

Technical Disclosure Commons

Defensive Publications Series

November 2020

Delivering Virtual Assistant Responses in a User-specified Alternate Language

Victor Carbune

Matthew Sharifi

Follow this and additional works at: https://www.tdcommons.org/dpubs_series

Recommended Citation

Carbune, Victor and Sharifi, Matthew, "Delivering Virtual Assistant Responses in a User-specified Alternate Language", Technical Disclosure Commons, (November 05, 2020)

https://www.tdcommons.org/dpubs_series/3745



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.

Delivering Virtual Assistant Responses in a User-specified Alternate Language

ABSTRACT

Many households have multiple smart devices, with each device being tagged or named to indicate the room in which it is situated. Users typically interact with smart devices via voice-based virtual assistant software that provides voice responses to spoken queries, e.g., in the same language as that of the query. This disclosure describes techniques that enable users to configure their smart devices such that the smart device always responds in a fixed language, independent of the language in which input commands are provided. Such an operation allows users to receive virtual assistant responses in a language of their choice. Similarly, users are provided with options to specify a fixed region that is applicable regardless of the region in which the device is actually located. Automatic query and/or response translation is performed to provide the responses.

KEYWORDS

- Virtual assistant
- Multilingual household
- Response personalization
- Virtual assistant language
- Output language
- Language learning
- Smart speaker
- Smart display

BACKGROUND

Many households have multiple smart speakers, smart displays, and/or other smart devices in different rooms. Typically, a separate smart speaker or other device that can respond to spoken commands is placed in each room within the home, such as the living room, kitchen, bedroom, etc. Often, each smart speaker is tagged or named to indicate the room in which it is situated.

People typically interact with smart speakers (and other such devices) via voice-based virtual assistant software provided via the device that is configured to receive spoken commands and, in turn, receive audio responses to their queries. The response from the virtual assistant to a spoken command in a particular language is typically delivered as audio in the same language as the spoken command. Some households are multilingual, with household members speaking multiple languages. Further, individuals within the household may be learning an additional language, such as children taking foreign language classes at school.

DESCRIPTION

There are certain situations in which it is useful for smart speakers or other devices in a household to have a notion of language and region stickiness. With such stickiness, the smart speaker responds in a fixed language (e.g., a preconfigured default) irrespective of the language in which a command is provided, via speech or another modality. Such a mode can enable language learners to learn the fixed language (since the smart speaker always responds in that language), or other users to explore “local answers” specific to the region, even when the users (and their smart speaker) are physically in a different region.

This disclosure describes techniques that enable users to configure their smart speakers (and other devices) such that the device always responds in a fixed language, independent of the language in which input commands are provided. Such an operation allows users to receive virtual assistant responses in a language of their choice. Similarly, users are provided with options to specify a fixed region that is applicable regardless of the region in which the device is actually located. For instance, a user can set the region specific to a city or other region setting that is independent of the language setting.

User specification of the output language and/or region for the virtual assistant responses is used for on-the-fly translation of the user input and/or virtual assistant output. If the query is executed in the same language as the user's input, the responses are typically obtained in the same language as the input language. In this case, the responses are translated to the fixed preconfigured output language before being delivered to the user. Results that are locally relevant to the region specified by the user may be obtained by translating the user's query from the original language to the language associated with the specified region and executing the query in the regional language to obtain and provide a regionally relevant response in the user-specified output language.

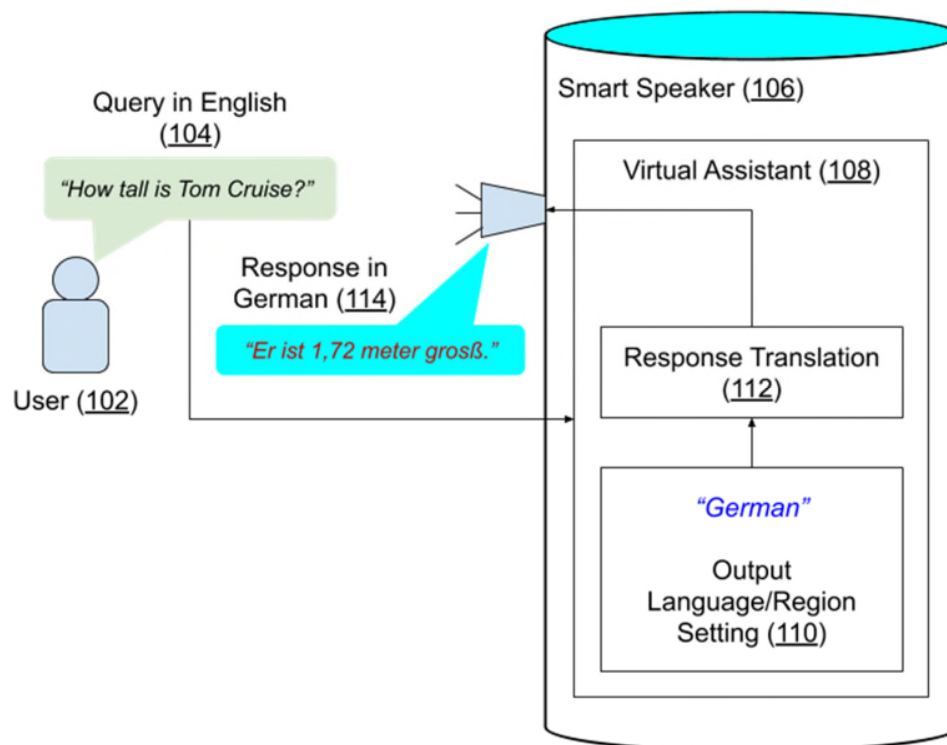


Fig. 1: Providing a virtual assistant response in a user-specified language

Fig. 1 shows an operational implementation of the techniques described in this disclosure.

