It Could Literally Change My Life: Exploring the Potential of Conversational Interaction for Indoor Wayfinding Among People with Visual Impairments

Segun J. Samuel, Mohammad Adnaan, Jeremy R. Cooperstock
Department of Electrical and Computer Engineering
McGill University
Montreal, Canada
e-mail: {segun.samuel | mohammad.adnaan}@mail.mcgill.ca, jer@cim.mcgill.ca

Ahmed Farooq
School of Information Sciences
Tampere University
Tampere, Finland
e-mail: {ahmed.farooq}@tuni.fi

***Abstract*—Navigating unfamiliar and complex indoor environments independently is challenging for People with visual Impairments (PVI). As a result, PVI rely on a variety of Assistive Technologies (AT) to preplan and execute their journeys indoors. Among these AT are the emerging Conversation Interaction (COI) systems that leverage Large Language Models (LLMs) to deliver engaging experiences. Specifically, for indoor location assistance, there is an increasing focus on the remote exploration of Points of Interest (POIs), such as shopping malls, before planning a visit. However, there are few studies on the use of COI systems for indoor wayfinding and related aspects, such as route rehearsals—a process of learning the required sequence of actions from one location to another. Consequently, it is not clear what kind of spatial information PVI would expect from conversational interaction agents, how they would use these aids, and the challenges that might arise from their use. To explore the potential of conversational interaction as a modality for indoor wayfinding assistance for PVI, we developed 'GeoChatre,' an interactive mobile COI app that enables contextual learning of unfamiliar indoor routes. Our study reveals the specific types of spatial information that PVI expect from conversational agents for indoor wayfinding, including step-based distance estimates, directional guidance, landmarks, obstacle awareness, and shows that contextual, progressive disclosure of information supports route recall. Additionally, our results show that voice-based interactions in shared indoor spaces raise privacy concerns, particularly on destination disclosure, highlighting the need for coded interactions and discreet input alternatives.**

Keywords-AI and Accessibility; Assistive Technologies; Navigation; Conversational Interaction; Visual Impairments; Privacy.

I. INTRODUCTION

Buildings with complex layouts present significant navigation barriers for People with Visual Impairments (PVI), particularly when visiting for the first time [1][2]. While a range of Assistive Technologies (AT) have been developed to support indoor mobility [3], the integration of Large Language Models (LLMs) into COI systems (e.g., ChatGPT) has opened a new avenue for delivering spatial information through natural dialogue [4][5], since LLMs provide a more accessible knowledge source than conventional Question and Answering (Q&A) systems [6].

For example, there is an increasing use of COI for indoor wayfinding assistance to support exploration of points of interest [7], such as in shopping malls [8], to provide users with information based on personal interests before making travel decisions. Conversational interaction is also being used to select or specify destinations [9]–[12], or for human localization in indoor environments through intelligent conversation between users and an agent [13]. While some studies have investigated COI for indoor wayfinding [13]–[16], there is limited work targeting indoor route rehearsals for PVI—the process of learning the required sequence of actions to reach destinations or return to their origin. Instead, blind individuals often acquire such route knowledge through the use of tactile maps [17] or with the assistance of Orientation and Mobility (O&M) specialists, especially when changing environments [18].

However, tactile maps assume tactile literacy on the part of the users and lack the capacity to provide detailed information [19], while O&M experts might not always be available. To overcome these limitations, this research explores the use of COI agents as wayfinding assistants to provide users with advance knowledge of POIs and facilitate their learning of indoor routes [5]. Our research considers the following questions:

RQ 1: What kind of spatial information do PVI expect from conversational interaction for indoor route learning?

RQ 2: How would PVI like to use conversational interaction for indoor route learning?

RQ 3: What challenges/reservations might PVI have with the use of conversational interaction for indoor wayfinding assistance?

The contributions of our study are twofold: First, we identify the specific types of spatial information that PVI expect from a conversational agent for indoor wayfinding assistance, including step-based distance estimates, directional guidance, landmarks, and obstacle awareness. Second, we uncover that voice-based interactions in public indoor spaces raise privacy concerns that might have often been overlooked in AT designs, with participants distinguishing between socially neutral and sensitive locations, thereby contributing new insights into the social dynamics of using COI systems in shared indoor environments.

