Towards Immersive Inclusivity for C2C: How Immersive Multimodal Interactions Can Make Online Customer-to-Customer Shopping More Inclusive

BEKTUR RYSKELDIEV, Mercari R4D, University of Tsukuba, Japan

MARIKO KOBAYASHI, Mercari R4D, Japan

KENTARO TERAMOTO, Mercari R4D, Japan
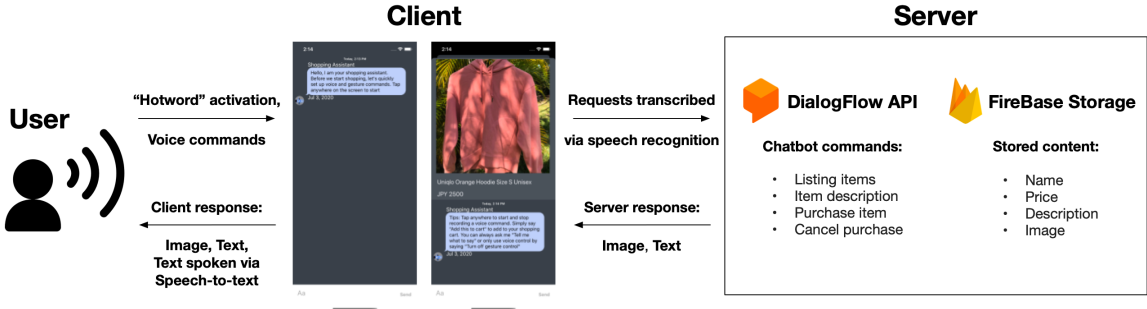
KOKI KUSANO, Mercari R4D, Japan

Fig. 1. Prototype application system architecture diagram

Online shopping is becoming ubiquitous in everyday life, with shopping experiences primarily delivered through either websites or specialized applications. This format may create inclusivity problems for users with different needs, as not all websites or applications are optimized for accessibility or have multiple alternative forms of interaction with the content. This can be especially challenging in customer-to-customer (C2C) platforms. While in online shopping a single platform can make merchants provide a unified display for description, standardized packaging, and ensure the item quality, in C2C platforms this responsibility falls mostly on users, thus creating a need for a more versatile way of processing, display, and interaction with C2C content. We are interested in how this issue can be alleviated through addition of online assistants that support immersive and multimodal forms of interactions. To gain a better understanding of the problem and needs of users, we have built a prototype application that uses a voice-activated assistant for item search and purchase and evaluated it in a pilot study. The findings suggest that participants value efficiency in C2C shopping and find current platforms problematic, while showing interest in voice search, immersive and multimodal interfaces to improve the overall shopping experience, despite some issues with voice recognition and response time.

CCS Concepts: • **Human-centered computing** → **Human computer interaction (HCI)**; **Interaction techniques**; **Interaction design**; **Accessibility**; • **Computing methodologies** → *Artificial intelligence.*

Additional Key Words and Phrases: artificial intelligence, inclusive design, human-computer interaction, immersive media

1 INTRODUCTION

Online shopping is becoming increasingly popular, with most of the interactions done through either specialized websites or mobile applications. This can potentially create an inclusivity problem: not all websites or applications are optimized for users with disabilities or are compatible with common accessibility tools (e.g., online screen readers). In case with customer-to-customer (C2C) markets, such as online auctions, this problem is especially noticeable as users can be trading previously owned items (and thus unable to confirm the condition of the item prior to or after the purchase) or have to use some form of user-to-user communication for shipping and price negotiation online.

At the same time, immersive multimodal forms of interaction with content, such as voice assistants or Augmented Reality (AR) are gaining popularity in online shopping, by enriching the browsing experience (e.g., 3D clothing viewers in AR[10]) or providing an alternative mode of communication (e.g., shopping via voice assistants[1]). Therefore, we are interested in how immersive technologies can address potential inclusivity issues for C2C environments and create a more natural form of information display and communication online.

