Journal of Digital Convergence
Vol. 16. No. 10, pp. 409-414, 2018

ISSN 1738-1916
https://doi.org/10.14400/JDC.2018.16.10.409

A Study on the User Experience of Smart Speaker in China
- Focused on Tmall Genie and Mi AI Speaker -

Xin-Ting Xiao[1], Seung-In Kim[2*]
[1]**Dept. of Digital Media Design, HongIk University, International Design School for Advanced Studies**
[2]**Prof. of Digital Media Design, HongIk University, International Design School for Advanced Studies**

중국 인공지능 스피커 사용자 경험에 관한 연구
- 티몰 지니와 샤오미 스마트 스피커를 중심으로-

소신정[1], 김승인[2*]
[1]홍익대학교 국제디자인전문대학원 디지털미디어디자인 전공
[2]홍익대학교 국제디자인전문대학원 교수

Abstract In China, the usage of smart speaker is continuously increasing. In this study, it is aimed to research on the user experience of the Chinese smart speaker. Therefore, we did literature research followed with theoretical background of smart speaker, and did case study of worldwide popular smart speaker brands. On this basis, we conducted in-depth interview with 8 users who have experienced with the top-selling Chinese smart speaker product "Tmall Genie" and "Mi AI speaker". The interview is based on 7 principles named Honeycomb model, which created by Peter Morville. As a result, users' discomfort was found in the functional part and the usability part of the smart speaker. Furthermore, the users were highly unsatisfied with the smart speaker in the credibility part. Accordingly, Chinese smart speaker should consider the user experience aspects to complement functional and usability parts for user.

Key Words : Smart Speaker, User Experience, Tmall Genie, Mi AI Speaker, Voice Command

요 약 최근 중국의 스마트 스피커 사용량이 지속해서 증가하고 있다. 본 연구는 중국 스마트 스피커 사용자의 사용자 경험을 조사하는 것에 목적이 있다. 따라서 본 연구는 스마트 스피커의 이론적 배경에 관한 문헌 연구와 세계적으로 인기 있는 스마트 스피커 브랜드에 대한 사례 연구를 하였다. 이에 따라 중국에서 가장 많이 팔린 스마트 스피커 제품인 티몰 지니와 샤오미 스마트 스피커 사용자 8명을 심층 인터뷰하였다. 인터뷰는 피터 모빌의 허니콤 모델의 7가지 원칙에 바탕으로 문제를 설정했다. 그 결과, 스마트 스피커의 유용성과 사용성에서 사용자들의 불편함을 알 수 있었다. 또한, 신뢰성에서는 사용자들이 스마트 스피커를 이용하면서 큰 불만을 느끼는 것을 확인할 수 있었다. 따라서 향후 스마트 스피커 사용자들에게 기능성과 사용성을 보완하기 위해 사용자 경험 측면을 고려해야 할 것이다.

주제어 : 인공지능 스피커, 사용자경험, 티몰 지니, 샤오미 스마트 스피커, 음성 명령

1. Introduction

1.1 Objectives and Background of the Research

With the rapid development of computer technology and artificial intelligence, voice interaction, as one of the most natural, harmonious and effective means of

* Corresponding Author : Seung-In Kim (r2d2kim@naver.com)
Received June 11, 2018 Revised July 19, 2018
Accepted October 20, 2018 Published October 28, 2018

human-computer interaction has been receiving universal attention[1]. Therefore, virtual voice assistant market is also increasing. According to survey from Creative Strategies, those who use voice assistant like Siri or Alexa are much more likely to use them at home[2]. In November 2014, Amazon, the US e-commerce company, released the first smart speaker product which called Echo. It is the product that works in the home environment with Alexa as its virtual voice assistant and can perform far-field interaction[3]. It successfully captured people's curiosity as soon as it went into the market.

The market for smart speakers began to expand as the successful of Echo. The Chinese market has also seized on this opportunity, and in 2015, JD.com, Inc., a Chinese e-commerce company released the first Chinese smart speaker. However, the reaction was less than expected. It did not arouse the concern of consumers, but the accuracy of automated speech recognition and natural language understanding are improved step by step. By 2017, dozens of e-commerce companies and tech companies led by Alibaba Group and Xiaomi Inc. have invested in smart speakers[4]. Smart speaker became a hot spot in China.

However, the reviews of Chinese users are not positive as overwhelming as foreign users. Whether the smart speakers that meet the needs of foreign users really adapt to the Chinese market. The purpose of this study is to explore the user experience and satisfaction of Chinese Smart speaker. Through comparison and analysis, we can find out the inconveniences of using process and propose the direction of product improvement.

1.2 Range of the Research

This study involved the investigation of Chinese and overseas' smart speaker brands, as well as the detailed comparison and analysis of the Chinese brand Tmall Genie and Mi AI speaker.

