

Voice.AI Gateway

Version 1.6



Notice

Information contained in this document is believed to be accurate and reliable at the time of printing. However, due to ongoing product improvements and revisions, AudioCodes cannot guarantee accuracy of printed material after the Date Published nor can it accept responsibility for errors or omissions. Updates to this document can be downloaded from <https://www.audiocodes.com/library/technical-documents>.

This document is subject to change without notice.

Date Published: June-15-2020

WEEE EU Directive

Pursuant to the WEEE EU Directive, electronic and electrical waste must not be disposed of with unsorted waste. Please contact your local recycling authority for disposal of this product.

Customer Support

Customer technical support and services are provided by AudioCodes or by an authorized AudioCodes Service Partner. For more information on how to buy technical support for AudioCodes products and for contact information, please visit our website at <https://www.audiocodes.com/services-support/maintenance-and-support>.

Documentation Feedback

AudioCodes continually strives to produce high quality documentation. If you have any comments (suggestions or errors) regarding this document, please fill out the Documentation Feedback form on our website at <https://online.audiocodes.com/documentation-feedback>.

Stay in the Loop with AudioCodes



Notes and Warnings



OPEN SOURCE SOFTWARE. Portions of the software may be open source software and may be governed by and distributed under open source licenses, such as the terms of the GNU General Public License (GPL), the terms of the Lesser General Public License (LGPL), BSD and LDAP, which terms are located at <https://www.audiocodes.com/services-support/open-source/> and all are incorporated herein by reference. If any open source software is provided in object code, and its



accompanying license requires that it be provided in source code as well, Buyer may receive such source code by contacting AudioCodes, by following the instructions available on AudioCodes website.

Related Documentation

Document Name
Voice.AI Gateway Bot API Reference Guide
Voice.AI Gateway Product Description
Voice.AI Gateway with One-Click Dialogflow Integration Guide

Document Revision Record

LTRT	Description
30920	Initial document release.
30921	Parameters added- sttContextId, , sttContextPhrases, sttContextBoost.
30922	"Amazon Lex" and "Google Dialogflow" terms added; AudioCodes API syntax example for initial sent message; hangupReason updated (CDR); handover changed to transfer; event type (e.g., string) added; sttEndpointID (description updated)' miscellaneous.
30923	Updated to Ver. 1.4. VOICE_AI_WELCOME event replaced by WELCOME.
30924	Dialogflow text length limitation; typo (Product Notice replaced by Product Description).
30926	Updated to Ver. 1.6; Nuance add for STT/TTS; transferSipHeaders (typo); transferReferredByURL (added); sttContextId (updated)
30927	Bot-controlled parameters added.

Table of Contents

1	Introduction	1
	Purpose	1
	Targeted Audience	1
2	Required Information	2
	Required Information of Bot Framework Provider	2
	Required Information of STT Provider	3
	Required Information of TTS Provider	4
3	Messages Sent by Voice.AI Gateway	7
	Initial Message	7
	End of Conversation Message	9
	Text Message	9
	DTMF Event	11
4	Messages Sent by Bot	13
	Basic Activity Syntax	13
	message Activity	13
	event Activities	14
	hangup	15
	transfer	15
	Adding SIP Headers on Call Transfer	16
	playUrl	17
	config	18
	Bot Framework Specific Details	18
	AudioCodes Bot API	18
	Microsoft Azure	19
	Google Dialogflow	19
	Parameters Controlled Also by Bot	21

1 Introduction

AudioCodes Voice.AI Gateway enhances chatbot functionality by allowing human communication with chatbots through voice (voicebot), offering an audio-centric user experience. Integrating the Voice.AI Gateway into your chatbot environment provides you with a single-vendor solution, assisting you in migrating your text-based chatbot experience into a voice-based chatbot.



- Prior to reading this document, it is recommended that you read the [Voice.AI Gateway Product Description](#) to familiarize yourself with AudioCodes Voice.AI Gateway architecture and solution.
- Most of the information provided in this document is relevant to all bot frameworks. Where a specific bot framework uses different syntax, a note will indicate this.

Purpose

This guide provides the following:

- Information that you need to supply AudioCodes for connecting the Voice.AI Gateway to the third-party cognitive services used in your chatbot environment - bot framework(s), speech-to-text (STT) engine(s), and text-to-speech (TTS) engine(s).
- Description of the messages sent by the Voice.AI Gateway to the bot, and messages sent by the bot to the Voice.AI Gateway to achieve the desired functionality. These descriptions allow the bot developer to adapt the bot's behavior to the voice and telephony engagement channels.

Targeted Audience

This guide is intended for IT Administrators and Bot Developers who want to integrate AudioCodes Voice.AI Gateway into their bot solution.

2 Required Information

This section lists the information that you need to supply AudioCodes for integrating and connecting the Voice.AI Gateway to the cognitive services of your chatbot environment. This includes information of the bot framework, Speech-to-Text (STT) provider, and Text-to-Speech (TTS) provider used in your environment.

Required Information of Bot Framework Provider

To connect the Voice.AI Gateway to bot frameworks, you need to provide AudioCodes with the bot framework provider's details, as listed in the following table.

Table 2-1: Required Information per Bot Framework

Bot Framework	Required Information
Microsoft Azure	<p>To connect to Microsoft Azure Bot Framework, you need to provide AudioCodes with the bot's secret key. To obtain this key, refer to Azure's documentation at https://docs.microsoft.com/en-us/azure/bot-service/bot-service-channel-connect-directline.</p> <p>Note: Microsoft Azure Bot Framework Direct Line Version 3.0 must be used.</p>
AWS	<p>To connect to Amazon Lex, you need to provide AudioCodes with the following:</p> <ul style="list-style-type: none"> ■ AWS account keys: <ul style="list-style-type: none"> ✓ Access key ✓ Secret access key <p>To obtain these keys, refer to the AWS documentation at https://docs.aws.amazon.com/general/latest/gr/managing-aws-access-keys.html.</p> <p>Note: The same keys are used for all Amazon services (STT, TTS and bot framework).</p> <ul style="list-style-type: none"> ■ Name of the specific bot ■ AWS Region (e.g., "us-west-2")
Google	<p>To connect to Google Dialogflow, you need to provide AudioCodes with the following:</p> <ul style="list-style-type: none"> ■ Private key of the Google service account. For information on how to create the account key, refer to Google's documentation at https://cloud.google.com/iam/docs/creating-managing-service-accounts

Bot Framework	Required Information
	<p>account-keys. From the JSON object representing the key, you need to extract the private key (including the "-----BEGIN PRIVATE KEY-----" prefix) and the service account email.</p> <ul style="list-style-type: none"> ■ Client email ■ Project ID (of the bot)
AudioCodes Bot API	To create the channel between the Voice.AI Gateway's Cognitive Service component and the bot provider, refer to the document Voice.AI Gateway API Reference Guide .

