



This PDF is generated from authoritative online content, and is provided for convenience only. This PDF cannot be used for legal purposes. For authoritative understanding of what is and is not supported, always use the online content. To copy code samples, always use the online content.

Genesys Voice Platform

Speech-Related Components

12/17/2025

Speech-Related Components

These components provide the interface to third-party speech servers that deliver speech recognition and text-to-speech services for VoiceXML. GVP uses the Media Resource Control Protocol to communicate with these services. The MCP includes client interfaces for MRCPv1 and MRCPv2. These clients are capable of load-balancing across pools of speech servers, and service monitoring of speech servers for availability purposes. GVP also includes an MRCPv1 proxy that provides global load balancing, peak reporting, and other services.

- **MRCP Proxy**

MRCP Proxy is an integral component that interfaces with the MRCP, Management Framework, and Operational Reporting API. The MRCP Proxy can be placed between the Media Control Platforms and the MRCPv1 resources within a GVP deployment. Deploying the MRCP Proxy enables ASR/TTS usage reporting data to be sent to the Reporting Server.

For more information about the MRCP Proxy application, see **MRCP Proxy** in the *GVP Deployment Guide*. For various MRCP Proxy configuration options, see **MRCP Proxy Options** in this document.

- **MRCPv1/v2 Services**

GVP provides transparent access to MRCP services from VoiceXML. GVP will acquire and releases ASR and TTS sessions as required for proper execution of a VoiceXML applications. Text to Speech services render audio streams from text, allowing playback of dynamic text for which pre-recorded prompts may be difficult or impossible to provide. Text rendering can be controlled using the Speech Synthesis Markup Language (SSML), and can support multiple languages.

Speech Recognition services can process caller audio and, using either pre-defined context free grammars or natural language models, return an interpretation of what the caller said. Speech recognition services support the Speech Recognition Grammar Specification (SRGS) for grammar specification, and the Natural Language Semantic Markup Language (NLSML) specification for describing the results of a recognition.