it is usually spoken as "twelve-thousand-three-hundred-forty-five dollars" as opposed to "dollar-symbol one-two (pause) three-four-five", which is how a computer might interpret it.

In TTS, *say-as* text normalization directs the speech synthesizer to speak text in a specific manner so that the listener understands it. Without the *say-as* functionality, the synthesizer might speak a time of 10:30 AM as "one-zero-three-zero am". With *say-as*, the synthesizer can say the time as "ten-thirty a-m", which the listener understands easier.

The following table lists the say-as normalization types that are available and their support within Interaction Text to Speech.

Text normalization type	Usage	Supported	Notes
address	Processes mailing addresses	Yes	<p>Processes US addresses, including military addresses, Post Office boxes, and rural routes. Abbreviations are expanded based on context:</p> <p><b>Abbreviation examples:</b></p> <p>123 Main St. "St." spoken as "street"</p> <p>Springfield, IN "IN" spoken as "Indiana"</p> <p>PO Box "PO" spoken as "post office"</p> <p>RR 2 "RR" spoken as "rural route"</p> <p><b>Tips:</b></p> <ul style="list-style-type: none"> <li>Do not include the name of the addressee inside the <code>&lt;say-asinterpret-as="address"&gt;</code> element. Instead, use regular text processing for all names.</li> <li>Use official United States Postal Service abbreviations for states, territories, and thoroughfare types.</li> <li>If you encounter the incorrect audio of "street" when "saint" is required, use the full word "Saint" instead of the "St." abbreviation, as seen in "St. Paul, MN".</li> <li>The <i>say-as</i> element for an address may contain <i>newline</i> characters</li> </ul> <p>Supported languages:</p> <ul style="list-style-type: none"> <li>en-US</li> </ul>

alphanumeric	Spells letters and numbers	Yes	Supported languages:	
			<ul style="list-style-type: none"> <li>• de-DE</li> <li>• en-AU</li> <li>• en-GB</li> <li>• en-US</li> <li>• es-US</li> <li>• fr-CA</li> </ul>	<ul style="list-style-type: none"> <li>• fr-FR</li> <li>• it-IT</li> <li>• ja-JP</li> <li>• nl-NL</li> <li>• pt-BR</li> </ul>
boolean	Understands yes and no	Yes	This type uses the VoiceXML 2.0 defined type for Boolean. ITTS supports usage of T, True, F, and False, and local variants for this normalization type.	
			Supported languages:	
currency	Processes currency amounts	Yes	In specifying currency values, you can use either the monetary symbol, such as \$, or the associated abbreviation defined by <a href="#">ISO-4217</a> , such as (USD).	
			ITTS processes values to the right of a decimal point for only four digits. ITTS ignores any additional digits.	
			Supported languages:	
			<ul style="list-style-type: none"> <li>• de-DE (€ or EUR)</li> <li>• en-AU (\$ or AUD)</li> <li>• en-GB (£ or GBP)</li> <li>• en-US (\$ or USD)</li> <li>• es-ES (€ or EUR)</li> <li>• es-US (\$ or USD)</li> <li>• fr-CA (\$ or CAD)</li> </ul> <p>You can use a comma or a period as the decimal mark for fr-CA.</p>	<ul style="list-style-type: none"> <li>• fr-FR (€ or EUR)</li> <li>• it-IT (€ or EUR)</li> <li>• ja-JP (¥ or JPY)</li> <li>• nl-NL (€ or EUR)</li> <li>• pt-BR (R\$ or BRL)</li> <li>• zh-CN (¥, CNY, or RMB)</li> </ul>