Required Information of STT Provider

To connect the Voice.AI Gateway to third-party, speech-to-text (STT) engines, you need to provide AudioCodes with the STT provider's details, as listed in the following table.

Table 2-2: Required Information per Supported STT Provider

STT Provider	Required Information from STT Provider	
	Connectivity	Language Definition
Microsoft Azure Speech Services	<p>To connect to Azure's Speech Service, you need to provide AudioCodes with your subscription key for the service. To obtain the key, see Azure's documentation at https://docs.microsoft.com/en-us/azure/cognitive-services/speech-service/get-started.</p> <p>Note: The key is only valid for a specific region.</p>	<p>To connect to Azure Speech Services, you need to provide AudioCodes with the following:</p> <ul style="list-style-type: none"> ■ Relevant value in the 'Locale' column in Azure's Text-to-Speech table (see below). <p>For example, for Italian (Italy), the 'Locale' column value is "it-IT".</p> <p>For languages supported by Azure's Speech Services, see the Speech-to-text table in Azure's documentation at https://docs.microsoft.com/en-us/azure/cognitive-services/speech-service/language-support.</p> <p>The Voice.AI Gateway can also use Azure's Custom Speech service. For more information, see Azure's documentation at https://docs.microsoft.com/en-us/azure/cognitive-services/speech-service/language-support.</p>

STT Provider	Required Information from STT Provider	
		us/azure/cognitive-services/speech-service/how-to-custom-speech-deploy-model . If you do use this service, you need to provide AudioCodes with the custom endpoint details.
Google Cloud Speech-to-Text	To connect to Google Cloud Speech-to-Text service, see Required Information of Bot Framework Provider on page 2 for required information.	To connect to Google Cloud Speech-to-Text, you need to provide AudioCodes with the following: <ul style="list-style-type: none"> ■ Relevant value in the 'languageCode' column in Google's Cloud Speech-to-Text table (see below). For example, for English (South Africa), the 'Language code' column value is "en-ZA". For languages supported by Google Cloud Speech-to-Text, see Google's documentation at https://cloud.google.com/speech-to-text/docs/languages .
Yandex	Contact AudioCodes for more information.	Contact AudioCodes for more information.
Nuance	Contact AudioCodes for more information.	Contact AudioCodes for more information.

Required Information of TTS Provider

To connect the Voice.AI Gateway to third-party, text-to-speech (TTS) engines, you need to provide AudioCodes with the TTS provider's details, as listed in the following table.

Table 2-3: Required Information per Supported TTS Provider

TTS Provider	Required Information from TTS Provider	
	Connectivity	Language Definition
Microsoft	To connect to Azure's Speech Service, you need	To connect to Azure Speech Services, you need to provide AudioCodes with the following:

TTS Provider	Required Information from TTS Provider	
<p>Azure Speech Services</p>	<p>to provide AudioCodes with your subscription key for the service. To obtain the key, see Azure's documentation at https://docs.microsoft.com/en-us/azure/cognitive-services/speech-service/get-started. Note: The key is valid only for a specific region.</p>	<ul style="list-style-type: none"> ■ Relevant value in the 'Locale' column in Azure's Text-to-Speech table (see below link). ■ Relevant value in the 'Short voice name' column in Azure's Text-to-Speech table (see below link). <p>For example, for Italian (Italy), the 'Locale' column value is "it-IT" and the 'Short voice name' column value is "it-IT-ElsaNeural".</p> <p>For languages supported by Azure's Speech Services, see the Text-to-Speech table in Azure's documentation at https://docs.microsoft.com/en-us/azure/cognitive-services/speech-service/language-support.</p>
<p>Google Cloud Text-to-Speech</p>	<p>To connect to Google Cloud Text-to-Speech service, see Required Information of Bot Framework Provider on page 2 for required information.</p>	<p>To connect to Google Cloud Text-to-Speech, you need to provide AudioCodes with the following:</p> <ul style="list-style-type: none"> ■ Relevant value in the 'Language code' column in Google's table (see below link). ■ Relevant value in the 'Voice name' column in Google's table (see below link). <p>For example, for English (US), the 'Language code' column value is "en-US" and the 'Voice name' column value is "en-US-Wavenet-A".</p> <p>For languages supported by Google Cloud Text-to-Speech, see Google's documentation at https://cloud.google.com/text-to-speech/docs/voices.</p>
<p>AWS Amazon Polly</p>	<p>To connect to Amazon Polly Text-to-Speech service, see Required Information of Bot Framework Provider on page 2 for required information.</p>	<p>To connect to Amazon Polly TTS service, you need to provide AudioCodes with the following:</p> <ul style="list-style-type: none"> ■ Relevant value in the 'Language' column in Amazon Polly TTS table (see below link). ■ Relevant value in the 'Name/ID' column in Amazon Polly TTS table (see below link). <p>For example, for English (US), the 'Language' column value is "English, US (en-US)" and the</p>

TTS Provider	Required Information from TTS Provider	
		'Name/ID' column is "Matthew". For languages supported by Amazon Polly TTS service, see the table in https://docs.aws.amazon.com/polly/latest/dg/voice-list.html .
Yandex	Contact AudioCodes for more information.	Contact AudioCodes for more information.
Almagu	Contact AudioCodes for more information.	Contact AudioCodes for more information.
Nuance	Contact AudioCodes for more information.	Contact AudioCodes for more information.