In order to investigate that we conducted a two-part pilot study. In the first part we have evaluated participants' general opinion towards online and C2C shopping in general with and without immersive interfaces. Then, in the second part we evaluated their interest in immersive interactions. Since, to our knowledge, at the time of the study there were no C2C shopping applications with conversational agents that would have a voice command support, we have developed a prototype mobile application that applied an online artificial intelligence (AI) assistant for information retrieval and communication, and supported voice-activated interactions with exhibited content. We asked participants to try our application and express their thoughts on presented form of interaction via questionnaire.

The preliminary results have shown interest among participants towards the voice assistant form of interaction presented in the prototype. Based on the initial positive feedback, we believe that our work has a potential to benefit both the future of inclusive design and immersive media research, as well as encourage improvement of user experience in C2C shops and services online.

2 BACKGROUND

While the topic of inclusive online user experience, and especially online shopping, is not particularly new – there are established recommendations for both mobile [2, 3] and web experiences [11] – the applications of inclusive online experiences in combination with immersive and mutlimodal interfaces have been gaining attention in Human-Computer Interaction (HCI) community.

Recent successful examples of using immersive (i.e., different from *unmediated reality* as defined by Skarbez et al. [8]) media for inclusive and accessible design include: applications of mobile AR for assistance with dynamic menus in vending machines and kiosks for users with disabilities [5], assistance with outdoor navigation by using spatial audio markers [6], head-mounted displays (HMD) for vision enhancement [12], and navigation through adaptive AR markers [7].

As discussed in work on assistive augmentation by Huber et al. [4], using immersive and multimodal interfaces gives users an opportunity to both supplement and enhance their senses. Thus, we believe that in the context of C2C

shopping, where users have to deal with potential challenges in communication, as well as different tasks related to item condition estimation, sizing, and packaging – providing alternative forms of search and interaction with the content could benefit all users.

To our knowledge, the problems of this particular type of inclusive online shopping are still relatively unexplored. Out of recent studies most similar to our goals, which would combine both elements of immersive interactions and web accessibility in the shopping context, our work bears resemblance to a project by Stangl et al. [9] which featured an online smart assistant which would process web store content and present it in an accessible manner through a conversational smart assistant.

3 IMPLEMENTATION

Based on these observations we have outlined two main goals of our investigation: *to gain a better understanding of users' needs when interacting with online C2C shopping applications,* and *how immersive multimodal interfaces can address these needs.* Our study hypotheses are that *the current implementation of C2C marketplace applications can make buying experience cumbersome for users*, and that *such inefficiencies could be alleviated through immersive and multimodal solutions.* We have designed a pilot study that would investigate both users' general preferences and issues in C2C shopping online, as well as their interest in immersive and multimodal interactions. The study consisted of two parts.

3.0.1 Part I: General Questionnaire. In the first part we investigated users' general background, online shopping preferences, and experience of using multimodal assistants. The questionnaire was based on a similar investigation by Stangl et al. [9] and included questions on: age, gender, which e-commerce services and platforms they prefer to use, and whether they have a previous experience of using voice assistants. In addition, we investigated whether participants are satisfied with current C2C marketplace interfaces by including six Likert scale (1 - Disagree, 5 - Agree) questions and one open comment question on the perceived importance of online shopping, as well as importance of having efficient interactions during search and purchase online, and have a possibility of using alternative means of accessing the shopping content (Table 1, questions 1-7).

3.0.2 Part II: Application Prototype. In the second part we investigated how users would react if there was an interface that addresses the shopping issues discussed in the first part of the study. Since to our knowledge at the time there were no commercial or experimental online C2C marketplace applications that used a voice-controlled conversation bot for shopping, we have implemented our own multimodal interface which featured one input modality (voice) and three output modalities (voice, text, pictures). In such case, participants' would have to reflect on their experience of using conventional (text, picture) C2C marketplace platforms in comparison with our prototype.