In first phrase, we learn about the characteristics of the most popular product in China and overseas through case study. In second phrase, we conduct in-depth interview to comprehend user experience and the inconvenience in use. Through the above research, the development and improvement direction of Chinese smart speaker is obtained.

2. Background

2.1 Definition and characteristics of Smart Speaker

A smart speaker is a type of wireless speaker and voice command device with an integrated virtual assistant (artificial intelligence) that offers interactive actions and handsfree activation with the help of one "hot word" (or several "hot words")[5].

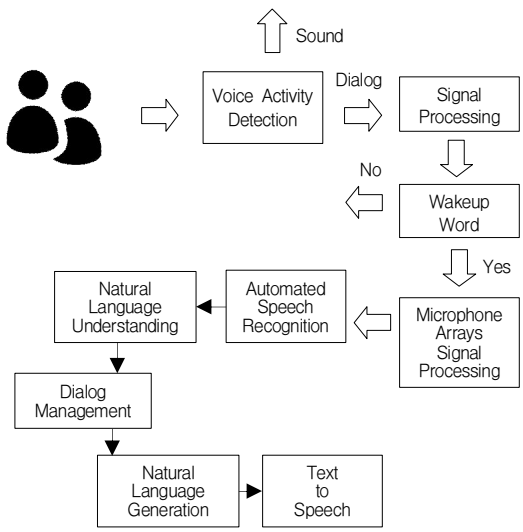


Fig. 1. Smart Speaker Interaction System [6]

Smart speakers are different from traditional speakers in two aspects. Firstly, it can work with WiFi connection and be used by voice interaction; Secondly, it can provide rich contents and services such as music playback, audio books as well as smart home control[7].

Table 1. Characteristics of Smart Speaker [8]

Hardware	Function
Appearance Loudspeaker Microphone Processer/Storage Bluetooth/Wifi ……	Music Playback Voice Service (Weather, Schedule, Stock) Third-party Service (Uber , Food Ordering) Home Automation ……

2.2 Case Study of Oversea Brands

2.2.1 Amazon "Echo"

Amazon Echo is the leader in the category of smart speakers. The first Echo was released by Amazon.com, Inc. in November 2014. Initially, it was just a smart player that could execute several voice commands. However, with the accumulation of user habits and continuous improvement of voice technology in the past several years, now the device is capable of voice interaction, music playback, making to-do lists, setting alarms, streaming podcasts, and providing weather, traffic and other real-time information. It can also control several smart devices acting as a home automation hub[9]. The greatest advantage of Echo is that by binding to amazon accounts, you can make a purchase directly online via voice commands[10].

2.2.2 Google "Google Home"

Google Home was released on May 19, 2016 by Google. Its advantage lies in the mature cloud computing capability and outstanding voice recognition function. In 2011, Google was formally established its AI department. At present, including Google search, Google Now, Gmail and so on, more than 100 teams are using machine learning, and to inject a large number of machine learning function to its open-source Android mobile phone system[11]. Compared with Alexa, Google Assistant dose better in dialog management.

2.2.3 Apple "HomePod"

Apple released HomePod on June 5, 2017. Although HomePod is also a smart entrance, HomePod pays more attention to the sound quality of music. HomePod is supported by Apple Music, which use Siri to identify and collect users' preferences and recommend Music based on users' interests[12]. In addition, HomePod is very easy to connect Apple devices such as Iphone. As long as Iphone is close to HomePod, the two devices will automatically connect to each other within a few seconds[13]. The collaborations between apple's products improve user experience, which solidify users' loyalty.

2.3 Case Study of Chinese Brands

2.3.1 Alibaba Group "Tmall Genie"

Tmall Genie is launched on July 5, 2017 by Alibaba Group, the largest Internet company in China. With strong support from the Alibaba industry chain, Tmall Genie have rich functional contents, such as music, radio, audio books, English learning resources. After binding Taobao(E-commerce platform owned by Alibaba Group) account, the user can shop and enjoy package tracking service. In addition, Tmall Genie will recommend products based on users' previous consumption habits. Through voice recognition, users can easily complete the payment process by saying a certain random voice verification code after adding their voice into a trusted list[14].

2.3.2 Xiaomi "Mi AI speaker"

Xiaomi Inc. released Mi AI speaker on July 26, 2017. It is now the second largest share of China's smart speakers market. After seven years of development, Xiaomi Inc. has grown into the world's largest intelligent hardware platform. Currently, the total amount of Xiaomi smart home networking devices has exceeded 60 million, which is still growing at a high speed[15]. Through Mi AI speaker, we can control Xiaomi smart home devices, such as TV, cleaning robot, air purifier and other devices. In addition, the virtual assistant of Mi AI speaker has its own cartoon image- a girl wearing a powersuit with short red hair[16].