3 Messages Sent by Voice.AI Gateway

This section describes the messages that are sent by the Voice.AI Gateway.

Initial Message

When the conversation starts, a message is sent with the details of the call. These details include (when available) the following:

Table 3-1: Description of Initial Message Sent by Voice.AI Gateway

Property	Type	Description
<code>callee</code>	String	Dialed phone number. This is typically obtained from the SIP To header.
<code>calleeHost</code>	String	Host part of the destination of the call. This is typically obtained from the SIP To header.
<code>caller</code>	String	Caller's phone number. This is typically obtained from the SIP From header.
<code>callerHost</code>	String	Host part of the source of the call. This is typically obtained from the SIP From header.
<code>callerDisplayName</code>	String	Caller's display name. This is typically obtained from the SIP From header.
<code><Additional attributes></code>	-	Defines additional attributes such as values from various SIP headers. These can be added by customization. The Voice.AI Gateway can be configured to extract values from the SIP INVITE message and then send them as additional attributes in the initial message to the bot.

The syntax of the initial message depends on the specific bot framework:

Table 3-2: Syntax of Initial Message Sent by Voice.AI Gateway

Bot Framework	Message Syntax
AudioCodes Bot API	<p>The message is sent as a <code>start</code> event, with the details inside the <code>parameters</code> property.</p> <p>Example:</p> <pre>{ "type": "event",</pre>

Bot Framework	Message Syntax
	<pre> "name": "start", "parameters": { "callee": "12345678", "calleeHost": "10.20.30.40", "caller": "12345678", "callerHost": "10.20.30.40" } } </pre>
Microsoft Azure	<p>The message is sent as a <code>channel</code> event, with the details inside the <code>channelData</code> property.</p> <p>Example:</p> <pre> { "type": "event", "name": "channel", "value": "telephony", "channelData": { "callee": "12345678", "calleeHost": "10.20.30.40", "caller": "12345678", "callerHost": "10.20.30.40" }, "from": { "id": "12345678" }, "locale": "en-US" } </pre>
Google Dialogflow	<p>The message is sent as a <code>WELCOME</code> event, with the details as <code>event</code> parameters.</p> <p>Example:</p> <pre> { "queryInput": { "event": { "languageCode": "en-US", "name": "WELCOME", "parameters": { "callee": "12345678", </pre>

Bot Framework	Message Syntax
	<pre data-bbox="517 309 1165 609"> "calleeHost": "10.20.30.40", "caller": "12345678", "callerHost": "10.20.30.40" } } } </pre> <p data-bbox="517 629 1385 696">Note: These parameters can be used when generating the response text, by using a syntax such as this:</p> <pre data-bbox="517 712 823 741">#WELCOME.caller" </pre>

End of Conversation Message

The syntax of the end-of-conversation message depends on the specific bot framework:

Table 3-3: Syntax of End-of-Conversation Message Sent by Voice.AI Gateway

Bot Framework	Message Syntax
AudioCodes Bot API	The conversation is terminated according to the AC Bot API documentation.
Microsoft Azure	<p>The conversation is terminated by sending an <code>endOfConversation</code> activity, with an optional <code>text</code> property with a textual reason.</p> <p>Example:</p> <pre data-bbox="504 1379 1040 1541"> { "type": "endOfConversation", "text": "Client Side" } </pre>
Google Dialogflow	Currently, no indication is sent for the end of conversation.

Text Message

When the speech-to-text engine detects user utterance, it is sent as a message to the bot. The message may contain details gathered by the speech-to-text engine. These details include:

Table 3-4: Description of Text Message Sent by Voice.AI Gateway

Property	Type	Description
<code>confidence</code>	Number	Numeric value representing the confidence level of the recognition.
<code>recognitionOutput</code>	Object	Raw recognition output of the speech-to-text engine (vendor specific).
<code>recognitions</code>	Array of Objects	If Continuous ASR mode is enabled, this array contains the separate recognition outputs.

The syntax of the text message depends on the specific bot framework:

Table 3-5: Syntax of Text Message Sent by Voice.AI Gateway

Bot Framework	Message Syntax
AudioCodes Bot API	<p>The message is sent as a <code>message</code> activity. Additional details are sent in the <code>parameters</code> property.</p> <p>Example:</p> <pre>{ "type": "message", "text": "Hi.", "parameters": { "confidence":0.6599681, } }</pre>
Microsoft Azure	<p>The message is sent as a <code>message</code> activity. Additional details are sent in the <code>channelData</code> property.</p> <p>Example:</p> <pre>{ "type": "message", "text": "Hi.", "channelData": { "confidence":0.6599681, } }</pre>
Google Dialogflow	<p>The message is sent as text input. Currently, additional details are not sent.</p> <p>Example:</p>

Bot Framework	Message Syntax
	<pre data-bbox="504 309 1150 656"> { "queryInput": { "text": { "languageCode": "en-US", "text": "Hi." } } } </pre> <p data-bbox="504 674 1385 824">Note: Dialogflow supports a maximum text input length of 256 characters. Therefore, if the input received from the speech-to-text engine is longer than 256 characters, the Voice.AI Gateway truncates the message before sending it to Dialogflow.</p>

DTMF Event

The syntax for DTMF tone signals (i.e., keys pressed on phone keypad by user) depends on the specific bot framework.

Table 3-6: Syntax of DTMF Sent by Voice.AI Gateway

Bot Framework	Message Syntax
AudioCodes Bot API / Microsoft Azure	<p data-bbox="652 1216 1353 1283">This message is sent as a DTMF event with the digits as the value of the event.</p> <p data-bbox="652 1301 762 1330">Example:</p> <pre data-bbox="652 1348 970 1559"> { "type": "event", "name": "DTMF", "value": "3" } </pre>
Google Dialogflow	<p data-bbox="652 1603 1353 1671">This message is sent as a DTMF event with the digits as the event parameters.</p> <p data-bbox="652 1688 762 1718">Example:</p> <pre data-bbox="652 1736 1190 1993"> { "queryInput": { "event": { "languageCode": "en-US", "name": "DTMF", "parameters": { </pre>

Bot Framework	Message Syntax
	<pre data-bbox="651 264 1034 481"> "digits": "3" } } }</pre> <p data-bbox="651 499 1362 571">Note: The digits can be used when generating the response text, by using a syntax such as this:</p> <pre data-bbox="651 584 906 613">#DTMF.digits"</pre>

4 Messages Sent by Bot

When the Voice.AI Gateway handles messages from the bot, it treats them as activities.