The prototype (working project name "ShopWithMe") runs on Google Dialogflow API[1] for chatbot replies, which gets the data for shopping items from our database hosted on Firebase[2] storage service, which is then delivered to a mobile application (Fig.1, 2). The mobile application is written in Swift 5.3 with iOS 14 SDK. It uses system-based frameworks for voice synthesis and recognition, and a 3rd-party framework Porcupine[3] for hotword (a word that starts the voice recognition session) detection. At the time of the testing the application supported such features as listing items, showing various item details, and provided purchase and cancellation of purchase functionality. All of the

---

[1]https://cloud.google.com/dialogflow/docs
[2]https://firebase.google.com
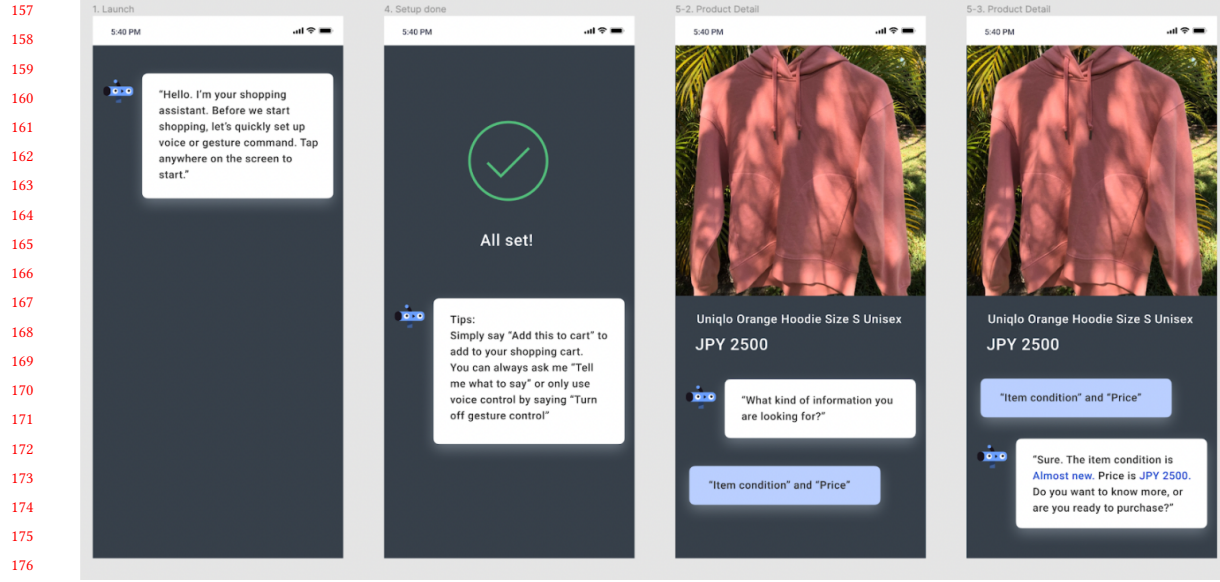[3]https://github.com/Picovoice/porcupine

Fig. 2. Prototype application interface

features except for the item listing were available both through UI and voice command (via conversational chatbot assistant).

In each session the process would feature a user saying a hotword "porcupine" (predefined by the hotword detection library), which would start the voice recognition session, then the user can pronounce commands which would be then transcribed into text and sent to the server running Google Dialogflow API. Once a command has been recognized, the information, retrieved from our database is displayed as text and (if available) picture, with the text being read out loud through iOS in-built text-to-speech feature.

After the users tested the prototype, we investigated whether they found provided modalities convenient, asked users to compare it with their previous experience of using C2C shopping applications, and provide additional open comment replies on immersive and multimodal features they would like to see in online shopping. The questions included five Likert scale questions (1 - Disagree, 5 - Agree) and three open comment questions (Table 1, questions 8-15).