The rest of the paper is structured as follows: Section II provides an overview of related work, followed by a brief explanation of system design in Section III. User recruitment, demographics, and study procedures are discussed in Section IV. Sections V and VI present the results and the implications of findings, respectively, while the limitations of the study are highlighted in Section VII. Section VIII concludes the paper, summarizes the central findings, and outlines directions for future work.

II. Related Work

Independent wayfinding in unfamiliar environments is challenging for most people, especially for PVI [11]. Accordingly, AT for navigation usually provide guidance or facilitate knowledge of the surroundings for PVI [20][21]. Previous research has shown the potential of traditional tactile maps to convey spatial environmental knowledge [22]. There are also digital interactive maps that provide dynamic spatial information and auditory feedback for spatial learning [17][23][24], but they impose hardware requirements [20] and are limited in their adaptability to different user needs or contexts without major modifications [17]. Other attempted solutions use commodity smartphones to provide virtual navigation via a sequence of turn-by-turn instructions, and render relevant POI information [11][12][20][25], but as these are passive systems, they do not support interactive Q&A as might be needed by PVI [8][17].

Since natural language dialogue is often the preferred interaction method by PVI [26], more AT are integrating COI agents to improve user experience [7][8][27][28]. Research on COI systems, particularly for indoor location assistance, has primarily focused on exploration of POIs. Such approaches allow remotely "probing" POIs for information based on individual interests [8][27][29]. There is also research on general scene understanding, mostly for outdoor settings [30], and on obstacle avoidance [31]. However, to the knowledge of the authors, only a few studies have been carried out on COI for wayfinding [8][13], and related aspects, such as preplanning indoor routes. Addressing these gaps in the literature is of value to PVI, since acquiring cognitive maps of both outdoor and indoor environments, before independent navigation, is crucial [5][17]. Moreover, there is no guidance available as to the kind of spatial information that PVI would expect from conversational interaction agents for indoor wayfinding, how such agents would be used, and the challenges that might arise in their use.

III. System Design of GeoChatre

To address the gaps in the literature identified above, we developed "GeoChatre," an interactive mobile app that enables contextual learning of unfamiliar routes in complex indoor environments, following the principles recommended for designing conversation agents for navigation [32]. GeoChatre is intended for a number of situations: 1) where PVI have a specific destination in mind for independent navigation, 2) scenarios in which users are temporarily disoriented, e.g., in complex building environments, and seek to regain orientation, and 3) where individuals simply want to know what POIs exist along a given route, either for the sake of their spatial awareness or to plan journeys there.

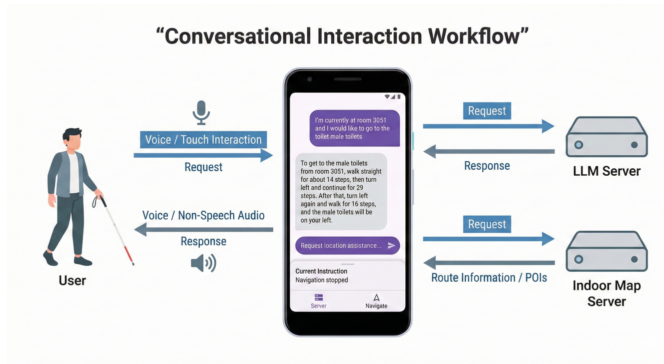

Figure 1. Conversational interaction architecture: GeoChatre.

The GeoChatre system, as shown in Figure 1, was implemented on Android, and tested on a Google Pixel 9 Pro smartphone. The app uses the "Take Me Out of Error" (TACME) indoor localization framework based on pedestrian dead reckoning that we developed as part of a larger study. TACME runs on a server built on indoor maps from floor plans to render routes and POIs information. The architecture comprises three principal components: a client-side, an LLM server, and an indoor map server. Communication between these components follows a request-response pattern over the REST Application Programming Interfaces (APIs).

