

# Technical Disclosure Commons

---

Defensive Publications Series

---

December 2021

## MULTILINGUAL VOICE SUPPORT FOR CONTACT CENTER AGENTS

Ananth Hegde

Amit Agnihotri

Nivedha Thangaraju

Follow this and additional works at: [https://www.tdcommons.org/dpubs\\_series](https://www.tdcommons.org/dpubs_series)

---

### Recommended Citation

Hegde, Ananth; Agnihotri, Amit; and Thangaraju, Nivedha, "MULTILINGUAL VOICE SUPPORT FOR CONTACT CENTER AGENTS", Technical Disclosure Commons, (December 28, 2021)  
[https://www.tdcommons.org/dpubs\\_series/4817](https://www.tdcommons.org/dpubs_series/4817)



This work is licensed under a [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/).

This Article is brought to you for free and open access by Technical Disclosure Commons. It has been accepted for inclusion in Defensive Publications Series by an authorized administrator of Technical Disclosure Commons.

## MULTILINGUAL VOICE SUPPORT FOR CONTACT CENTER AGENTS

### AUTHORS:

Ananth Hegde  
Amit Agnihotri  
Nivedha Thangaraju

### ABSTRACT

In a contact center, when a caller elects to speak to an agent the call will be routed (based on the language that the caller chose) to an agent who can speak the language that was selected by the caller. A global company with operations or offices in many different countries faces the challenge of communicating in multiple languages, something that is a key to providing excellent customer service. In particular, it can be challenging for a company to achieve such a capability when they simply cannot hire a customer support agent for every possible language. To address such challenges, techniques are presented herein that support eliminating the language barrier between agents and callers, providing a single contact center expert catering to customer needs across the globe without any region and locale barriers. The presented techniques support extracting a caller language code (during both traditional interactive voice response (IVR) and conversational IVR phases) and using same to eliminate the language barrier between a caller and an agent by integrating a language translation service into contact center products. Aspects of the presented techniques leverage the Session Initiation Protocol (SIP) and the Real-time Transport Protocol (RTP). The presented techniques offer a number of benefits, including providing a cost reduction to a contact center, reducing the need to hire multilingual agents, etc.

### DETAILED DESCRIPTION

In a contact center, when a caller elects to speak to an agent the call will be routed (based on the language that the caller chose) to an agent who can speak the language that was selected by the caller.

A global company with operations or offices in many different countries faces the challenge of communicating in multiple languages, something that is a key to providing excellent customer service. In particular, it can be challenging for a company to achieve

such a capability when they simply cannot hire a customer support agent for every possible language.

To address the types of challenges that were described above, techniques are presented herein that support an automatic language translator for an agent, through which the language barrier for an agent may be eliminated and one agent can serve multilingual callers.

Aspects of the techniques presented herein may be explicated with reference to an illustrative scenario. Such a scenario will be described below.

When a caller connects to a contact center, they proceed through interactive voice response (IVR) treatment. When the caller requests to be transferred to an agent, a Session Initiation Protocol (SIP) connection will be established between the caller's telephone and the agent's desktop. Through that connection will be passed the caller's language code (e.g., as extracted through the IVR process) in a SIP header. The agent's telephone will store the customer's preferred language code (as sent in the SIP header) and when the language of an agent and the language of a caller do not match a language translator may be enabled by the agent on the agent's desktop. Upon enabling such an option, an agent desktop may connect to a translation service and receive a translated message as voice, on the agent's desktop and in their preferred language, so that the agent can hear and understand the customer's query. When an agent speaks in their own language, the agent's voice may be intercepted and then translated into the customer's preferred language before any Real-time Transport Protocol (RTP) packets are transmitted to the caller.

Aspects of the techniques presented herein, as explicated by the illustrative scenario that was presented above, support different use cases including, for example, an agent's softphone supporting translation service integration, a mobile agent supporting an application for performing language translation, etc.

According to aspects of the techniques presented herein, a caller's language may be extracted through IVR interaction before a call is handed over to an agent. If a caller is joining a virtual agent (in IVR) then an available recognition application programming interface (API) may support language detection. If the caller is going through traditional IVR then by collecting dual-tone multi-frequency (DTMF) input the caller's preferred language may be extracted.

Figure 1, below, presents elements of an exemplary live language translator integration for the voice portion of a contact center agent according to aspects of the techniques presented herein.