The syntax for sending the activities in the different bot frameworks is described in Section [Bot Framework Specific Details](#) on page 18.

Activities sent by the bot contain actions to be performed and parameters. The parameters can affect the current action or change the behavior of the whole conversation. A list of the configurable parameters are described in Section [Parameters Controlled by Bot](#).

The Voice.AI Gateway handles activities synchronously and therefore, an activity is not executed before the previous one has finished. For example, when the Voice.AI Gateway receives two activities—to play text to the user and to hang up the call—the `hangup` activity is only executed after it has finished playing the text.

Basic Activity Syntax

Each activity is a JSON object that has the following properties:

Table 4-1: Properties of JSON Object Activities

Property	Type	Description
<code>type</code>	String	Either <code>message</code> or <code>event</code> .
<code>name</code>	String	Name of event for the <code>event</code> activity. For supported events, see event Activities on the next page.
<code>text</code>	String	Text to be played for the <code>message</code> activity.
<code>activityParams</code>	Params object	Set of parameters that affect the current activity.
<code>sessionParams</code>	Params object	Set of parameters that affect the remaining duration of the conversation.

The Params object is comprised of key-value pairs, where the key is the parameter name and the value is the desired value for the parameter. For a list of the supported parameters, see [Parameters Controlled by Bot](#).

message Activity

The most common activity is the `message` activity, which indicates to the Voice.AI Gateway to play the given text to the user.

Example:

```
{
  "type": "message",
  "text": "Hi, how may I assist you?"
}
```

A `message` activity can also contain parameters that affect its handling. For example, to disable caching of the text-to-speech generated voice for the current activity, the following activity can be sent:

```
{
  "type": "message",
  "text": "I have something sensitive to tell you.",
  "activityParams": {
    "disableTtsCache": true
  }
}
```

The `text` field can contain Speech Synthesis Markup Language (SSML). The SSML can be one of the following:

- A full SSML document, for example:

```
<speak>
  This is <say-as interpret-as="characters">SSML</say-as>.
</speak>
```

- Text with SSML tags, for example:

```
This is <say-as interpret-as="characters">SSML</say-as>.
```



- The SSML is parsed by the text-to-speech engine. Refer to their documentation for a list of supported features.
- When using SSML, all invalid XML characters, for example, the ampersand (&), must be properly escaped.

event Activities

This section lists the supported events. Each event is shown with a list of associated parameters. These parameters can be set either in the configuration of the bot or by sending them as part of the `activityParams` (to be used once) or as part of the `sessionParams` (to be used for the remaining duration of the conversation).

The list only includes parameters that are specific to that event, but other parameters can also be updated by the event. For example, the `language` parameter can be updated by `playUrl`, by adding it to the `activityParams` or `sessionParams` properties.

hangup

The `hangup` event disconnects the conversation.

The following table lists the parameters associated with this event.

Table 4-2: Parameters for hangup Event

Parameter	Type	Description
<code>hangupReason</code>	String	Conveys a textual reason for hanging up. This reason appears in the CDR of the call.

Example:

```
{
  "type": "event",
  "name": "hangup",
  "activityParams": {
    "hangupReason": "conversationCompleted"
  }
}
```

transfer

The `transfer` event transfers the call to a human agent or to another bot. The `handover` event is a synonym for the `transfer` event.

The following table lists the parameters associated with this event.

Table 4-3: Parameters for transfer Event

Parameter	Type	Description
<code>transferTarget</code>	String	URI to where the call must be transferred call to. Typically, the URI is a "tel" or "sip" URI.
<code>handoverReason</code>	String	Conveys a textual reason for the transfer.
<code>transferSipHeaders</code>	Array of Objects	Array of objects listing SIP headers that should be sent to the transferee. Each object comprises a <code>name</code> and a <code>value</code> attribute. For more information, see Adding SIP Headers on Call Transfer on the next page.

Parameter	Type	Description
<code>transferReferredByURL</code>	String	Defines the party (URL) who initiated the call referral. If this parameter exists, the SBC adds a SIP Referred-By header to the outgoing INVITE or REFER message (according to the 'Remote REFER Mode' parameter). If the SBC handles locally (termination), the SBC adds it to a new outgoing INVITE. If not handled locally (regular), the SBC adds it to the forwarded REFER message.

Example:

```
{
  "type": "event",
  "name": "transfer",
  "activityParams": {
    "handoverReason": "userRequest",
    "transferTarget": "tel:123456789"
    "transferReferredByURL": "sip:456@ac.com",
  }
}
```

Adding SIP Headers on Call Transfer

When the bot performs a call transfer using the `transfer` event, it can add data to be sent as SIP headers in the generated SIP message (REFER or INVITE). This is done by the `transferSipHeaders` parameter. This parameter contains an array of JSON objects with the following attributes:

Table 4-4: Attributes of transferSipHeaders Parameter

Attribute	Type	Description
<code>name</code>	String	Name of the SIP header.
<code>value</code>	String	Value of the SIP header.

For example, the following `transfer` event can be used to add the header "X-My-Header" with the value "my_value":

```
{
  "type": "event",
```

```

"name": "transfer",
"activityParams": {
  "transferTarget": "sip:john@host.com",
  "transferSipHeaders": [
    {
      "name": "X-My-Header",
      "value": "my_value"
    }
  ]
}
}

```

If the Voice.AI Gateway is configured to handle transfer by sending a SIP INVITE message, it will contain the header, for example:

```
X-My-Header: my_value
```

If the Voice.AI Gateway is configured to handle transfer by sending a SIP REFER message, it will contain the value in the URI of the Refer-To header, for example:

```
Refer-To: <sip:john@host.com?X-My-Header=my_value>
```

playUrl

The `playURL` event plays audio to the user from a given URL.