3. In-depth Interview

3.1 Experiment Method

The purpose of this study is to evaluate the usability of Tmall Genie and Mi AI speaker. In-depth interviews was conducted with 8 users who are from 20 years old to 30 years old.

3.2 Research Methods and Subjects

From April 10, 2018 to April 18, 2018., we conducted in-depth interview with 4 male interviewees and 4 female interviewees from 20 to 30 years old who have fully used one of the ai speaker product. We found that we can get enough content through limited time and cost to summarize and conclude by interviewing four costumers of each brand.

Table 2. In-depth interview profile

	Gender	Age	Job	Brand
A	Male	24	Programmer	Tmall Genie
B	Male	30	Financial Planner	Tmall Genie
C	Female	26	Designer	Tmall Genie
D	Female	29	Designer	Tmall Genie
E	Male	25	Office Worker	Mi AI Speaker
F	Male	30	Programmer	Mi AI Speaker
G	Female	24	Designer	Mi AI Speaker
H	Female	29	Office Worker	Mi AI Speaker

3.3 Research process

In-depth interview will explore the usability of smart speaker with Honeycomb Model (Useful, Usable, Findable, Credible, Accessible, Desirable, Valuable), which provides extensive coverage in evaluation categories.

By analyzing the evaluation content, we can find the consumers' satisfaction and the inconvenience through user experience, as well as their needs. Therefore, the in-depth interview questionnaire consists of 15 items in total, with 3 main questions for each element. The follow-up questions and probe will be different based on the specific answer of each main question.

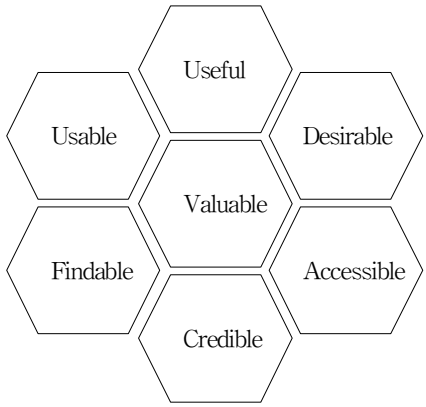


Fig. 2. Honeycomb Model

Table 3. In-depth interview Questionnaire Topics

Components	Questions
Useful	Does it function well as a smart speaker?
	Do you feel frustrated during the operation?
	Does it work as expected?
Usable	Is there any difficulties to start with using a smart speaker?
	Is there any inconvenience in using it?
	Can it recognize your instruction correctly?
Findable	How long does one operation take?
	Can the key word be recognized easily?
	Will it recognized your instructions when you speak with accent?
Credible	Is there any error in use ?
	Do you feel credible using it?
	Do you trust the information provided?
Accessible	Where do you use it?
	Does it work well under any surrounding?
	Can you do mostly operation in voice control without the assistance of the mobile app?
Desirable	Do you enjoy the appearance of your smart speaker?
	Are you satisfied with its response?
	Do you usually use it?
Valuable	Is it irreplaceable?
	Is it worth buying?
	Do you enjoy using it?

3.4 Results

3.4.1 Useful

Both Tmall Genie and Mi AI speaker have reached the user's expectations to a certain extent. Users hold positive attitude on exploring all the functions of the product. However, the users of Mi AI speakers complained more on the copyright issues than the users of Tmall Genie.

3.4.2 Usable

Both of the products are easy to operate. Users can do most operation without reading the instruction. However, they have weakness in the process of voice interaction. When speaking to Tmall Genie, using a certain speaking speed will be better. If user speaks too quickly or too slowly, it may cause a incorrect identify. Compared with Tmall Genie, Mi AI speaker is weaker to do multi-round dialog. Users need to say the wake up word before every sentence.

3.4.3 Findable

Vast majority of voice operations with smart speaker are faster than on mobile phones. However, only Mandarin can be recognized, What is more, even Mandarin can not be 100% well-recognized.

3.4.4 Credible

Interviewees feel a strong sense of distrust of both two products in totally different aspects. The selling point of Tmall Genie is that user can use voice control to recharge their mobile phone bills and make a purchase, but most of the interviewees mentioned that they will not try this function. And one interviewee even had a bad experience with wrong voice recognition, which cause a loss of money. The core function of Mi AI speaker, the access to the smart home control, also does not act like expectation. Lots of Xiaomi smart device still can not be controlled by voice.

3.4.5 Accessible

The products can be easily woke up by everywhere of room. Moreover, users almost do not need a mobile app to assist it. However, the Tmall Genie is more sensitive than Mi AI speaker. Sometimes users need to speaker loudly or repeat the wake up word several times to Mi AI speaker.