A. Input Processing and User Intent Classification

Users interact with GeoChatre through voice or text input. For voice interaction, users double-tap the screen, triggering distinct audio cues (speech and non-speech) to signal the start and end of voice registration. Voice input is captured using Android's built-in SpeechRecognizer. The system also supports textual input. Once a verbal request is transcribed, it is sent to the LLM server, which runs Llama 3.2 3B. The LLM is prompted to perform intent classification to determine whether the input represents a new route request, a follow-up inquiry, or an ambiguous query requiring clarification. For new route requests, the mobile application sends the extracted source and destination to the indoor map server via a REST API call. The server provides spatial data, including relevant safety cues for each route segment.

B. Data Retrieval and Response Generation

The route data retrieved from the map server is combined with the user request and passed to the LLM. The LLM, by prompts, synthesizes route information into a natural-language response tailored for PVI. For initial direction requests, the

LLM generates concise summaries by combining consecutive similar actions (e.g., aggregating multiple short segments into a concise instruction such as *Walk straight for 15 steps, then turn right at the water fountain and walk for an additional 6 steps. Main restroom will be on your left*. Users can request more detailed information through follow-up queries. Following this approach to information rendering, our information design can be situated within some established theoretical frameworks: the Landmark-Route-Survey model of spatial knowledge acquisition [33][34] and the information-seeking mantra of "overview first, zoom and filter, then details on demand" [35]. In other words, when a user first makes a direction request, GeoChatre delivers a concise route summary that aggregates consecutive navigation actions into an overview. Users can then progressively access richer spatial information through follow-up questions—effectively achieving details on demand through natural dialogue. The generated response is delivered to the user through Android's text-to-speech engine, with audio feedback beeps providing interaction cues for different user request and response states. The LLM is also customized through prompting to provide multilingual indoor navigation assistance required by PVI.

IV. User Study

We secured the approval of the McGill University Research Ethics Board (REB #19-10-041) for our study. Seven participants, ($N = 7$; 2 female, 5 male, mean age of $M = 41.9$ years ($SD = 10.2$) took part in the experiment, and were compensated CA $15 per hour for their time. One completely blind person participated in the study. The remaining six participants had corrected-to-normal vision. The study was conducted on the third floor of a multi-story university building. The indoor map server is developed from the floor plan. The map indexed 28 POIs across the navigable area, organized into categories, including classrooms, offices, and amenities such as a printing station, a lounge area, study areas, emergency exits, and vertical circulation points. The building features an L-shaped main corridor spanning approximately 200 meters, with multiple perpendicular branches extending to classrooms, offices, and amenities.

We collected demographic information from the subjects, including their vision level, time of blindness onset, use of wayfinding aids, and whether they have any hearing impairments. A short training session was conducted to familiarize them with the control settings for the app, such as for the loudness and speed of text-to-speech. Also, during this phase, participants received names of different indoor locations and learned to request direction guidance from their current position. They were also informed that they can ask for additional details, such as nearby landmarks, estimated walking time, safety cues, and any other information they deemed helpful for constructing mental representations. In the experiment, subjects were requested to choose from indoor locations that they would visit and use the COI tool to explore and familiarize themselves with the route information prior to undertaking the journey. The goal was to explore how COI systems would support the development of spatial knowledge for real-world navigation. After the participants carried out these tasks, they completed a post-questionnaire survey. We also conducted semi-structured interviews to gain deeper insights into the experiences of the participants. The sessions were audio recorded, and each study session lasted approximately 60 minutes.

V. Results

A. RQ1: What kind of spatial information do PVI expect from conversational interaction for indoor route learning?

The analysis reveals that subjects expected a rich set of spatial information from GeoChatre. Their expectations clustered around four distinct categories of spatial information, each serving a complementary role for independent wayfinding indoors.