date	Processes chronological dates	Yes	<p>ITTS supports the following date formats:</p> <ul style="list-style-type: none"> <li>• mdy - specify as <i>mm/dd/yyyy</i> in the text</li> <li>• dmy - specify as <i>dd/mm/yyyy</i> in the text</li> <li>• ymd - specify as <i>yyyy/mm/dd</i> in the text</li> <li>• md - specify as <i>mm/dd</i> in the text</li> <li>• dm - specify as <i>dd/mm</i> in the text</li> <li>• ym - specify as <i>yyyy/mm</i> in the text</li> <li>• my - specify as <i>mm/yyyy</i> in the text</li> <li>• y - specify as <i>yyyy</i> in the text</li> <li>• m - specify as <i>mm</i> in the text</li> <li>• d - specify as <i>dd</i> in the text</li> </ul> <p>Example (en-US):</p> <pre>&lt;say-as interpret-as="date" format="mdy"&gt;01/01/1984&lt;/say-as&gt;</pre> <p>Output (en-US):</p> <p>"January first, nineteen eighty-four"</p> <div style="border: 1px solid black; padding: 5px;"> <p><b>Tip:</b>  You can use the following delimiters when specifying dates:  / (slash)  - (hyphen)  . (period)</p> <p>You can also type single-digits for months and days, and two-digit years.</p> </div> <p>Supported languages:</p> <table border="0" style="width: 100%;"> <tr> <td style="vertical-align: top;"> <ul style="list-style-type: none"> <li>• de-DE</li> <li>• en-AU</li> <li>• en-GB</li> <li>• en-US</li> <li>• es-ES</li> <li>• es-US</li> <li>• fr-CA</li> </ul> </td> <td style="vertical-align: top;"> <ul style="list-style-type: none"> <li>• fr-FR</li> <li>• it-IT</li> <li>• ja-JP</li> <li>• nl-NL</li> <li>• pt-BR</li> <li>• zh-CN</li> </ul> </td> </tr> </table>	<ul style="list-style-type: none"> <li>• de-DE</li> <li>• en-AU</li> <li>• en-GB</li> <li>• en-US</li> <li>• es-ES</li> <li>• es-US</li> <li>• fr-CA</li> </ul>	<ul style="list-style-type: none"> <li>• fr-FR</li> <li>• it-IT</li> <li>• ja-JP</li> <li>• nl-NL</li> <li>• pt-BR</li> <li>• zh-CN</li> </ul>
<ul style="list-style-type: none"> <li>• de-DE</li> <li>• en-AU</li> <li>• en-GB</li> <li>• en-US</li> <li>• es-ES</li> <li>• es-US</li> <li>• fr-CA</li> </ul>	<ul style="list-style-type: none"> <li>• fr-FR</li> <li>• it-IT</li> <li>• ja-JP</li> <li>• nl-NL</li> <li>• pt-BR</li> <li>• zh-CN</li> </ul>				
digits	Reads strings digit-by-digit	Yes	<p>Supported languages:</p> <table border="0" style="width: 100%;"> <tr> <td style="vertical-align: top;"> <ul style="list-style-type: none"> <li>• de-DE</li> <li>• en-AU</li> <li>• en-GB</li> <li>• en-US</li> <li>• es-US</li> <li>• es-ES</li> <li>• fr-CA</li> </ul> </td> <td style="vertical-align: top;"> <ul style="list-style-type: none"> <li>• fr-FR</li> <li>• it-IT</li> <li>• ja-JP</li> <li>• nl-NL</li> <li>• pt-BR</li> <li>• zh-CN</li> </ul> </td> </tr> </table>	<ul style="list-style-type: none"> <li>• de-DE</li> <li>• en-AU</li> <li>• en-GB</li> <li>• en-US</li> <li>• es-US</li> <li>• es-ES</li> <li>• fr-CA</li> </ul>	<ul style="list-style-type: none"> <li>• fr-FR</li> <li>• it-IT</li> <li>• ja-JP</li> <li>• nl-NL</li> <li>• pt-BR</li> <li>• zh-CN</li> </ul>
<ul style="list-style-type: none"> <li>• de-DE</li> <li>• en-AU</li> <li>• en-GB</li> <li>• en-US</li> <li>• es-US</li> <li>• es-ES</li> <li>• fr-CA</li> </ul>	<ul style="list-style-type: none"> <li>• fr-FR</li> <li>• it-IT</li> <li>• ja-JP</li> <li>• nl-NL</li> <li>• pt-BR</li> <li>• zh-CN</li> </ul>				
number	Reads strings as a value (not digit-by-digit)	Yes	Supported languages:		

