

# Technical Disclosure Commons

---

Defensive Publications Series

---

July 08, 2019

## VOICE UI FOR MULTIPLE AI AGENTS

HP INC

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

---

### Recommended Citation

INC, HP, "VOICE UI FOR MULTIPLE AI AGENTS", Technical Disclosure Commons, (July 08, 2019)  
[https://www.tdcommons.org/dpubs\\_series/2330](https://www.tdcommons.org/dpubs_series/2330)



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.

## **Voice UI for multiple AI agents**

**Situation: Multiple voice agents on a PC device**

**Idea: Multiple voice agents with a single UI that will standardize the user experience.**

**Hypothesis: User will use voice agent(s) to accomplish tasks on a PC device.**

### **Abstract of the invention**

We are standardizing UI for multiple voice agents on a PC device with a lightweight Voice App which will run as a background service on all consumer and commercial PC devices.

This idea was brainstormed and thought through in-depth to address a critical pain point for users interacting with multiple voice agents. This idea was also built into a demo prototype and validated with 12 users in a 1:1 interview format. From our assumptions and subsequent research, we learnt that users will need to learn each voice agent behavior and response within their specific app. For example, Alexa will respond to a user command from inside the Alexa app with a specific audio and visual theme. Xaowei will respond to a user command from a separate Xaowei app with a completely different audio and visual theme. And Google Assistant and Cortana will also have their own separate apps as well as separate visual and audio themes. This will place a lot of cognitive load on the user. So, while interacting with multiple voice agents, user will be confronted with many different combinations of feedback locations, visual and audio themes. Trying to comprehend and memorize each VA's feedback system can be daunting, and these challenges may prevent user from having a smooth and frictionless experience with voice agents.

To address the above user pain point, we created a standardized UI and feedback system, which is an on-screen UI, with an animated feedback bar. The user may command any one of the multiple VA's on a PC, but the feedback will be consistent and standardized from a visual and audio perspective and will live within a Voice app and will feel like existing at the top level, almost at an OS level. Because of this standardization of the UI and bringing it to the top, the user does not have to think who they are interacting with, but they can just be focused on the outcome, the end state. So we reduce the user's cognitive load and enable a frictionless experience with voice agents.

This standardization of multiple voice agents will also extend into configurations and settings. User will be able to manage/configure multiple voice agents on the PC device in one centralized place. In this global settings space in HP Voice app, user will, for example, be able to globally change configurations for all VA's with a single click, with a single checkbox or a single toggle or switch functionality.

## **What problems we are solving for the user:**

1. User pain point 1: Interacting with multiple voice agents that are available and having to learn the way each VA responds on the PC places a lot of cognitive load on the user. Trying to comprehend each VA's feedback system can be confusing.
2. User pain point 2: Opening 4 different VA apps to configure voice agents. (even if the user wants to apply the same configuration for the all VA's.)
3. User pain point 3: Using voice agent on a PC in conference rooms or noisy environments. On screen animated feedback bar will draw user's attention.

## **What problems did we overcome to solve the problems for the user**

### **1. Solving a dual problem for multiple voice agents on PC and multi-language voice agents on PC.**

This dual situation adds complexity to standardizing not just across multiple VA's but across multiple-language VA's as well. At the end we created a framework that supports both, multiple VA's and multiple-language VA's.

### **2. Complexity around creating a hardware LED feedback system on HP PC's or devices**

We iterated on a hardware LED feedback system, but that was a challenge because it would be difficult to accommodate across multiple different assistant requirements on a hardware level, as well as in select devices (like premium Notebook PC's) we don't have space on the device itself to add more LEDs. So, we concluded that software (on screen) is the best way to provide a flexible, but also consistent experience across the PC portfolio.

**3. Diverse branding:** Each VA has a different visual and audio branding and users have an emotional association, preference and comfort level with specific brands. This includes light rings on a physical device (eg: Alexa) or LED lights with dots and colors (eg: Google Assistant) and of course the wake word, the voice and other audio effects. We resolved this complexity by giving users a wide range of feedback bar choices to select from, that included color and audio theme from their favorite VA's that they are familiar with and comfortable with.