1) Step-count Information: Step-based distance estimation provided by GeoChatre emerged as one of the most widely discussed types of spatial information. Step counts were valued as actionable cues that support self-location awareness. As some participants explained: *It is able to calculate the distance between where I am and where I am going to by steps. That means I am able to count my steps and get to where I am going to.* (P001); *... it can also tell me how many steps to take to get there ...* (P002). Some expressed that the step information can help to determine how much effort is required to complete wayfinding tasks. For example, *it can also help me to know how much effort it will take me to get there ...* (P002). Importantly, participants demonstrated an understanding of the limitations of step-count cues. One acknowledged that *the number of steps might not be accurate, since the number of paces for each person may be different.* (P002). However, they still valued it as *a rough idea of what the distance is.* (P007). This suggests that PVI may not require absolute precision, but rather a reliable distance estimate that supports spatial knowledge of the journey ahead.

2) Landmark Identification and Sequential Ordering: Landmarks emerged as prominent spatial features expected with their sequential presentation along the routes. There were several expressions, such as, *it is able to tell me the different landmarks I will encounter on my path... the water fountain, study area, the fire extinguisher, the elevator, then the female toilet..., the female restroom before...* (P001). Some highlighted the value of the order of information presented: *the aspect I find interesting is the sequence of landmarks starting from where you are leaving to... the way it arranged it sequentially... it is very interesting.* (P004). It appears the sequential ordering of landmarks creates a narrative structure of the journey that supports the construction of mental maps, just as expressed by one participant that *... being able to map the indoor area is quite interesting.* (P006)·

3) Turn-by-Turn Directional Information: Participants valued explicit Turn-by-Turn instructions (TbT) as directional cues. Most of them successfully recalled the routes. For example, a participant recounted one of the route instructions saying, *It said from here, then I go... take a right 14 steps,* *take left 29 steps, left again 16 steps, and my destination will be at the left.* (P001). This demonstrates the degree to which these instructions were internalized and retained. Another account reveals how TbT and landmarks work in concert: *when they say you go straight, you turn left and you continue a bit forward and I think they gave indications also you're going to see... bins, you're going to see fountain and you're going to pass one of the labs.* (P007). The integration of directional, landmark, and metric information constitutes the comprehensive spatial information expected by PVI.

4) Obstacle Awareness and Safety Cues: Obstacle and safety information emerged as an important expectation, particularly regarding the dynamic nature of indoor spaces. Participants valued information about obstacles and how to avoid them. One expressed satisfaction for being told to *be careful of the obstacles on the way* (P002) while some raised concerns about how the system would *adapt quickly to understanding when...* objects are moved around. This suggests that PVI expect spatial information from COI agents to reflect the current state of the environment.

B. RQ2: How would people with visual impairments like to use conversational interaction for indoor route learning?

1) Journey Preplanning: A striking finding is the strong emphasis participants placed on using the COI app before undertaking a journey. Being able to plan before traveling will enable PVI to build cognitive maps of the environment. This was clearly articulated by one of the participants who said: *It will help before one embarks on a journey, you can have an idea of the path, and how to get there.* (P002). This remote environmental awareness capability with conversational agents extends to exploration of POIs and querying for their existence in an interactive manner: *even without standing up from my seat, I can know if a particular object or location exists, and if it exists, it can also help me to know how much effort to take me to get there* (P002). Some of the participants expressed satisfaction with knowing what to expect by simply interacting with the COI agent, much more like with human assistance: *knowing what to expect when going... before going to a place* (P003). They also commented on the advantage of being aware of certain information from COI aids that would not have otherwise been available: *Even for people..., who can see, you may not notice some of the information it has given you, you may not even have noticed them or even pay attention to them.* (P001). The COI agent offers a fundamentally different experience from conventional prejourney tools, which lack adaptability and flexibility to user requests [8].

2) Duality Modes, Active-Journey Error Correction and Integrated Navigation: Subjects articulated a clear desire for the COI agent to serve both planning and real-time goals, and critically, to integrate with existing mobility aids. Participants (P002, P003, P005, and P007) expressed satisfaction with the dual role—the possibility of using the app for remote planning and in-situ navigation: *It can give information beforehand and when you are in the process of carrying out the action.* (P005). A real-time use case could even be for when individuals at certain indoor locations (e.g., at the reception) wanting to find directions to a destination of interest. They would simply interact with the agent in such situations and receive similar assistance as that from humans: *I have gone to a number of places where you don't know where certain things are, where certain things are arranged, or even where the restroom is; you will be asking people around, but with this, it will guide you.* (P001).

