

---

# Tell Me More: Understanding User Interaction of Smart Speaker News Powered by Conversational Search

**Hyunhoon Jung**  
Naver Clova AI  
hyunhoon.j@navercorp.com

**Gilhwan Hwang**  
Seoul National University  
hwang.gilhwan@snu.ac.kr

**Joonhwan Lee**  
Seoul National University  
joonhwan@snu.ac.kr

**Changhoon Oh**  
Seoul National University  
yurial@snu.ac.kr

**Cindy Yoonjung Oh**  
Seoul National University  
yoonjungoh16@snu.ac.kr

**Bongwon Suh**  
Seoul National University  
bongwon@snu.ac.kr

## ABSTRACT

In this study, we apply a “conversational search” to the design of the news service of smart speakers so that users can actively get richer information while listening to the news. We designed a research prototype called “Anchor,” where a smart speaker news assistant provides users with news about specific keywords and responds to users’ questions. We recruited 21 participants and conducted a user study where they consumed the news with Anchor, followed by post hoc interviews. The results of the qualitative analysis revealed the following. (1) People preferred interactive news to news briefings. (2) People found it useful to get answers on their questions by talking with the assistant. (3) Although

---

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

*CHI’19 Extended Abstracts, May 4–9, 2019, Glasgow, Scotland Uk*

© 2019 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-5971-9/19/05.

<https://doi.org/10.1145/3290607.3312979>

users were allowed to ask questions anytime, they often hesitated, as they did not want to miss the whole flow of the news. (4) However, they had difficulty recalling the questions they had not asked. Based on these findings, we discuss the implications for news design in a voice-only user interface.

### CCS CONCEPTS

• **Human-centered computing** → **Human computer interaction (HCI)**.

### KEYWORDS

Smart speaker news; conversational search; voice user interaction

### ACM Reference Format:

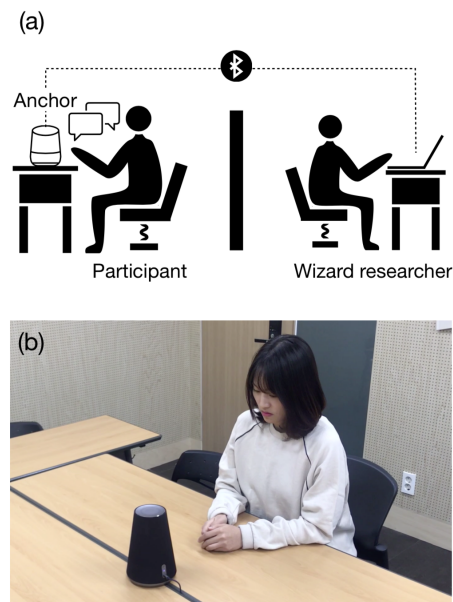
Hyunhoon Jung, Changhoon Oh, Gilhwan Hwang, Cindy Yoonjung Oh, Joonhwan Lee, and Bongwon Suh. 2019. Tell Me More: Understanding User Interaction of Smart Speaker News Powered by Conversational Search. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts (CHI'19 Extended Abstracts)*, May 4–9, 2019, Glasgow, Scotland Uk. ACM, New York, NY, USA, 6 pages. <https://doi.org/10.1145/3290607.3312979>

### INTRODUCTION

Smart speakers, such as Amazon Echo and Google Home, are increasingly being used in daily life. Although some devices have a screen, in most cases, users have to interact with the systems' agents with voice only. In the IoT environment, voice agents are expected to be applied to various home appliances, and users' information activity through voice interaction is expected to increase [3].

News media is also being consumed in new ways through smart speakers. Many smart speakers provide users with news briefings or play news presented by existing media. Most of these, however, merely convey the news headlines or deliver the news unilaterally to users. People cannot consume news that includes in-depth content. Existing news media does not take into account the various possibilities that smart speakers can provide to users, such as conversations between the user and the agent and quick access to the information the user desires. Users tend to lower their expectations and trust of agents when they encounter a problem interacting with the agents of smart speakers [2]. Therefore, it is necessary to create a news service design that is more suitable for the voice interaction environment.

In this study, we aimed to introduce a conversational search [4] into the design of a news service using smart speakers and explore its possibilities. In a conversational search, the user talks to the device using a sentence, and the device responds with a complete human-like sentence. It not only grasps the intent of the user's question but also gives immediate feedback to the user's question. We have applied this to the news design of smart speakers, enabling both users and intelligent agents to communicate with each other and allowing users to answer their questions related to the news.



**Figure 1:** (a) The system consisted of a smart speaker (Anchor), a researcher’s laptop computer, and a wizard (a researcher who controls the computer). The speaker was placed on a table in the middle of the lab, and the wizard and the laptop were located in a separate room that was hidden from users’ view. The researcher’s laptop stored the audio files of the corpus, which were played through the smart speaker connected to the computer via Bluetooth with the wizard’s operation. (b) A photo of the actual experiment. We used Naver “WAVE” (<https://clova.ai/ko/ko-product-wave.html>) as the experimental smart speaker.