The format of the file must match the format specified by the `playUrlMediaFormat` parameter; otherwise, the audio will be played corruptly.

The following table lists the parameters associated with this event.

Table 4-5: Parameters for playURL Event

Parameter	Type	Description
<code>playUrlUrl</code>	String	URL of where the audio file is located.
<code>playUrlCaching</code>	Boolean	Enables caching of the audio: <ul style="list-style-type: none"> <code>true</code>: Enables caching <code>false</code>: (Default) Disables caching
<code>playUrlMediaFormat</code>	String	Defines the format of the audio: <ul style="list-style-type: none"> <code>wav/lpcm16</code> (default)

Parameter	Type	Description
		■ <code>raw/lpcm16</code>
<code>playUrlAltText</code>	String	Defines the text to display in the transcript page of the user interface while the audio is played.

Example:

```
{
  "type": "event",
  "name": "playUrl",
  "activityParams": {
    "playUrlUrl": "https://example.com/my-file.wav",
    "playUrlMediaFormat": "wav/lpcm16"
  }
}
```

config

The `config` event updates the session parameters, regardless of specific activity.

There are no parameters that are associated with this event.

The following is an example of the `config` event, enabling the Barge-In feature:

```
{
  "type": "event",
  "name": "config",
  "sessionParams": {
    "bargeln": true
  }
}
```

Bot Framework Specific Details

This section provides details specific to bot frameworks.

AudioCodes Bot API

For AudioCodes Bot API, the activities can be sent as is, with the addition of the attributes `id` and `timestamp`, as defined in the [AudioCodes Bot API](#) documentation.

Microsoft Azure

For Azure bots, the `sessionParams` and `activityParams` properties should be placed inside the `channelData` property.

Example:

```
{
  "type": "event",
  "name": "transfer",
  "channelData": {
    "activityParams": {
      "handoverReason": "userRequest",
      "transferTarget": "tel:123456789"
    }
  }
}
```

Google Dialogflow

For Google Dialogflow, the activities are derived from intent's response (the "Default" response, which is the response to PLATFORM_UNSPECIFIED platform).

The response's text is used to construct a `message` activity for playing the text to the user.

To send additional parameters or activities, Custom Payload must be added to the response (see <https://cloud.google.com/dialogflow/docs/intents-rich-messages>).

The Custom Payload can contain a JSON object with the following properties:

Table 4-6: Google Dialogflow Custom Payload Properties

Property	Description
<code>activityParams</code>	This is applied when playing the text of the response (i.e., of the <code>message</code> activity).
<code>sessionParams</code>	This is applied when playing the text of the response (i.e., of the <code>message</code> activity).
<code>activities</code>	Array of activities to be executed after playing the text of the response.

For example, if the text response is "I'm going to transfer you to a human agent" and the Custom Payload contains the following JSON object:

```
{
  "activityParams": {
```

```
"disableTtsCache": true
},
"activities": [
  {
    "type": "event",
    "name": "transfer",
    "activityParams": {
      "transferTarget": "tel:123456789"
    }
  }
]
}
```

Then the audio of the text "I'm going to transfer you to a human agent." is played without caching (due to the `disableTtsCache` parameter). After it has finished playing, the `transfer` activity is executed.

The above example can be configured through the Dialogflow user interface, as follows:

Table 4-7: Custom Payload Configuration Example through Dialogflow User Interface

Text or SSML Response
🔍 🗑️

1 I'm going to transfer you to a live agent.

2 Enter a text or SSML response variant ⬆️

Custom Payload
🔍 🗑️

```

1 {
2   "activityParams": {
3     "disableTtsCache": true
4   },
5   "activities": [
6     {
7       "type": "event",
8       "name": "transfer",
9       "activityParams": {
10        "transferTarget": "tel:123456789"
11      }
12    }
13  ]
14 }
```

Parameters Controlled Also by Bot

These parameters can be configured on the Voice.AI Connector, but they can also be determined and updated by the bot dynamically. The bot takes precedence (i.e., overrides Voice.AI Connector configuration). Parameters that are specific to a single event type are documented in Section [event Activities](#) on page 14. As explained in Section [Basic Activity Syntax](#) on page 13, these parameters can be included in the `activityParams` or the `sessionParams` of any activity sent by the bot.

Table 4-8: Bots Section Parameter Descriptions (Also Controlled by Bot)

Parameter	Type	Description
<code>azureSpeechRecognitionMode</code>	String	Defines the Azure STT recognition mode.

Parameter	Type	Description
		<ul style="list-style-type: none"> ■ <code>conversation</code> (default) ■ <code>dictation</code> ■ <code>interactive</code> <p>Note: The parameter is applicable only to the Microsoft Azure STT service.</p>
<code>bargeIn</code>	Boolean	<p>Enables the Barge-In feature.</p> <ul style="list-style-type: none"> ■ <code>true</code>: Enabled, When the bot is playing a response to the user (playback of bot message), the user can "barge-in" (interrupt) and start speaking. This terminates the bot response, allowing the bot to listen to the new speech input from the user (i.e., Voice.AI Gateway sends detected utterance to the bot). ■ <code>false</code>: (Default) Disabled. The Voice.AI Gateway doesn't expect speech input from the user until the bot has finished playing its response to the user. In other words, the user can't "barge-in" until the bot message response has finished playing.
<code>bargeInOnDTMF</code>	Boolean	<p>Enables the Barge-In on DTMF feature.</p> <ul style="list-style-type: none"> ■ <code>true</code>: (Default) Enabled. When the bot is playing a response to the user (playback of bot message), the user can "barge-in" (interrupt) with a DTMF digit. This terminates the bot response, allowing the bot to listen to and process the digits sent from the user. ■ <code>false</code>: Disabled. The Voice.AI Connector doesn't expect DTMF input from the user until the bot has finished playing its response to the user. In other words, the user can't "barge-in" until the bot message response has finished playing. <p>Note: When the parameter is enabled, you also need to enable <code>sendDTMF</code>.</p>
<code>bargeInMinWordCount</code>	Integer	<p>Defines the minimum number of words that the user must say for the Voice.AI Gateway to consider it a barge-in. For example, if configured to 4 and the</p>