			<ul style="list-style-type: none"> <li>• de-DE</li> <li>• en-AU</li> <li>• en-GB</li> <li>• en-US</li> <li>• es-US</li> <li>• es-ES</li> <li>• fr-CA</li> </ul>	<ul style="list-style-type: none"> <li>• fr-FR</li> <li>• it-IT</li> <li>• ja-JP</li> <li>• nl-NL</li> <li>• pt-BR</li> <li>• zh-CN</li> </ul>	
ordinal	Processes ordinal numbers, such as "first", "second", and so on.	Yes	<p>You can specify only the digits or the digits and an ordinal ending for the supported language.</p> <p><b>Note:</b> If the language uses gendered forms for ordinals, and the gender is not specified in the text, ITTS hypothesizes the most-likely gender for the ordinal.</p> <p>Supported languages:</p>	<ul style="list-style-type: none"> <li>• de-DE</li> <li>• en-AU</li> <li>• en-GB</li> <li>• en-US</li> <li>• es-ES</li> <li>• es-US</li> <li>• it-IT</li> <li>• ja-JP</li> <li>• nl-NL</li> <li>• pt-BR</li> </ul>	<ul style="list-style-type: none"> <li>• fr-CA and fr-FR</li> </ul> <p>1 results in "premier" or "première" with the ordinal.</p> <p>1<sup>re</sup> results in "première"</p> <p>1 femme results in "première femme"</p> <p>1 cœur results in "premier"</p>
spell	Processes characters within a string, such as a word	Yes	<p>The synthesizer speaks each character individually. Any punctuation characters in the string are read and named, such as ampersand (&amp;), space (), and pound sign (#).</p> <p>Supported languages:</p>	<ul style="list-style-type: none"> <li>• de-DE</li> <li>• en-AU</li> <li>• en-GB</li> <li>• en-US</li> <li>• es-US</li> <li>• es-ES</li> <li>• fr-CA</li> </ul>	<ul style="list-style-type: none"> <li>• fr-FR</li> <li>• it-IT</li> <li>• jp-JP</li> <li>• nl-NL</li> <li>• pt-BR</li> <li>• zh-CN</li> </ul>
telephone	Processes telephone numbers	Yes	<p>The synthesizer reads telephone numbers, including SIP, according to local conventions in supported languages.</p>	<ul style="list-style-type: none"> <li>• de-DE</li> <li>• en-AU</li> <li>• en-GB</li> <li>• en-US</li> <li>• es-US</li> <li>• es-ES</li> <li>• fr-CA</li> </ul>	<ul style="list-style-type: none"> <li>• fr-FR</li> <li>• it-IT</li> <li>• ja-JP</li> <li>• nl-NL</li> <li>• pt-BR</li> <li>• zh-CN</li> </ul>

time	Processes time statements, such as "12:45 AM"	Yes	<p>Supports hours, minutes, seconds, and 12 or 24-hour clock.</p> <p>This normalization type does not support any <code>format</code> options.</p> <p>This normalization type does not support durations, such as "60 minutes".</p> <p>For all languages, the separator for the hours and minutes of the time is a colon (:). For top-of-the-hour times, such as "2 o'clock", you can omit the colon and minutes. Example: 2 pm</p> <p>Supported languages:</p> <table border="0" data-bbox="748 415 1510 741"> <tr> <td data-bbox="748 415 1239 741"> <ul style="list-style-type: none"> <li>• de-DE</li> <li>• en-AU</li> <li>• en-GA</li> <li>• en-US</li> <li>• es-ES</li> <li>• es-US</li> <li>• fr-CA/fr-FR</li> </ul> <p>These languages support usage of <code>h</code> as a separator between the hours and minutes, such as <code>1h30</code>.</p> </td> <td data-bbox="1239 415 1510 741"> <ul style="list-style-type: none"> <li>• it-IT</li> <li>• ja-JP</li> <li>• nl-NL</li> <li>• pt-BR</li> <li>• zh-CN</li> </ul> </td> </tr> </table>	<ul style="list-style-type: none"> <li>• de-DE</li> <li>• en-AU</li> <li>• en-GA</li> <li>• en-US</li> <li>• es-ES</li> <li>• es-US</li> <li>• fr-CA/fr-FR</li> </ul> <p>These languages support usage of <code>h</code> as a separator between the hours and minutes, such as <code>1h30</code>.</p>	<ul style="list-style-type: none"> <li>• it-IT</li> <li>• ja-JP</li> <li>• nl-NL</li> <li>• pt-BR</li> <li>• zh-CN</li> </ul>
<ul style="list-style-type: none"> <li>• de-DE</li> <li>• en-AU</li> <li>• en-GA</li> <li>• en-US</li> <li>• es-ES</li> <li>• es-US</li> <li>• fr-CA/fr-FR</li> </ul> <p>These languages support usage of <code>h</code> as a separator between the hours and minutes, such as <code>1h30</code>.</p>	<ul style="list-style-type: none"> <li>• it-IT</li> <li>• ja-JP</li> <li>• nl-NL</li> <li>• pt-BR</li> <li>• zh-CN</li> </ul>				