4 PILOT EVALUATION

We have recruited 8 participants, age 20-39 (SD = 7.15, Mean age = 27.14), all participants self-identified as male. Out of 8 participants, 6 spoke Japanese as a native language, and 2 spoke English as a second language. Participants were recruited via word of mouth, and were offered a 3000 Japanese Yen compensation at the end of the experiment. Prior to the experiment, the procedure was also reviewed and approved by an internal ethics review committee at our institution (Mercari, Inc), with all participants giving us a written agreement on participating in the experiment and data collection. Each session was held in-person in an empty ventilated office space, with social distancing and mask wearing measures implemented.

Table 1. Questions **1-7** are the first part of evaluation, while **8-15** are the second part

	First part
1	I often shop via customer-to-customer (C2C) services (such as C2C marketplaces, auctions, etc)
2	Online C2C shopping is important to me in daily life (such as finding better deals for items, or selling goods)
3	Finding items I want to purchase efficiently (quickly) is important to me
4	Being able to easily (quickly) list items for sale is important to me
5	Being able to efficiently (quickly) communicate with a seller/buyer is important to me (for example, such as being able to send quick/automated replies for trading or user rating)
6	I would like to be able to have an alternative way of shopping on C2C marketplace services (for example, by using voice commands, or using other methods such as Augmented Reality, etc.)
7	(Open comment) What is your main inconvenience with online shopping? What features would you like to see in C2C shopping applications?
	Second part
8	I found using voice commands convenient
9	I found using chatbot for item search convenient
10	I found using chatbot for setting up a purchase (checking the details of the item, price, condition, etc)
11	I would like to use a chatbot feature for C2C shopping in the future
12	I would like to use a voice command feature for C2C shopping in the future
13	(Open comment) What part of interaction with the ShopWithMe application did you like or find interesting?
14	(Open comment) What part of interaction with the ShopWithMe application did you dislike?
15	(Open comment) How would you improve the ShopWithMe application? What features would you like to see?

5   RESULTS AND DISCUSSION

Based on the survey results, it appears that participants find C2C shopping important in their daily life and are valuing efficiency in item purchase, search, and seller communication. This correlates with our hypothesis that the current implementation of C2C shopping platforms where sellers have to provide (sometimes inconsistent) item details is problematic for buyers. It is also reflected in open comments. For instance participant 3 notes that *"The website is complicated for finding specific functions,"* while participants 5, 6, 7, 8 have reported having a fatigue from reading item description when browsing for new items.

Voice search seemed to have interested the participants. As participant 4 has noted *"It would be nice if it could handle imperfect commands, such as being able to communicate even with words (even if the accuracy drops a bit)."* Given the previous feedback on description fatigue, a more convenient way of information representation, such as using a conversational assistant can improve the C2C shopping experience.

This was also reflected in the second part of the survey. After testing the application, all participants have noted the convenience of using conversational agent for item search and purchase and were interested in using such in the future. In open comments, the participants have noted the speed and ease of access of information in comparison with existing C2C shopping experiences, while participant 4 have also noted that *"The chatbot spoke in great detail when I asked for more information. It was fresh to see that reviews were also included."*