Some individuals specifically noted the value of the app for getting directions: *It is also good for a shopping mall; you want to shop and you are looking for the directions ...* (P004); *It would be nice in navigating new places, new environment like malls, indoor marketplaces, new spaces... you have never been before, you are looking for a particular location inside that building.* (P006) and *I like to use this application in a complex environment.* (P005). They also want to use the COI app for wayfinding error corrections as part of real-time navigation support. This reflects a practical case when people are lost or temporarily disoriented in buildings and would like to regain their orientation: *I also find the part that you enter a wrong direction and it guides you back... so it doesn't waste your time.* (P001).

Most of the subjects explained the need to integrate the COI app with existing mobility systems: *it can not be standalone. It will need other mobility aids. Integrating this with a mobility aid will make navigation seamless.* (P002). One participant said *I will use this to plan and then use the mobility aid for the real-time movement.* (P003). Another PVI further explained that *if you can combine it with..., it would just be like okay, it gives you an overview of the route first, and you start walking, and it kind of updates you as you work. So I think it would be very good.* (P007). All these suggest that PVI see the conversational agent as a builder of a cognitive map that should complement other mobility tools.

3) Diverse Indoor Environments and Application Scenarios: All seven participants discussed other potential use cases, indicating a remarkably broad applicability. Although the most consistently recurring scenario was navigating unfamiliar environments (P001, P002, P003, P005, P006, P007), participants also identified specific indoor venues including shopping malls, subway stations (P004, P006, P007), workspaces (P006), university campuses (P007), or finding location of washrooms, cardio stations, cafetaria, gyms and fitness centres (P001, P007) within buildings. Based on their prior working experience, one participant offered a particularly "rare" application in maintenance and inventory management for finding missing items: *I have seen where some maintenance organizations are unable to use some spare parts because they are unable to find their location, and because they are unable to find them, they won't be able to use them until those spare parts are expired.* (P004). Some also attested to the value of the COI agent in familiar environments: *even places where you have been, but you need direction..., to the same place.* (P006) and even for the sighted: *That information will help both somebody with visual impairments and somebody who can really see.* (P001).

C. RQ3: What challenges/reservations might PVI have with the use of conversational interaction for indoor wayfinding assistance?

Despite an overwhelmingly positive reception from participants, particularly on the perceived usefulness and enjoyment of COI for indoor navigation experience as shown in Figure 2, participants identified several key issues that could impede the effective use of COI for indoor navigation.

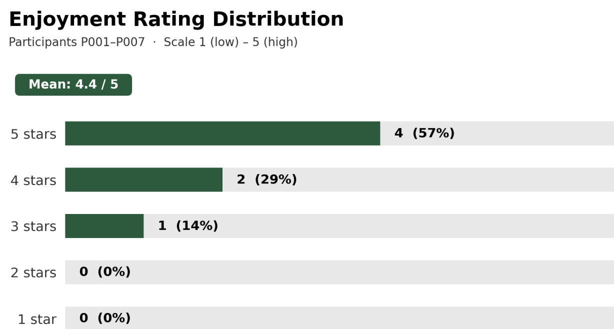

Figure 2. Participant enjoyment rating with conversational interaction: GeoChatre.

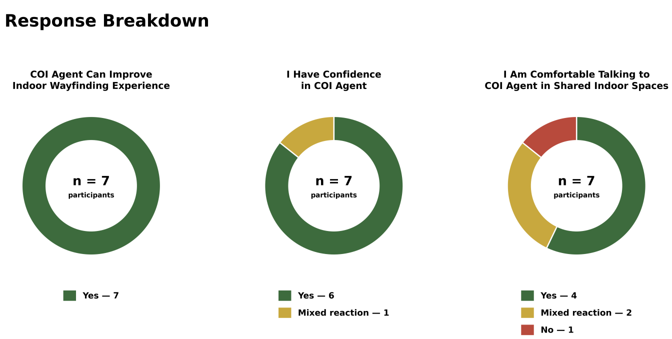

Figure 3. User perception of conversational wayfinding agent.