## User-defined Dictionaries

Beginning with CIC 2019 R3, Interaction Text to Speech (ITTS) supports user-defined dictionaries. This feature allows you to define specific pronunciations of names and words to use for speech recognition and in the playing of prompts in speech synthesis.

Pronunciation Lexicon Specification (PLS) is a definition that allows automated speech recognition and text-to-speech engines to use external dictionaries during speech recognition and speech synthesis. For more information, see <https://www.w3.org/TR/pronunciation-lexicon/>.

You can set specific word and name pronunciations in a lexicon file overriding default pronunciation, when it improves customer understanding or experience.

### VoiceXML

The VoiceXML standard supports a feature that can use external dictionaries with speech recognition and TTS engines. For speech recognition, you specify the dictionary in the grammar file in a specially-defined lexicon element. You can specify multiple dictionaries at different points in the IVR dialog flow. For TTS prompts, you specify the dictionary in the lexicon element defined inside a prompt element. For more information about definitions and the VoiceXML standard, see <https://www.w3.org/TR/voicexml20/>.

In CIC, reference the VoiceXML document containing the lexicon in the **Document URI** field of a handler subroutine. For more information about handler setup, see the [https://help.genesys.com/cic/mergedProjects/wh\\_tr/desktop/pdfs/voicexml\\_tr.pdf](https://help.genesys.com/cic/mergedProjects/wh_tr/desktop/pdfs/voicexml_tr.pdf). For more information about specifying lexicons in grammar files, see <https://www.w3.org/TR/speech-grammar/#S4.10>.

**SSML:** ITTS can receive Speech Synthesis Markup Language (SSML) documents or strings directly as input and process the specified prompts. You specify the user dictionary in a specially-defined lexicon element that must comply with SSML syntax. In CIC, the SSML string referencing the lexicon is a string variable in the **Optional Parameters** field of a handler subroutine. For more information, see <https://www.w3.org/TR/speech-synthesis/#S3.1.4>.

### Example of a lexicon file with an alternate pronunciation for the name Anna

```
<?xml version="1.0" encoding="UTF-8"?>
<lexicon version="1.0"
  xmlns="http://www.w3.org/2005/01/pronunciation-lexicon"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.w3.org/2005/01/pronunciation-lexicon
    http://www.w3.org/TR/2007/CR-pronunciation-lexicon-20071212/pls.xsd"
  alphabet="x-inin-arpabet" xml:lang="en-US">
  <lexeme>
    <grapheme>Anna</grapheme>
    <phoneme>ah n ah</phoneme>
  </lexeme>
</lexicon>
```

### Example of lexicon usage in SSML

```

<speak version="1.0" xml:lang="en-US"
xmlns="http://www.w3.org/2001/10/synthesis"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.w3.org/TR/speech-synthesis/synthesis.xsd" xml:base="file:///path/to/file/">
    <lexicon uri="file:///path/to/file/lexicon_pron.pls" xml:id="alt_pron_for_Anna"/>
    <lookup ref="alt_pron_for_Anna"> Hello Anna</lookup>
</speak>

```

### Example of lexicon usage in VoiceXML

```

<?xml version="1.0" encoding="UTF-8"?>
<!--
ASR of name Anna followed by a greeting
-->
<vxml xmlns="http://www.w3.org/2001/vxml"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.w3.org/2001/vxml
http://www.w3.org/TR/voicexml20/vxml.xsd"
version="2.0">
<form>
<field name="name">
<prompt>
Please say your name
</prompt>
<grammar type="application/srgs+xml" src="name.grxml"/>
</field>
<filled>
<prompt>
<lexicon uri="file:///path/to/file/alt_pron_for_anna.pls"/>
Hello <value expr="name"/></prompt>
<exit/>
</filled>
</form>
</vxml>