We did however notice certain issues, for instance using a hotword detection has proven to be an issue since the framework could not recognize different English accents. We believe that it would be more convenient to either remove
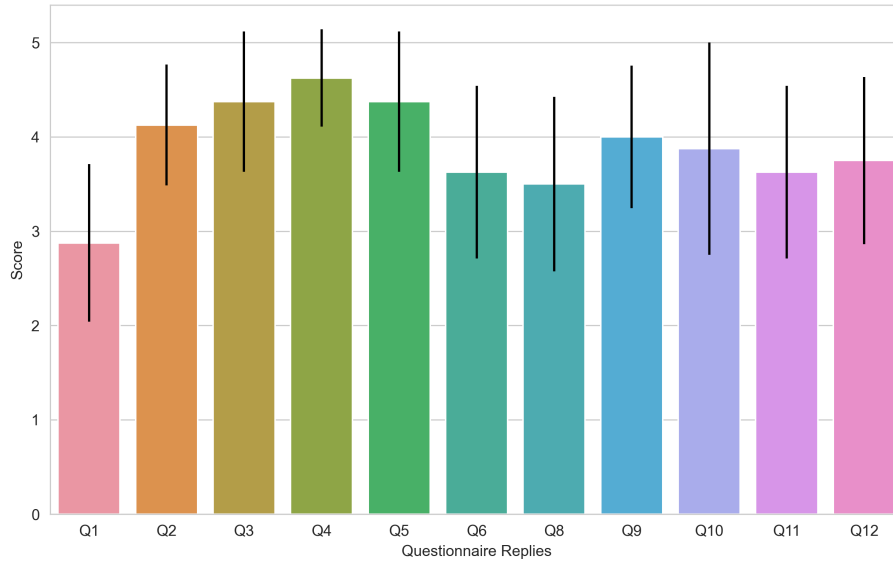
Fig. 3. User questionnaire results (error bars represent standard deviation)

Table 2. Questionnaire scores (Part 1)

	Q1	Q2	Q3	Q4	Q5	Q6
Mean	2.88	4.12	4.38	4.62	4.38	3.62
Standard deviation	0.83	0.64	0.74	0.52	0.74	0.92

Table 3. Questionnaire scores (Part 2)

	Q8	Q9	Q10	Q11	Q12
Mean	3.5	4.0	3.88	3.62	3.75
Standard deviation	0.93	0.76	1.13	0.92	0.89

the hotword detection from the workflow (by for instance replacing it with an onscreen button) or explore alternative frameworks (e.g., using built-in system frameworks). Similarly, voice recognition was also imperfect, as it did not work well with different English accents.

Besides voice recognition, the participants have also commented on the conversation bot. While they found it useful, they commented on response time (need faster response time), and requested to have more conversational features. For instance, participant 6 have noted *"I wanted information on who was selling it and when it would be delivered. I also felt that shipping and payment methods were necessary information for c2c."*

Issues with voice recognition and conversation bot might have influenced participants' opinions on voice interactions. All participants were not sure about the necessity of voice control or voice interaction in general, but were interested in trying it again in the future. This can be also perhaps explained that 5 out of 8 participants have used voice control

features on their phones, but they have never used it for shopping (only using it for voice commands within the system or application).

Aside from the voice features, some participants have also noted the other aspects of C2C shopping, such as shipping. For instance, there was a request for size estimation, where participant 5 noted: *"This may out of scope of the purpose on ShopWithMe, but I thought it would be very useful when buying clothes, etc., if there was a function that estimated the size of the product from the image and told me when the product did not have a satisfactory description."*

Overall, we found the initial feedback promising, as it outlined participants' needs and interest in online C2C shopping experiences as well as how immersive multimodal interfaces can solve the issue. We are currently working on conducting a similar user evaluation on a larger and a more diverse pool of participants. Furthermore, we are narrowing down the group of targeted users, and are considering focusing the next update of our system on users with low vision disabilities.

6 CONCLUSION

We have investigated how immersive multimodal interfaces can positively affect interactions with C2C shopping content. Our preliminary results indicate that:

- Users are interested in a more optimal item search and purchase process which is at the moment a noticeable problem on C2C shopping platforms. Immersive interfaces can assist this by making interactions more natural
- Information representation (such as item description) is of a particular problem for users when presented in text form. Providing a way for a more convenient processing, retrieval, and representation of this information can improve the shopping process
- Having a multimodal way of interacting with content is important: for instance instead of using hotword detection, using a combination of a UI button and voice recognition could be more convenient for users
- Aside from the voice search, there is also interest towards other aspects of immersive interfaces, such as for instance using AR applications for item size estimation

We believe that making the marketplace accessible and inclusive is an especially important issue. The current C2C marketplace has still complicated accessibility issues in buying and selling phases, items listed on the market having varied quality and different level of description detail. In this work in progress study we have demonstrated the possibility to encourage further insight into a more inclusive C2C marketplace by using immersive interactions, which empower various individuals to participate freely in C2C market. We hope our work would be beneficial both for researchers and industry practitioners working on inclusivity issues in online marketplaces.