Parameter	Type	Description
		<p>user only says 3 words during the bot's playback response, no barge-in occurs.</p> <p>The valid range is 1 to 5. The default is 1.</p>
<code>botFailOnErrors</code>	Boolean	<p>Defines what happens when the Azure bot error "retry" occurs.</p> <ul style="list-style-type: none"> ■ <code>true</code>: The error is printed to the log and the call is disconnected. ■ <code>false</code>: (Default) The error is printed to the log, but the call is not disconnected.
<code>botNoInputGiveUpTimeoutMS</code>	Integer	<p>Defines the maximum time that the Voice.AI Connector waits for a response from the bot. If no response is received when the timeout expires, the Voice.AI Connector disconnects the call with the SBC.</p> <p>In this scenario (disconnects), you can also configure the Voice.AI Connector to perform specific activities, for example, playing a prompt to the user or transferring the call (see <code>generalFailoverActivities</code> parameter). The default is 0 (i.e., feature disabled).</p> <p>If the call is disconnected, the SIP BYE message sent by the SBC to the user indicates this failure, by prefixing the value in the Reason header with "Bot Err:".</p>
<code>botNoInputRetries</code>	Integer	<p>Defines the maximum number of allowed timeouts (configured by the <code>botNoInputTimeoutMS</code> parameter) for no bot input. If you have configured a prompt to play (see the <code>userNoInputSpeech</code> or <code>userNoInputUrl</code> parameter), the prompt is played to the user each time the timeout expires. The default is 0 (i.e., only one timeout – no retries). For more information on the no bot input feature, see the <code>botNoInputTimeoutMS</code> parameter.</p> <p>Note: If you have configured a prompt to play upon timeout expiry, the timer is triggered only after playing the prompt to the user.</p>
<code>botNoInputSpeech</code>	String	<p>Defines the textual prompt to play to the user when no input has been received from the bot</p>

Parameter	Type	Description
		<p>framework when the timeout expires (configured by <code>botNoInputTimeoutMS</code>).</p> <p>By default, the parameter is not configured.</p> <p>For example:</p> <pre>{ "name": "LondonTube", "provider": "my_azure", "displayName": "London Tube", "botNoInputTimeoutMS": 5000, "botNoInputSpeech": "Please wait for bot input" }</pre> <p>For more information on the no bot input feature, see the <code>botNoInputTimeoutMS</code> parameter.</p> <p>Note: If you have also configured to play an audio prompt (see the <code>botNoInputUrl</code> parameter), the <code>botNoInputSpeech</code> takes precedence.</p>
<code>botNoInputTimeoutMS</code>	Integer	<p>Defines the maximum time (in milliseconds) that the Voice.AI Connector waits for input from the bot framework.</p> <p>If no input is received from the bot when this timeout expires, you can configure the Voice.AI Connector to play a textual (see the <code>botNoInputSpeech</code> parameter) or an audio (see the <code>botNoInputUrl</code> parameter) prompt to the user.</p> <p>The default is 0 (i.e., feature disabled).</p>
<code>botNoInputUrl</code>	String	<p>Defines the URL from where the audio prompt is played to the user when no input has been received from the bot when the timeout expires (configured by <code>botNoInputTimeoutMS</code>).</p> <p>By default, the parameter is not configured.</p> <p>For more information on the no bot input feature, see the <code>botNoInputTimeoutMS</code>.</p> <p>Note: If you have also configured to play a textual prompt (see the <code>botNoInputSpeech</code> parameter), the <code>botNoInputSpeech</code> takes precedence.</p>
<code>continuousASR</code>	Bool	Enables the Continuous ASR feature. Continuous

Parameter	Type	Description
	boolean	<p>ASR enables the Voice.AI Gateway to concatenate multiple STT recognitions of the user and then send them as a single textual message to the bot.</p> <ul style="list-style-type: none"> ■ <code>true</code>: Enabled ■ <code>false</code>: (Default) Disabled <p>For an overview of the Continuous ASR feature, refer to the Product Description.</p>
<code>continuousASRDigits</code>	String	<p>This parameter is applicable when the Continuous ASR feature is enabled.</p> <p>Defines a special DTMF key, which if pressed, causes the Voice.AI Gateway to immediately send the accumulated recognitions of the user to the bot. For example, if configured to "#" and the user presses the pound key (#) on the phone's keypad, the device concatenates the accumulated recognitions and then sends them as one single textual message to the bot.</p> <p>The default is "#".</p> <p>Note: Using this feature incurs an additional delay from the user's perspective because the speech is not sent immediately to the bot after it has been recognized. To overcome this delay, configure the parameter to a value that is appropriate to your environment.</p>
<code>continuousASRTimeoutInMS</code>	Integer	<p>This parameter is applicable when the Continuous ASR feature is enabled.</p> <p>Defines the automatic speech recognition (ASR) timeout (in milliseconds). When the device detects silence from the user for a duration configured by this parameter, it concatenates all the accumulated STT recognitions and sends them as one single textual message to the bot.</p> <p>The valid value is 2,500 (i.e., 2.5 seconds) to 60,000 (i.e., 1 minute). The default is 3,000.</p>
<code>disableTtsCache</code>	Boolean	<p>Enables caching of TTS (audio) results from the bot. Therefore, if the Voice.AI Connector needs to send a request for TTS to a TTS provider and this text has been requested before, it retrieves the result from</p>