```



## Configure the TTS engine in Interaction Administrator

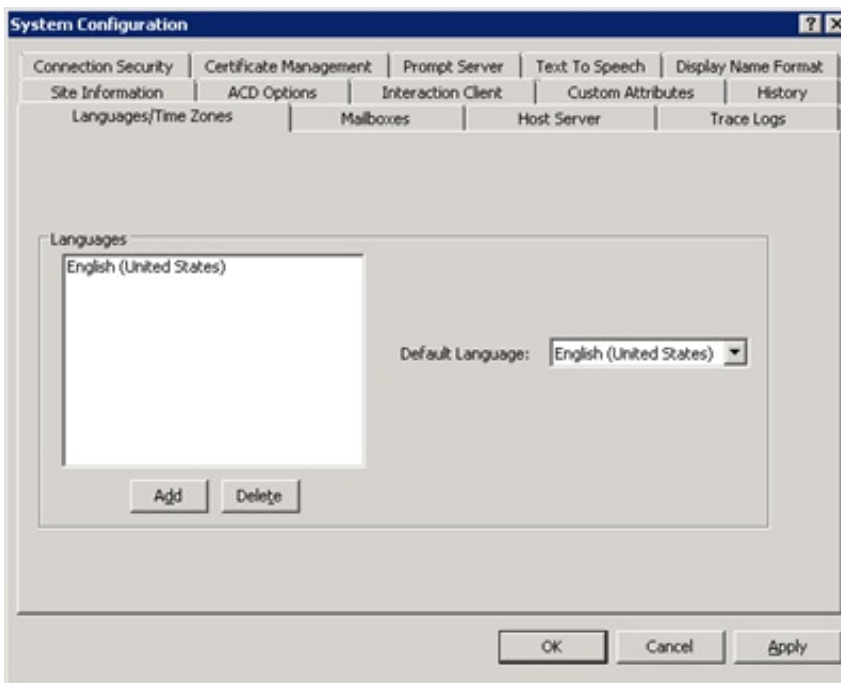
Use Interaction Administrator to configure TTS Engine features. For more information about this configuration, see the *Interaction Administrator Help* at [https://help.genesys.com/cic/mergedProjects/wh\\_ia/desktop/interaction\\_administrator\\_help.htm](https://help.genesys.com/cic/mergedProjects/wh_ia/desktop/interaction_administrator_help.htm).

### To configure the TTS Engine

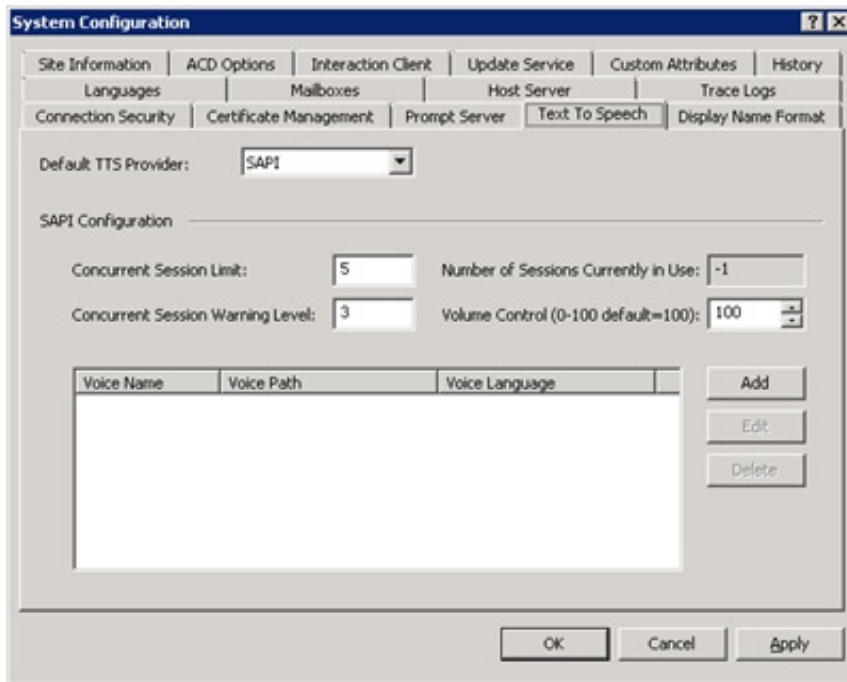
1. Log on to Interaction Administrator with administrator credentials.
2. In the navigation pane, click the **System Configuration** container.