3.4.6 Desirable

Users of the Tmall Genie highly appreciate its appearance. On the contrary, users of the Mi AI speaker are not satisfied with its appearance. They gave more praise on its beautiful voice and humorous conversation, which gave them a sense of kindness. The response by Tmall Genie is shorter and more formal.

3.4.7 Valuable

All the interviewees gave positive attitudes to these product. At the mean time they expected the function to be further improved.

4. Conclusion

The purpose of this study is to compare and analyze the user experience of Tmall Genie and Mi AI speaker, two of the most popular smart speakers in China. In this regard, we conducted in-depth interview with 4 users for each brand and reached the following conclusions.

Although the users focus on different functions, all of them have a strong feeling of distrust while using. The reasons for the sense of mistrust are the insecurity of the payment module and the inadequate contents. To solve this problem, the brands must continue to enrich the support content of each function while implementing and improving the existing functionality. In addition, fewer participants use it to control smart home appliances. The reason is that the penetration rate of smart home appliances in China is low. Therefore, this function, which is supposed to be a selling point, is not the demand of Chinese consumers at present stage. Finding more precise needs of Chinese consumer should be the next research phrase.

REFERENCES

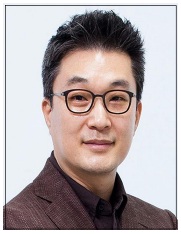
[1] R. H. Wang. (2004). *The Progress of Standardization for Chinese speech Interactive Technology.* Information Technology& Standardization.

[2] Carolina Milanesi. (2016). *Voice Assistant Anyone? Yes please, but not in public!* . Creative Strategies(online). http://creativesstrategies.com/voice-assistant-anyone-yes-please-but-not-in-public/
[3] (2017). *Research Report of Amazon Echo in 2017.* Woshipm(online). http://www.woshipm.com/evaluating/781540.html
[4] (2018). *Seven Predictions of Smart Speaker in 2018: Will it Become a Hot Item or Maintain the Status.* Zhidx(online). http://zhidx.com/p/104936.html
[5] Wikipedia. *Smart Speaker.* https://en.wikipedia.org/wiki/Smart_speaker
[6] Y. O. (2018). *Process of Voice Interaction.* 36kr(online). http://36kr.com/p/5114712.html
[7] C. F. (2018). *The current position of Smart Speaker in China in 2017.* Analyse(online). https://www.analysys.cn/analysis/trade/detail/1000830/
[8] Technology Supports of Smart Speaker. https://blog.csdn.net/jackytintin/article/details/62040823
[9] Kriti Mishra. (2018). *What is Amazon Echo?.* Nimbuzz(online). http://www.nimbuzz.com/et/articles/technology/what-is-amazon-echo-817
[10] (2017). *Research on Smart Life : Smart Speaker Industry.* Baijiahao(online). http://baijiahao.baidu.com/s?id=1583655448655151997&wfr=spider&for=pc
[11] (2017). *Music is another way to go. Apple HomePod is not going to compete with Amazon Echo.* Baijiahao(online). http://baijiahao.baidu.com/s?id=1569457828496388&wfr=spider&for=pc
[12] Q. H. Zhao. (2017). *Apple released the smart speaker, HomePod, without comparing it to the Amazon Echo.* Leiphone(online). https://www.leiphone.com/news/201706/0kTzBmaVKTr5pS1o.html
[13] *How to evaluate HomePod?* Zhihu(online). https://www.zhihu.com/question/60727115
[14] G. Y. Yang. (2017). *Tmall Genie: To Lead a New Round of Human-computer Interaction Revolution with High Starting Point and Strong Ecology.* Baijiahao(online). http://baijiahao.baidu.com/s?id=1572359990224005&wfr=spider&for=pc
[15] (2017). *Lei jun announced that xiaomi has become the world's largest intelligent hardware IoT platform.* Huanqiu(online). http://tech.huanqiu.com/it/2017-11/11407780.html
[16] Q. W. Fei. (2017). *Mi AI Speaker is a 17-year-old Girl with Red Hair.* DoNews(online). https://news.znds.com/article/news/26128.html

소 신 정(Xiao, Xin Ting) [학생회원]

・2016년 6월 : Jiangnan University, School of Digital Media, Animation, Bachelor
・2018년 3월 ~ 현재 : 홍익대학교 국제디자인전문대학원 디지털미디어디자인 전공
・관심분야 : 사용자경험디자인, 미디어 아트
・E-Mail : coevonne@gmail.com

김 승 인(Kim, Seung In) [종신회원]



・2001년 3월 ~ 현재 : 홍익대학교 국제디자인전문대학원 교수
・2006년 3월 ~ 현재 : 홍익대학교 디자인혁신센터 센터장
・관심분야 : 사용자경험디자인, 서비스디자인, 시각디자인
・E-Mail : r2d2kim@naver.com