Parameter	Type	Description
		<p>its cache instead of requesting it again from the TTS provider.</p> <ul style="list-style-type: none"> ■ <code>true</code>: Enabled ■ <code>false</code>: (Default) Disabled
<code>googleInteractionType</code>	String	<p>Defines the Google STT interaction type. For more information, see https://cloud.google.com/speech-to-text/docs/reference/rest/v1p1beta1/RecognitionConfig#InteractionType.</p>
<code>handoverReason</code>	String	<p>Defines the textual reason when the call is transferred to another party (e.g., another bot or a human agent).</p> <p>By default, the parameter is not defined.</p>
<code>hangupReason</code>	String	<p>Conveys a textual reason for hanging up (disconnecting call). This reason appears in the CDR of the call.</p> <p>Example message:</p> <pre>{ "type": "event", "name": "hangup", "activityParams": { "hangupReason": "conversationCompleted" } }</pre>
<code>language</code>	String	<p>Defines the language (e.g., "en-ZA" for South African English) of the bot conversation and is used for TTS and STT functionality. The value is obtained from the service provider.</p> <ul style="list-style-type: none"> ■ STT: <ul style="list-style-type: none"> ✓ Azure: The parameter is configured with the value from the 'Locale' column in Azure's Speech-Text table (e.g., "en-GB"). ✓ Google: The parameter is configured with the value from the 'languageCode' (BCP-47)

Parameter	Type	Description
		<p>column in Google's Cloud Speech-to-Text table (e.g., "nl-NL").</p> <p>For more information, refer to section Required Information of STT Provider on page 3.</p> <p>■ TTS:</p> <ul style="list-style-type: none"> ✓ Azure: The parameter is configured with the value from the 'Locale' column in Azure's Text-to-Speech table (e.g., "it-IT"). ✓ Google: The parameter is configured with the value from the 'Language code' column in Google's Cloud Text-to-Speech table (e.g., "en-US"). ✓ AWS: The parameter is configured with the value from the 'Language' column in Amazon's Polly TTS table (e.g., "de-DE"). <p>For more information, refer to section Required Information of TTS Provider on page 4.</p> <p>Note: This string is obtained from the TTS or STT service provider by the Customer and must be provided to AudioCodes. For more information, see the Voice.AI Gateway Integration Guide.</p>
<code>playUrlAltText</code>	String	Defines the text to display in the transcript page of the user interface while the audio is played.
<code>playUrlCaching</code>	Boolean	Enables caching of the audio in the TTS cache: <ul style="list-style-type: none"> ■ <code>true</code>: Enables caching ■ <code>false</code>: (Default) Disables caching
<code>playUrlMediaFormat</code>	String	Defines the format of the audio: <ul style="list-style-type: none"> ■ <code>wav/lpcm16</code> (default) ■ <code>raw/lpcm16</code>
<code>playUrlUrl</code>	String	Defines the HTTP-based server by URL where the audio file to be played is located. This allows the play of pre-recorded prompts (audio file) to the user from a remote third-party server.
<code>resumeRecognitionT</code>	Inte	When Barge-In is disabled, speech input is not

Parameter	Type	Description
<code>timeoutMS</code>	Integer	<p>expected before the bot's response has finished playback. If no reply from the bot arrives within this configured timeout (in milliseconds), the Voice.AI Gateway expects speech input from the user and STT recognition is re-activated.</p> <p>The valid value is 0 (i.e., no automatic resumption of recognition) to 600,000 (i.e., 10 minutes). The default is 10,000.</p>
<code>sendDTMF</code>	Boolean	<p>Enables the sending of DTMF events to the bot.</p> <ul style="list-style-type: none"> ■ <code>true</code>: Enabled ■ <code>false</code>: (Default) Disabled
<code>sttContextBoost</code>	Number	<p>Defines the boost number for context recognition of the speech context phrase configured by <code>sttContextPhrases</code>. Speech-adaptation boost allows you to increase the recognition model bias by assigning more weight to some phrases than others. For example, when users say "weather" or "whether", you may want the STT to recognize the word as weather.</p> <p>For more information, see https://cloud.google.com/speech-to-text/docs/context-strength.</p> <p>Note:</p> <ul style="list-style-type: none"> ■ The parameter can be used by all bot providers when the STT engine is Google. ■ When using other STT engines, the parameter has no affect.
<code>sttContextId</code>	String	<p>Defines the STT context. This is used for the DNN server, and as custom context for Azure's STT service.</p>
<code>sttContextPhrases</code>	Array of Strings	<p>When using Google's Cloud STT engine, this parameter controls Speech Context phrases. The parameter can list phrases or words that is passed to the STT engine as "hints" for improving the accuracy of speech recognitions.</p> <p>For more information on speech context (speech adaptation) as well details regarding tokens (class</p>

Parameter	Type	Description
		<p>tokens) that can be used in phrases, go to https://cloud.google.com/speech-to-text/docs/speech-adaptation.</p> <p>For example, whenever a speaker says "weather" frequently, you want the STT engine to transcribe it as "weather" and not "whether". To do this, the parameter can be used to create a context for this word (and other similar phrases associated with weather):</p> <pre>"sttContextPhrases": ["weather"]</pre> <p>Note:</p> <ul style="list-style-type: none"> ■ The parameter can be used by all bot providers when the STT engine is Google. ■ When using other STT engines, the parameter has no affect.
<code>sttDisablePunctuation</code>	Boolean	<p>Prevents the STT response from the bot to include punctuation marks.</p> <ul style="list-style-type: none"> ■ <code>true</code>: Enabled. Punctuation is excluded. ■ <code>false</code>: (Default) Disabled. Punctuation is included. <p>Note: This requires support from the STT engine.</p>
<code>sttEndpointID</code>	String	A synonym for the <code>sttContextId</code> parameter.
<code>transferReferredByURL</code>	String	<p>Defines the party (URL) who initiated the referral. If this parameter exists, the SBC adds a SIP Referred-By header to the outgoing INVITE/REFER message (according to the 'Remote REFER Mode' parameter). If the SBC handles locally (termination), the SBC adds it to a new outgoing INVITE. If not handled locally (regular), the SBC adds it to the forwarded REFER message.</p>
<code>transferSipHeaders</code>	Array of Objects	<p>Array of objects listing SIP headers that should be sent to the transferee. Each object comprises a <code>name</code> and a <code>value</code> attribute.</p>