3. In the right pane, double-click **Configuration**. The **System Configuration** dialog box appears.



4. Click the **Text To Speech** tab. The **Text To Speech** tab appears.



5. In the **Default TTS Provider** list box, click the TTS engine to use:
  - **SAPI** (default)- Uses the Microsoft Speech API (SAPI) of the Windows Server operating system on the CIC server.
  - **MRCP** - Uses a third-party TTS engine, such as Nuance or Loquendo.
  - **Media Server** - Uses the Interaction Text to Speech engine of Interaction Media Server.

**Note:**

The **Media Server** item appears only if you applied the Interaction Text to Speech feature license to your CIC server.

6. If you are using SAPI for your TTS solution, do the following:
  - a. In the **Concurrent Session Limit** box, type the maximum number of concurrent sessions allowed. The limit is either a license-enforced limit or a load-enforced limit. For example, if you have a 20-port license, the system cannot connect to more than 20 sessions.
  - b. In the **Concurrent Session Warning Level** box, type the minimum number of concurrent sessions that can be active before a warning message appears. The warning message indicates that you are close to exceeding the concurrent session limit.
  - c. In the **Volume Control** box, type or select the volume level for the voice. 100 is the default value and the maximum value.
  - d. Click **OK**.

**Note:**

To add a voice and language for SAPI, see [Add Voices and Languages for SAPI](#).

# Add Voices and Languages for SAPI

You can write custom applications for multiple voices and languages by creating a voice name parameter for each voice and modifying the handler to use these SAPI voice name parameters.

If you downloaded and installed Microsoft Speech Runtime Platform on the CIC server and want to add voices that it provides, define the voices in Interaction Administrator and reference the Registry location where the tokens are. The base Registry path for the Microsoft Speech Runtime Platform voices is as follows:

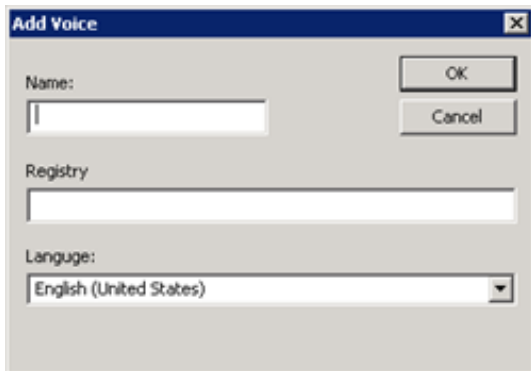
```
HKEY_LOCAL_MACHINE
  \SOFTWARE
    \Wow6432Node
      \Microsoft
        \Speech Server
          \v11.0
            \Voices
              \Tokens
```

The **Text To Speech** tab in the **System Configuration** dialog box in Interaction Administrator allows you to add multiple voices and languages. You can add an unlimited number of voices; however, you can associate each language to only one voice. Voice configuration settings on this tab override the voice configuration settings in the **Windows Speech** applet.

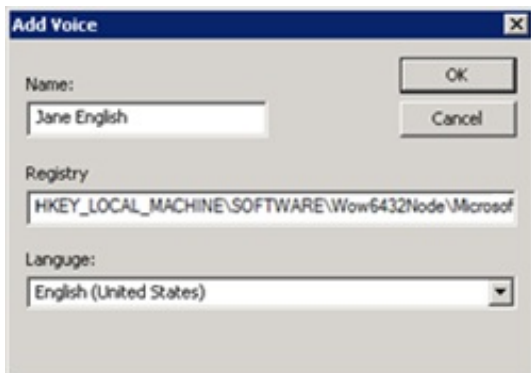
After defining the voice, you can pass the voice name parameter (for example, "Jane English") to the TTS-defined tool.