In what follows, we present our research prototype “Anchor,” a smart speaker news system based on the Wizard-of-Oz method with a conversational search, and its user study with 21 participants. Finally, we report the results of the experiment, followed by design implications and avenues for future work.

## RESEARCH PROTOTYPE DESIGN

In order to explore the possibility of a smart speaker-based news system capable of a conversational search, we needed to design a research prototype that users could experience without hindrances. Accordingly, we designed a research prototype “Anchor” operated by the Wizard-of-Oz method.

### Selecting news articles and preparing expected questions and answer corpus

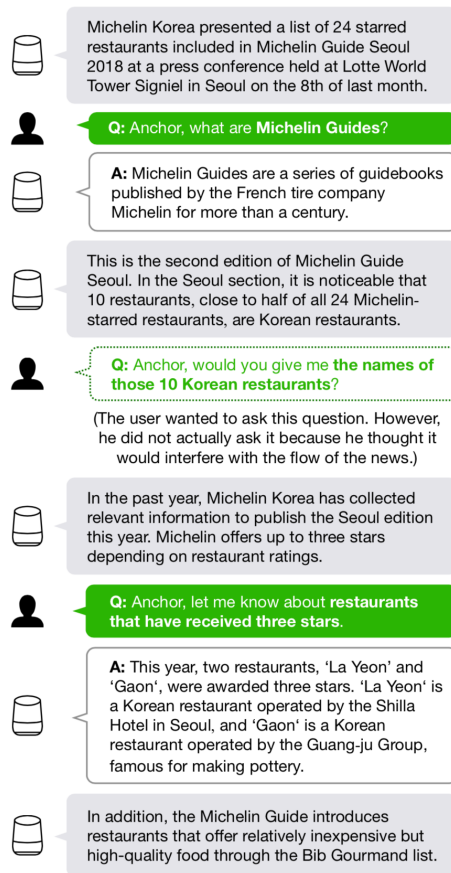
First, we selected the news articles that the system would provide to users. Rather than selecting one topic only, we prepared three articles to meet the various interests of multiple users. We also sought to select topics that could be structured in depth, requiring special expertise, rather than described with just a few sentences. We finally selected Bitcoin, Michelin, and the 2018 Winter Olympics and designed news stories about each of these topics. Then, through the pilot study, we created a corpus of anticipated questions that could be raised for each article and answers to them. As a result, we were able to prepare an average of 12 expected questions and answers per news story.

### Generating speech synthesis audio files

After creating the corpus of news articles and expected questions and answers, we turned them into audio files. In this process, we used the speech synthesis API of Naver Clova (<https://developers.naver.com/products/clova/tts/>). In addition to the corpus, the sentences needed for the introduction of the dialogue, including greetings, were also converted into audio files.

### Setting up the Wizard-of-Oz-based system in the lab

After preparing the speech synthesis files, we installed the research prototype in our lab for the user study. The system consisted of a smart speaker, a researcher’s laptop computer, and a wizard (a researcher who controls the computer) (Figure 1). We used Naver “WAVE” as the experimental smart speaker. The speaker was placed on a table in the middle of the lab, and the wizard and the laptop were located in a separate room that was hidden from users’ view. The researcher’s laptop stored the audio files of the corpus, which were played through the smart speaker connected to the computer via Bluetooth with the wizard’s operation. The user could start the interaction by saying a wake-up word “Anchor” and greeting to the assistant, and the wizard researcher then played the file with the appropriate response. Then, news articles were presented in random order, and when the user asked questions, the researcher selected and played the appropriate answers.



**Figure 2: An example of a user scenario for a situation where a user receives news about the Michelin Guide from Anchor and continues conversational search interactions with it**

## USER STUDY DESIGN

To investigate how users interacted with the system in detail, we conducted a user study, which consisted of a series of news consumption tasks, followed by semi-structured interviews.

### Participants

We recruited participants who had conversation experience with smart speaker or smartphone voice assistants so that they would not have difficulty interacting with the smart speakers by voice in the experiment. We recruited a total of 21 users (10 women and 11 men, mean age of 27.81, SD of 4.26).

### Procedure

Participants visited a laboratory with a research prototype for the experiment. Before each experiment started, the researchers explained the purpose and procedures of the study and told the participants how to use Anchor. They were allowed to interact with it several times to get used to it. Then, Anchor provided the three news stories in random order, and participants could ask questions about the news stories anytime (Figure 2). After the news listening and conversational search, we conducted semi-structured interviews. The participants were mainly asked about the pros and cons of the system and ways in which the system could be improved. All interviews were recorded and transcribed.