Parameter	Type	Description
<code>transferTarget</code>	String	Defines the URI to where the call must be transferred. Typically, the URI is a "tel" or "sip" URI.
<code>userNoInputRetries</code>	Integer	<p>Defines the maximum number of allowed timeouts (configured by the <code>userNoInputTimeoutMS</code> parameter) for no user input. If you have configured a prompt to play (see the <code>userNoInputSpeech</code> or <code>userNoInputUrl</code> parameter), the prompt is played each time the timeout expires.</p> <p>The default is 0 (i.e., only one timeout).</p> <p>For more information on the no user input feature, see the <code>userNoInputTimeoutMS</code> parameter.</p> <p>Note: If you have configured a prompt to play upon timeout expiry, the timer is triggered only after playing the prompt to the user.</p>
<code>userNoInputSendEvent</code>	Boolean	<p>Enables the Voice.AI Connector to send an event message to the bot if there is no user input for the duration configured by the <code>userNoInputTimeoutMS</code> parameter, indicating how many times the timer has expired ('value' field):</p> <pre>{ "type": "event", "name": "noUserInput", "value": 1 }</pre> <ul style="list-style-type: none"> ■ <code>true</code>: Enabled. ■ <code>false</code>: (Default) Disabled.
<code>userNoInputSpeech</code>	String	<p>Defines the textual prompt to play to the user when no input has been received from the user when the timeout expires (configured by <code>userNoInputTimeoutMS</code>).</p> <p>By default, the parameter is not configured.</p> <p>For example:</p> <pre>{</pre>

Parameter	Type	Description
		<pre data-bbox="770 271 1374 600"> { "name": "LondonTube", "provider": "my_azure", "displayName": "London Tube", "userNoInputTimeoutMS": 5000, "userNoInputSpeech": "Hi there. Please say something" } </pre> <p data-bbox="775 613 1390 685">For more information on the no user input feature, see the <code>userNoInputTimeoutMS</code>.</p> <p data-bbox="775 698 1378 808">Note: If you have also configured to play an audio prompt (see the <code>userNoInputUrl</code> parameter), the <code>userNoInputSpeech</code> takes precedence.</p>
<code>userNoInputTimeoutMS</code>	Integer	<p data-bbox="775 853 1355 960">Defines the maximum time (in milliseconds) that the Voice.AI Connector waits for input from the user.</p> <p data-bbox="775 976 1382 1364">If no input is received when this timeout expires, you can configure the Voice.AI Connector to play a textual (see the <code>userNoInputSpeech</code> parameter) or an audio (see the <code>userNoInputUrl</code> parameter) prompt to ask the user to say something. If there is still no input from the user, you can configure the Voice.AI Connector to prompt the user again. The number of times to prompt is configured by the <code>userNoInputRetries</code> parameter.</p> <p data-bbox="775 1379 1382 1610">If the <code>userNoInputSendEvent</code> parameter is configured to <code>true</code> and the timeout expires, the Voice.AI Connector sends the following event to the bot (supported only by Azure, Google, and AC-bot-API), indicating how many times the timer has expired:</p> <pre data-bbox="770 1624 1374 1877"> { "type": "event", "name": "noUserInput", "value": 1 } </pre> <p data-bbox="775 1890 1230 1921">The default is 0 (i.e., feature disabled).</p> <p data-bbox="775 1937 847 1968">Note:</p>

Parameter	Type	Description
		<ul style="list-style-type: none"> ■ DTMF (any input) is considered as user input (in addition to user speech) if the <code>sendDTMF</code> parameter is configured to <code>true</code>. ■ If you have configured a prompt to play when the timeout expires, the timer is triggered only after playing the prompt to the user.
<code>userNoInputUrl</code>	String	<p>Defines the URL from where the audio prompt is played to the user when no input has been received from the user when the timeout expires (configured by <code>userNoInputTimeoutMS</code>). By default, the parameter is not configured. For more information on the no user input feature, see the <code>userNoInputTimeoutMS</code>.</p> <p>Note: If you have also configured to play a textual prompt (see the <code>userNoInputSpeech</code> parameter), the <code>userNoInputSpeech</code> takes precedence.</p>
<code>voiceName</code>	String	<p>Defines the voice name for the TTS service.</p> <ul style="list-style-type: none"> ■ Azure: The parameter is configured with the value from the 'Short voice name' column in Azure's Text-to-Speech table (e.g., "it-IT-ElsaNeural"). ■ Google: The parameter is configured with the value from the 'Voice name' column in Google's Cloud Text-to-Speech table (e.g., "en-US-Wavenet-A"). ■ AWS: The parameter is configured with the value from the 'Name/ID' column in Amazon's Polly TTS table (e.g., "Hans"). ■ Almagu: The parameter is configured with the value from the 'Voice' column in Almagu's TTS table (e.g., "Osnat"). <p>Note: This string is obtained from the TTS service provider by the Customer and must be provided to AudioCodes. For more information, refer to Section Required Information of TTS Provider on page 4.</p>

This page is intentionally left blank.

International Headquarters

1 Hayarden Street,

Airport City

Lod 7019900, Israel

Tel: +972-3-976-4000

Fax: +972-3-976-4040

AudioCodes Inc.

200 Cottontail Lane

Suite A101E

Somerset NJ 08873

Tel: +1-732-469-0880

Fax: +1-732-469-2298

Contact us: <https://www.audiocodes.com/corporate/offices-worldwide>

Website: <https://www.audiocodes.com/>

©2020 AudioCodes Ltd. All rights reserved. AudioCodes, AC, HD VoIP, HD VoIP Sounds Better, IPmedia, Mediant, MediaPack, What's Inside Matters, OSN, SmartTAP, User Management Pack, VMAS, VoIPerfect, VoIPerfectHD, Your Gateway To VoIP, 3GX, VocaNom, AudioCodes One Voice, AudioCodes Meeting Insights, AudioCodes Room Experience and CloudBond are trademarks or registered trademarks of AudioCodes Limited. All other products or trademarks are property of their respective owners. Product specifications are subject to change without notice.

Document #: LTRT-30927