## To add a voice for a language

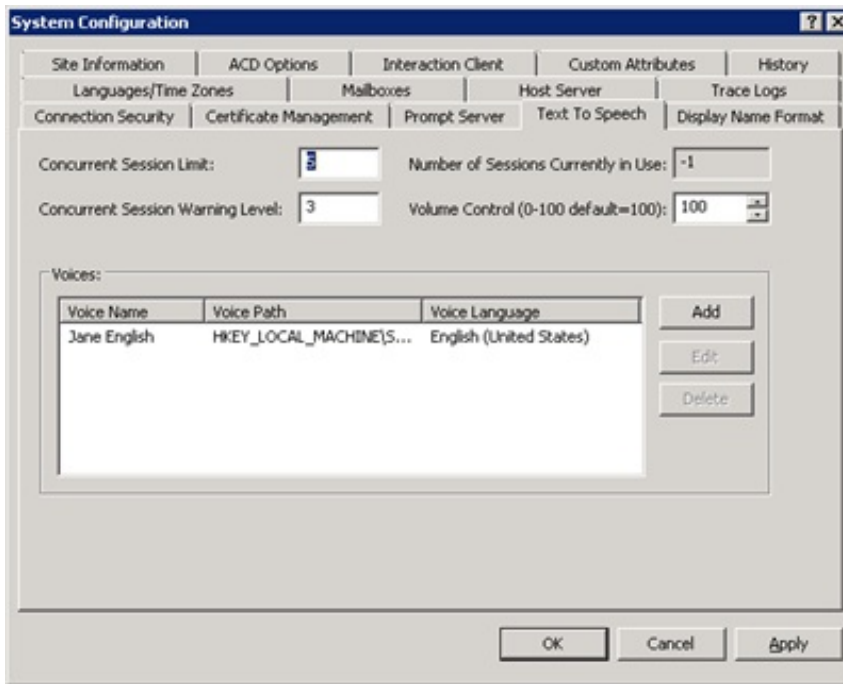
1. Log on to Interaction Administrator with administrator credentials.
2. In the navigation pane, click the **System Configuration** container.
3. In the right pane, double-click **Configuration**. The **System Configuration** dialog box appears.
4. Click the **Text To Speech** tab.
5. On the **Text to Speech** tab, click **Add**. The **Add Voice** dialog box appears.



6. In the **Name** box, type the name to assign to the voice.
7. In the **Registry** box, type the registry path to the voice token.
8. In the **Language** list box, click the language in which the voice is spoken.



9. Click **OK**. The voice appears in the **Voices** section of the **Text to Speech** tab.



10. In the **Volume Control** box, type or select the volume level for the voice. 100 is the default value and the maximum value.
11. Click **OK**.

For more information about these options, see the *Interaction Administrator Help* at [https://help.genesys.com/cic/mergedProjects/wh\\_ia/desktop/interaction\\_administrator\\_help.htm](https://help.genesys.com/cic/mergedProjects/wh_ia/desktop/interaction_administrator_help.htm).

# Change Log

The following table lists the changes to the *CIC Text to Speech Engines Technical Reference* since its initial release.

Date	Changes
12-October-2012	Updated for CIC 4.0 SU3; removed references to HMP.
29-April-2013	<ul style="list-style-type: none"> <li>• Updated title page and copyright notice.</li> <li>• Updated reference to Microsoft text-to-speech website (now Tellme).</li> <li>• Added reference to <i>MRCP Technical Reference</i>.</li> <li>• Updated reference to IETF document RFC 6787.</li> </ul>
13-February-2014	<ul style="list-style-type: none"> <li>• Updated Copyright notice.</li> <li>• Updated registry path for the Jane English voice.</li> </ul>
30-July-2014	Updated documentation to reflect changes required in the transition from version 4.0 SU# to CIC 2015 R1, such as updates to product version numbers, system requirements, installation procedures, references to Interactive Intelligence Product Information site URLs, and copyright and trademark information.
01-July-2015	<ul style="list-style-type: none"> <li>• Updated cover page to reflect new color scheme and logo.</li> <li>• Updated copyright and trademark information.</li> </ul>
09-October-2015	Updated the document to reflect the CIC 2016 R1 version.
09-February-2016	<ul style="list-style-type: none"> <li>• Updated copyright and trademark information</li> <li>• Added content for Interaction Text-to-Speech (ITTS)</li> <li>• Updated content to conform to latest TTS offerings and compatibilities</li> <li>• Applied general edits for clarity and conformity</li> </ul>
13-June-2016	<ul style="list-style-type: none"> <li>• IC-135402 - Made various edits, clarifications, and improvements throughout the document for Interaction Text to Speech (ITTS)</li> <li>• IONMEDIA-2412 - Added French -France (fr-FR) as a supported language model for ITTS</li> <li>• IONMEDIA-2631 - Added Mandarin Chinese (zh-CN) as a beta version of a language model for ITTS</li> <li>• IONMEDIA-2630 - Added Dutch (nl-NL) as a supported language model for ITTS</li> <li>• IONMEDIA-2483, 2472 - Added Brazilian Portuguese (pt-BR) as a supported language model for ITTS</li> <li>• IONMEDIA-2400, 2401,2424: Say-as support improvements - Added support for more language models in text normalizations</li> </ul>
22-September-2016	IONMEDIA-2754 - Added support for <code>address</code> say-as text normalization for en-US in "Supported say-as text normalization."
12-December-2016	Updated "Supported say-as text normalization" with the following changes: <ul style="list-style-type: none"> <li>• IONMEDIA-2832 - Added zh-CN support for <code>currency</code>, <code>date</code>, <code>digits</code>, <code>number</code>, <code>telephone</code>, and <code>time</code> say-as text normalizations</li> <li>• IONMEDIA-2826 - Added pt-BR support for <code>alphanumeric</code> say-as text normalization</li> <li>• IONMEDIA-2827 - Added pt-BR support for <code>Boolean</code> say-as text normalization</li> <li>• IONMEDIA-2828 - Added pt-BR support for <code>digits</code> say-as text normalization</li> <li>• IONMEDIA-2829 - Added pt-BR support for <code>number</code> say-as text normalization</li> <li>• IONMEDIA-2830 - Added pt-BR support for <code>spell</code> say-as text normalization</li> </ul>
24-May-2017	Added language support for Italian.
01-June-2017	Added language support for Dutch