A user (102) specifies "German" as the output language/region setting (110) for a virtual

assistant (108) provided via a smart speaker (106). When the user issues a voice query in English (104) to the virtual assistant, the answer is determined (including optional information retrieval from online sources). The answer (which may be in English, due to the query being in English) is translated (112) to the user-specified language (German) and the translated German answer is delivered as the response (114). The smart speaker may provide a visual indication of the default language, e.g., in the example of Fig. 1, the top of the smart speaker is shown in blue (e.g., which may be a light built-in to the device) that indicates that this device speaks “German.”

Implementation of the described techniques that enable a smart speaker to have language/region stickiness can support a variety of playful, exploratory, and useful virtual assistant interactions in a household setting. Some examples of such interactions are provided below.

1. Obtaining answers directly translated in another language.
 - English input: *“How tall is Tom Cruise?”*
 - German output: *“Er ist 1,72 meter groß.”*
2. Implicit translation based on answers in another language.
 - English input: *“How do I say chair?”*
 - German output: *“Stuhl.”*
3. Obtaining information related to the region connected to the output language/region (set to Romanian/Bucharest in this example).
 - English input: *“What’s the lake temperature?”*
 - Romanian output: *“Lacul Herăstrău are 20 degrees.”* (providing information about the most likely lake reference for the Bucharest region in Romania.)
4. Marking devices based on output language.

- English input: *“Mark all smart speakers set to German output with green lights.”*
- Output: Light indicators on all smart speakers that provide answers in German are turned green.

Users can set the desired output language and/or the region for the virtual assistant via any suitable mechanism such as voice commands (e.g., “Set kitchen speaker to Romanian”), a graphical user interface (GUI) (e.g., available for a smart display), settings configuration files, etc. With user permission, the virtual assistant responses can be personalized based on the region set by the users, thus allowing the user to obtain local answers relevant to the region as if the user were physically located in that region. Such user-permitted simulated localization can enable localization/ translation features available in the user’s specified region even when these are otherwise disabled or unavailable in the region where the device is located. The translation can be performed on-the-fly, for either or both the input query (e.g., to translate the query into target language and obtain local results in the associated region) as well as the answer. For instance, if a user located in Zurich sets their virtual assistant to the Bucharest region and to provide answers in Romanian, the region settings for Bucharest are applied instead of those for Zurich.

Certain smart devices include broadcasting features, e.g., “Dinner is ready.” If users broadcast a message to speakers configured with different languages, the message is automatically translated to the device-specific language and is provided as spoken output.

When two devices in close proximity, such as two smart speakers in adjacent areas of a home, are configured to different output languages and/or regions, users can enable coordination between the devices to provide answers to the same query in a complementary manner. For instance, one smart speaker can provide the answer to the user’s query in German followed by the other smart speaker that provides the same answer in Romanian prefaced by “Here’s how

you'd say this in Romanian.” In some cases, differences in user-specified region settings can result in different answers to the same user query, such as “What is the country’s population?” In such cases, the answer provided by the smart speaker can be augmented to include the relevant regional context. For instance, the answers to the user’s query can be “În România, populația este de 19,3 milioane.”

The techniques can be implemented on any device that provides virtual assistant capability, such as smart speakers, smartphones, tablets, etc. Further, the techniques can support coordinated operation across multiple devices present within a physical environment, such as a home. A default language or region setting can be helpful to users with language learning, to set particular language preferences for particular devices can help users in multilingual households always receive answers in their preferred language, etc.

Further to the descriptions above, a user may be provided with controls allowing the user to make an election as to both if and when systems, programs or features described herein may enable collection of user information (e.g., information about a user’s language/region preferences, spoken queries or commands, a user’s devices, or a user’s current location), and if the user is sent content or communications from a server. In addition, certain data may be treated in one or more ways before it is stored or used, so that personally identifiable information is removed. For example, a user’s identity may be treated so that no personally identifiable information can be determined for the user, or a user’s geographic location may be generalized where location information is obtained (such as to a city, ZIP code, or state level), so that a particular location of a user cannot be determined. Thus, the user may have control over what information is collected about the user, how that information is used, and what information is provided to the user.

CONCLUSION

This disclosure describes techniques that enable users to configure their smart speakers (and other devices) such that the device always responds in a fixed language, independent of the language in which input commands are provided. Such an operation allows users to receive virtual assistant responses in a language of their choice. Similarly, users are provided with options to specify a fixed region that is applicable regardless of the region in which the device is actually located.