REFERENCES

[1] Amazon. 2023. Alexa Voice Shopping. Website. Retrieved February 25, 2023 from https://www.amazon.com/alexa-voice-shopping/b?ie=UTF8&node=14552177011.

[2] Apple. 2023. Accessibility. Article. Retrieved February 25, 2023 from https://developer.apple.com/design/human-interface-guidelines/accessibility/overview/introduction/.

[3] Mars Ballantyne, Archit Jha, Anna Jacobsen, J. Scott Hawker, and Yasmine N. El-Glaly. 2018. Study of Accessibility Guidelines of Mobile Applications. In *Proceedings of the 17th International Conference on Mobile and Ubiquitous Multimedia* (Cairo, Egypt) *(MUM 2018)*. Association for Computing Machinery, New York, NY, USA, 305–315. https://doi.org/10.1145/3282894.3282921

[4] Jochen Huber, Roy Shilkrot, Pattie Maes, and Suranga Nanayakkara. 2018. *Assistive augmentation.* Springer.

[5] Junhan Kong, Anhong Guo, and Jeffrey P. Bigham. 2019. Supporting Older Adults in Using Complex User Interfaces with Augmented Reality. In *The 21st Int. ACM SIGACCESS Conf. on Computers and Accessibility* (Pittsburgh, PA, USA) *(ASSETS '19)*. Association for Computing Machinery, New York, NY, USA, 661–663. https://doi.org/10.1145/3308561.3354593

[6] Microsoft. 2018. Soundscape. Article. Retrieved February 25, 2023 from https://www.microsoft.com/en-us/research/product/soundscape/.

[7] Sebastião Rocha and Arminda Lopes. 2020. Navigation Based Application with Augmented Reality and Accessibility. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI EA '20)*. Association for Computing Machinery, New York, NY, USA, 1–9. https://doi.org/10.1145/3334480.3383004

[8] Richard Skarbez, Missie Smith, and Mary C Whitton. 2021. Revisiting milgram and kishino's reality-virtuality continuum. *Frontiers in Virtual Reality* 2 (2021), 647997.

[9] Abigale J. Stangl, Esha Kothari, Suyog D. Jain, Tom Yeh, Kristen Grauman, and Danna Gurari. 2018. BrowseWithMe: An Online Clothes Shopping Assistant for People with Visual Impairments. In *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility* (Galway, Ireland) *(ASSETS '18)*. Association for Computing Machinery, New York, NY, USA, 107–118. https://doi.org/10.1145/3234695.3236337

[10] Vyking. 2023. AR Try-on Technology For Footwear. Website. Retrieved February 25, 2023 from https://www.vyking.io/.

[11] WCAG. 2023. Web Content Accessibility Guidelines (WCAG) Overview. Article. Retrieved February 25, 2023 from https://www.w3.org/WAI/standards-guidelines/wcag/.

[12] Yuhang Zhao, Sarit Szpiro, and Shiri Azenkot. 2015. ForeSee: A Customizable Head-Mounted Vision Enhancement System for People with Low Vision. In *Proceedings of the 17th International ACM SIGACCESS Conference on Computers Accessibility* (Lisbon, Portugal) *(ASSETS '15)*. Association for Computing Machinery, New York, NY, USA, 239–249. https://doi.org/10.1145/2700648.2809865