24-October-2017	<ul style="list-style-type: none"> <li>• Rebranded this document to apply Genesys styles and terminology.</li> <li>• Updated "Supported languages for Interaction Text to Speech", to note, for zh-CN, only Latin characters are supported in say-as "alphanumeric", not Chinese characters. Improved TTS reading of dates for English languages: Text to Speech (TTS) better interprets text to read back a date instead of a fraction, where applicable. For example, "12/15" is ambiguous, since that could represent a fraction, a date (month/year or month/day), or a number sequence. Previously, "12/15" would have been read back as a fraction. Following this update, outside of a say-as environment, TTS reads back a date unless it finds additional context to indicate that a fraction or number is intended. The format used to read back dates is locale-specific. For example, "1/2" in an es-US date context will be read "January 2". In an en-GB/AU context, "February 1" is read back instead. In addition, TTS better interprets years. It reads back "nineteen eighty five" instead of "one thousand nine hundred eighty five". Plural years or decade references such as "1930s" are read back as "nineteen thirties", but not "one thousand nine hundred thirty S".</li> <li>• Say-as-alphanumeric TTS improvements: To improve the customer experience, say-as-alphanumeric input is retained, even if some characters are not included in a language's character set. For example, Mandarin doesn't include the Polish letters or . If someone inputs "Lech Wasa" in zh-CN as say-as input, TTS will input "Lech Wasa" into character normalizer. Previously, character normalizer stripped out any non-Mandarin characters, resulting in "Lech Wasa". Starting with this release, TTS will output "Lech Walesa" in the case of zh-CN. By retaining more of the input, better pronunciation can be attained.</li> </ul>
08-January-2018	<p>Updated a note in "Supported languages for Interaction Text to Speech" which said, "To ensure that Interaction Media Server does not use too many CPU resources, PureConnect recommends that you use no more than four languages with Interaction Text to Speech. Overuse of Interaction Media Server resources can result in defects or failures in audio processing." The note now says:</p> <p style="padding-left: 40px;">"To ensure that Interaction Media Server does not exceed memory resources, customers should test the performance and memory usage of Media Servers when using more than 4 TTS languages. Older Media Servers may not be able to handle more than 4 languages. Overuse of Interaction Media Server resources can result in defects or failures in audio processing."</p>
06-March-2019	<ul style="list-style-type: none"> <li>• IONMEDIA-4052- Upgraded Mandarin Chinese, China (zh-CN) to a fully supported language model for ITTS</li> <li>• IONMEDIA-4052: Added Spanish, Spain (es-ES) as a supported language model for ITTS</li> </ul>
06-June-2019	Reorganized the content only, which included combining some topics and deleting others that just had an introductory sentence such as, "In this section...".
11-June-2019	Added support for user-defined dictionaries.