### Analysis method

The qualitative data from the user study was analyzed using thematic analysis [1]. In the process, we used Reframer, a qualitative research software tool provided by Optimal Workshop. Through a collaborative, iterative process, we conceptualized and revised main categories.

## RESULTS

Through the user study, we found four main issues.

### Interactive news on topics of interest

First, participants expressed satisfaction with listening to interactive news rather than simply news briefings. They noted that while it could be useful to have a variety of summarized news stories, they preferred news that covers topics in which they are interested. When they heard the interactive news, they described the smart speaker's assistant as smarter and more credible. P04 said, "I usually do not listen to the news with my speakers. But it was nice to hear a relatively long news story written around specific keywords." P17 mentioned, "I felt like it was a professional reporter ... I think the credibility of the news has increased." This shows that voice news needs to be provided in a way that is selective and focused on content that fits the user's interests.

**Questions help users understand the news**

One of the most important points was that the conversational search increased the users' "interest and understanding" of the news content. They went beyond simply listening passively to the news to actively seeking the information needed to understand the news. P08 said, "Usually when I hear 'Bitcoin' news, I don't listen to it because I don't know much about it. But I asked him to explain this term, and Anchor explained it to me so that it was easy to understand." They also found this search method to be more appropriate for voice interaction with a smart speaker assistant. They felt that this interaction reduced the disconnect between them and their assistants and felt that they continued to interact naturally with the assistants.

**Tendency to avoid breaking the news flow**

On the other hand, users did not always ask their assistant questions when they wanted to, even though they were aware that the questions helped them to understand the news. They wanted to answer their detailed questions related to the news, but at the same time, they did not want to miss the overall content the news was intended to deliver. They thought that their questions could interfere with the overall flow of the news. P20 said, "Even though I wanted to ask questions whenever I thought of them, I thought that I should listen to the news until the end." In addition, P13 mentioned, "I thought my question was interfering with the news." This shows that the news needs to be designed so that users can ask questions without hesitation at a desired point in time.

**Out of sight, out of mind**

Participants stated that when the news was over, it was difficult to recall the questions they had hesitated to ask previously. They often could not even remember what they were wondering about after missing the time to ask questions, even though they were very curious at first. They noted that if they asked questions properly, their interest and understanding of the news increased, but if not, they felt their interest in the news fade quickly. P15 said, "I had a question to ask, but I could not think of what it was after the news was over." Moreover, P08 mentioned, "After the news, my interest suddenly disappeared." This shows that information acquisition in voice interactions is very transient and that news needs to be designed to address this point.

**DESIGN IMPLICATIONS**

Based on the findings, we suggest implications for applying a conversational search to audio news.

- Select news that meets users' interests, and provide in-depth content.
- Leave room for users to ask questions to understand the news.
- Design a "come back" conversation so users can re-engage with the news after the Q and A.

- Provide keywords that can be further explained to the user after the news is over.

### LIMITATIONS AND FUTURE WORK

There are several limitations to this study. First, the experiment was done using the Wizard-of-Oz method (i.e., not fully automated). Second, although we selected news stories on a variety of topics, they did not reflect the real interests of the users. Finally, because the experiment was conducted in a lab environment, it did not address how users would use the news service in real life. In future research, we will increase the automation level of the system, generate personalized news tailored to the users' interests, and examine how users use the system in their daily lives. We will also improve the system to reflect our proposed design applications.

### CONCLUSION

In this study, we applied a “conversational search” to the design of the news service of smart speakers so that users can actively seek information while consuming the news. We designed a research prototype called “Anchor,” in which a news assistant provides users with news and responds to users' questions about the news content. Through the user study, we identified that people prefer interactive news to news briefings. Although they found it useful to get answers on their questions by talking with the agent, they often hesitated, as they did not want to miss the whole flow of the news. They also had difficulty recalling the questions they had not asked.

The contributions of this study to the HCI community are as follows. We designed a new type of service in which a conversational search system was combined with audio news, thus pioneering new applications. In addition, through a user study, we empirically identified users' experiences of interacting with these systems. Finally, we induced design implications for user interaction when the conversational search system is applied to audio news. We hope that this work will provide guidance on the application of the conversational search system to voice-only environments.

### REFERENCES

- [1] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.
- [2] Benjamin R Cowan, Nadia Pantidi, David Coyle, Kellie Morrissey, Peter Clarke, Sara Al-Shehri, David Earley, and Natasa Bandeira. 2017. What can i help you with?: infrequent users' experiences of intelligent personal assistants. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services*. ACM, 43.
- [3] Pablo Valerio. 2018. Voice assistants at home to grow 1,000 percent. *IoT Times* (2018). Retrieved January 3, 2018 from <https://iot.eetimes.com/voice-assistants-at-home-to-grow-1000-percent/>.
- [4] Alexandra Vtyurina, Denis Savenkov, Eugene Agichtein, and Charles LA Clarke. 2017. Exploring conversational search with humans, assistants, and wizards. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. ACM, 2187–2193.