**4. Windows desktop clutter:** Trying to help user navigate multiple VA feedback in the midst of all the visual, audio and input clutter of a Windows PC desktop. The feedback had to be visually prominent and visible to the user. We wanted to create something unique that will rise above the noise and notifications of Windows OS.

**5. Windows OS:** Overcome constraints of the Windows operating system where we cannot use the taskbar to present any feedback animations. We had to improvise, so that we were not interfering with any of the Windows OS framework. We also had to work around Windows OS constraints.

**6. In commercial PC scenarios, most particularly in conference room scenario, VA audio is often turned off,** so we cannot completely rely on VA audio feedback. In this situation, on-screen animated feedback bar will draw user's attention.

7. Research has indicated that users get confused between muting while on a PC conference call (device mute) vs. muting VA's on a PC. By creating a prominent and consistent 'Stop listening' state for multiple VA's we helped differentiate between VA 'stop listening' state and device mute state and therefore removed confusion.

8. Security issues and need for global 'stop listening' state: We have learnt through research that tech security can be a source of anxiety and uncertainty for PC customers in general and HP customers as well. They want clear communication on how and when they are secure. A physical button/key to turn VA off to 'Stop listening' will give our customers confidence that the device supports their privacy needs. There is a constant anxiety that the VA's are constantly listening for the prompt or spying.

## **Description, innovation goals and how they tie in with user motivation**

### **Primary innovation**

**Primary innovation goal:** To create a unified consistent feedback system for a user who is commanding multiple voice agents that are installed on their PC device.

**User intent:**

User is ready to command a voice agent... for example, to launch an upcoming conference call

**User Motivation:** Using voice command helps user reduce a number of steps to complete a task with hands-free or far-field interaction.

**User pain point:** Interacting with multiple voice agents and having to learn the way each VA responds (inside VA's specific app), places a lot of cognitive load on the user. Trying to comprehend and memorize each VA's feedback system can be daunting, and these challenges may prevent user from having a smooth and frictionless experience with voice agents.

**How are we solving the problem:** By creating a standardized feedback system, which is an on-screen UI, with an animated feedback bar. The user may command any one of the multiple VA's on a PC but the feedback will be consistent and standardized from a visual and audio perspective. Because of this standardization of the UI, the user does not have to think who they are interacting with, but they can just be focused on the outcome, the end state. So we reduce the user's cognitive load and enhance their voice agent experience.

### **Secondary innovation**

**Secondary innovation goal:** Manage/configure multiple voice agents on the PC device in one centralized place.

**Users intent:** Change a configuration for all the voice agents on PC

**Pain point:** Opening 4 different VA apps to configure voice agents. (even if the user wants to apply the same configuration for the all VA's.)

**User Motivation:** Easily configure all VA's

**How are we solving the problem:** We offer a global settings space in the HP Voice app for all VA's... which will have, for example, a single checkbox or toggle function to globally change configurations for all VA's with a single click.

## **How this innovation will work in the Voice App**

Both the above ideas/functionalities will be integrated into a Voice App, which is a light weight app service running in the background.

What this lightweight app service can do:

1. Voice app (background service) will start when user signs in to PC [Default ON, can be turned off by user]
2. Voice app (background service) can be accessed from All apps menu and System tray
3. Voice app (background service) will ping a voice agent when appropriate wake word is matched
4. Voice app will provide prominent standardized visual feedback when user commands a voice agent that will stand out from the visual clutter of PC desktop screen.
5. Voice app will provide some settings/controls/configurations, so user can customize visual feedback bar location and color theme based on their preferred voice agent theme.
6. Voice app will enable user to change multiple VA configurations with a single toggle or switch action.

## **Our innovation vs. past innovations**

To our knowledge there are no similar inventions in this space.

What exists today is that each device with voice agent has a dedicated voice agent... for example smart speakers have their own proprietary voice agents. There are not many devices that have 3 or 4 voice agents installed on the device giving user a choice of talking to multiple voice agents. Therefore, the very premise is new and uncharted. We did not have any past references or examples or benchmarks to work from. There was no opportunity for any competitor analysis or study. We innovated from the ground up.

We did not find a PC device with multiple VA's and additionally with multi-language VA's.

*Disclosed by Deborani Duttagupta and Jae Woo Chung, HP Inc.*