1) Accent and Speech Recognition Barriers: A few subjects had challenges with the speech recognition. This likely results from accents, dialectal variation, or low-resource language influences that affected their interaction with the app. One participant framed this as an equity issue and said *it should be accessible for all people irrespective of their accents.* (P005). There appears to be a sort of semantic substitution: *It's misplacing some words for other words, like, restroom for restaurants.* (P005). This finding carries significant weight, given that visual impairments may necessitate voice-based interactions. As depicted in Figure 3, the "accent-bias" elicited mixed reactions, affecting user confidence in the agent.

2) Privacy and Social Concerns in Indoor Public Spaces: A nuanced challenge emerged around the social dynamics of speaking to a COI agent in public. While several participants were generally comfortable interacting with GeoChatre, some provided insights into the situational nature during such interactions, distinguishing between socially neutral destinations: *I don't mind if people hear that I need to go to the cafeteria...* and socially sensitive ones: *But ... stuff like the washroom or even like, yeah, like even I need the bin, well, nobody needs to know that I'm going to throw my dirty t-shirt in the bin.* (P007). Some proposed solutions, such as coded destination options: *it would be nice to just use your phone and be able to click. Or option one, washroom; ...* (P007, P006) to allow navigation requests without explicitly disclosing destinations. In other words, participants want COI systems to automatically prompt with number-coded options for destinations or support the customization of location labeling when users explicitly request such interactions. These findings reveal the desire and challenge to control the information that becomes publicly audible through voice-based interactions for wayfinding in shared indoor environments.

3) Noise and Environmental Interference: Some participants (P003, P007) raised concerns about performance in noisy environments. They expressed discomfort using the COI agent in large indoor spaces *because there could be noise, like, interference, and it might not be able to, like, pick it.* (P003). Another participant drawing on experience with voice assistants: *if you're like Siri or all the vocal assistants, they tend to glitch when there are many noises.* (P007). This challenge is particularly critical because the environments where PVI might most need wayfinding assistance—shopping malls, subway stations, university buildings—are often among the noisiest. In those situations, PVI might be required to speak louder, which is often undesirable: *if you yell out loud and there's a lot of people.* (P007) while planning navigation to socially sensitive areas (e.g., washrooms).

VI. Discussion

A. Multi-Layered Spatial Information Requirements

Our findings from this study suggest that COI systems for indoor wayfinding should deliver spatial information in an integrated manner, combining steps, landmarks, turn-by-turn directions, and obstacle warnings. By leveraging the reasoning capabilities of LLMs, this information should be delivered progressively. In one of the accounts from the participants, they like that GeoChatre *did not output all those details* (P006) at once but provided information based on contexts and as requested.

There is also a strong desire for conversational interaction agents to support both journey planning and in-situ navigation. Planning would allow users to remotely explore the environments, assess effort, and build mental models of routes before traveling. This knowledge can be further "rebuilt" at the sites (e.g., within the building) just before navigation, or even during wayfinding, to self-correct from disorientation. Additionally, the consistent framing of conversational agents as complementary aids by PVI means they should be designed and used alongside existing mobility tools.

B. Contextual and Progressive Information Delivery

Results revealed that participants value the COI agent for its mode of information delivery, which was both contextual—

tailored to the request—and progressive—structured from overview to detail. This finding aligns with prior research on the standard principles of presenting information—overview-first, and detail-on-demand [35][36]. Participants captured both dimensions of information presentation: *in the first go, it did not put all those details, which might not be necessary until if I have to ask, and when I want details, it gives me details . . .* (P006); *when I asked for landmarks, it provided landmarks . . .* (P007). This combined delivery mode appears to support route recall as subjects were able to recount spatial elements and describe the routes after they interacted with the agent.