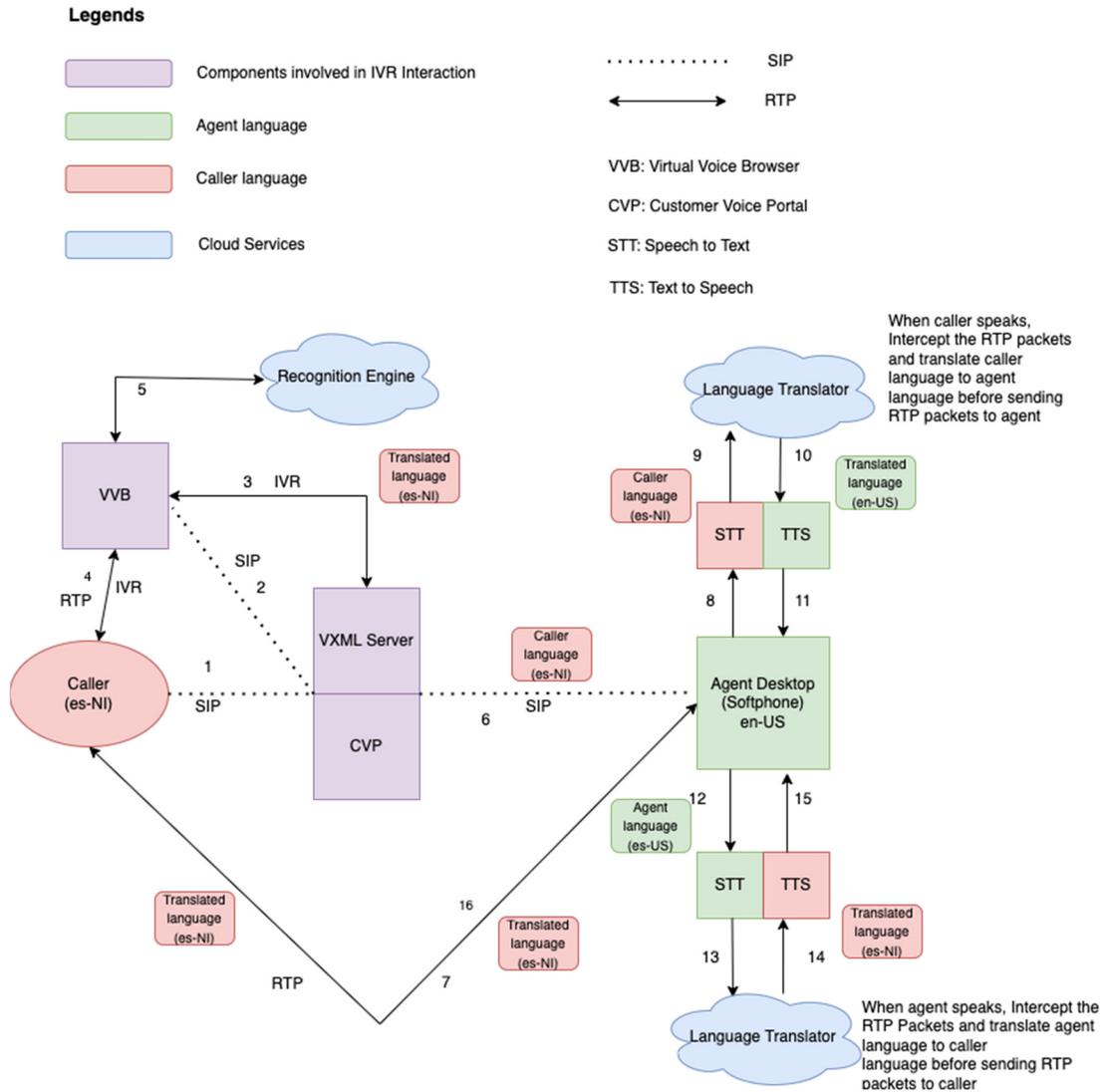


Figure 1: Exemplary Live Language Translator Integration

The exemplary integration that is depicted in Figure 1, above, identifies a series of steps. Those steps, which are labeled 1 through 15 in the figure, will be described below.

Step 1 encompasses the establishment of SIP signaling from a Caller to a vendor's Customer Voice Portal (CVP). Next, Step 2 encompasses the establishment of SIP signaling from the CVP to a vendor's Virtual Voice Browser (VVB). Under Step 3, the

VVB communicates with a Voice Extensible Markup Language (VXML) IVR facility to execute an IVR flow from the VXML server. Thereafter, Step 4 encompasses an exchange of RTP packets between the Caller and the VVB (comprising the actual IVR communication between the Caller and the VVB).

Under Step 5, the Caller's language is detected by sending RTP samples to a Recognizer Engine. Next, Step 6 encompasses the establishment of SIP signaling from the CVP to an Agent Desktop (comprising a transfer of the call from the Caller to the Agent and establishing RTP connection following IVR treatment). Identification of the Caller's language (as detected during the IVR phase as described in Step 5, above) is sent in a SIP header to the Agent to enable language translation.

A connection can be established to a Speech to Text (STT) module in support of the conversion of voice from the Caller to a text format, as shown in Step 8. During Step 8, RTP received from caller in step 7 is converted to text by the STT module is translated into the language that is preferred by the Agent and the translated language is converted into voice, as shown in Step 10.

As shown at Step 11, the translated Caller's message is received, as voice, by the Agent and a connection to a STT module is established at Step 12 in order to convert the Agent's voice to a text format.

During Step 13, the Agent's message (in a text format) is translated into the language that is preferred by the Caller and the translated text is converted into voice (i.e., speech), as shown in Step 14. Thereafter, the translated voice is received by the Agent Desktop, as shown at Step 15, and an exchange of RTP packets between the Caller and the Agent in their preferred language can be performed, as shown in Step 15 and Step 16.

It is important to note that some existing solutions attempt to address elements of the multilingual challenges that were described above through the use of chat and email. In contrast, aspects of the techniques presented herein solve the challenges through the use of voice during a live interaction by employing SIP and RTP.

In summary, techniques have been presented herein that support eliminating the language barrier between agents and callers, providing a single contact center expert catering to customer needs across the globe without any region and locale barriers. The presented techniques support extracting a caller language code (during both traditional IVR

and conversational IVR phases) and using same to eliminate the language barrier between a caller and an agent by integrating a language translation service into contact center products. Aspects of the presented techniques leverage SIP and RTP. The presented techniques offer a number of benefits, including providing a cost reduction to a contact center, reducing the need to hire multilingual agents, etc.