C. Accessibility and Inclusivity Implications

For PVI who rely on voice as one of their primary input modalities, speech recognition failure is technically equivalent to an inaccessible interface. One explanation for this recognition issue lies in the composition of COI training data [37]. Although many AI models are trained on vast datasets, these often lack linguistic and cultural diversity because they are disproportionately composed of Western-centric data [16]. COI systems should therefore be trained on diverse accents from the outset, with strong consideration for potential influences from low-resource languages [37].

The findings around social comfort suggest that systems should incorporate privacy-aware features such as coded POIs. The distinction between socially neutral and sensitive locations should inform interaction designs that preserve the privacy of use in shared spaces. In addition, the recurrent feedback that COI can benefit sighted individuals as well as PVI promotes a universal design that could enlarge the user base and reduce stigma associated with such assistive technologies. In deed, participants expressed strong willingness to adopt the technology, with one stating *I would just use it all the time. Because it's super practical . . .* and describing it as something that *could literally change my life.* (P007).

VII. LIMITATIONS

While this study provided insights into the expectations of PVI, desired use scenarios, and reservations regarding "conversational" indoor wayfinding, several limitations should be acknowledged. First, the work drew on interviews with seven participants. This might limit the generalization of findings to the broader community of PVI. Future studies should recruit a larger sample. In addition, the present study evaluates spatial knowledge acquisition rather than actual navigation performance. While our results demonstrate that participants can form cognitive maps of the indoor environment using COI, we did not measure how well this knowledge translates to successful real-world indoor wayfinding. This will be addressed in a planned follow-up study involving in-situ navigation tasks with PVI participants. Furthermore, analysis of results relies on self-reported data, so there might be differences between the reports of participants and their actual behaviour in real-world settings.

VIII. CONCLUSION AND FUTURE WORK

Our study on conversational interaction for indoor wayfinding assistance shows that people with visual impairments have multi-dimensional expectations. PVI desire contextualized and multilayered spatial information and the protection of personal privacy when interacting with COI agents. They prefer using COI systems flexibly across planning and in-situ navigation, in diverse indoor environments, and as a complement to existing mobility aids. Generally, "conversational wayfinding" is perceived as promoting independence while offering cross-ability benefits. However, there remain significant problems, including "linguistic-bias", privacy concerns, and environmental interference in crowded environments. Future research should pursue privacy-preserving interaction designs and scalable deployment strategies to advance this emerging technology toward real-world impact for indoor navigation.

ACKNOWLEDGMENT

This research was supported by NSERC Discovery Grant. The authors thank the participants for their time and valuable insights.

REFERENCES

[1] C. Engel et al., "Travelling more independently: A requirements analysis for accessible journeys to unknown buildings for people with visual impairments," in *Proceedings of the 22nd International ACM SIGACCESS Conference on Computers and Accessibility*, 2020, pp. 1–11.

[2] W. Jeamwatthanachai, M. Wald, and G. Wills, "Indoor navigation by blind people: Behaviors and challenges in unfamiliar spaces and buildings," *British Journal of Visual Impairment*, vol. 37, no. 2, pp. 140–153, 2019.

[3] A. T. Parker et al., "Wayfinding tools for people with visual impairments in real-world settings: A literature review of recent studies," in *Frontiers in Education*, Frontiers Media SA, vol. 6, 2021, p. 723 816.

[4] X. Tang, A. Abdolrahmani, D. Gergle, and A. M. Piper, "Everyday uncertainty: How blind people use genai tools for information access," in *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, 2025, pp. 1–17.

[5] R. Adnin and M. Das, "I look at it as the king of knowledge: How Blind People Use and Understand Generative AI Tools," in *Proceedings of the 26th International ACM SIGACCESS Conference on Computers and Accessibility*, 2024, pp. 1–14.

[6] J. Choi, D. Choi, S. Jeong, H. Hong, and J. Seering, "How far i'll go: Imagining futures of conversational ai with people with visual impairments through design fiction," *arXiv preprint arXiv:2510.12268*, 2025.

[7] P. Karmaker, D. Korre, M. H. u. Rehman, and M. Khodadadzadeh, "AI-Enhanced Landmark Recognition For Self-Guided Tour Application Using Large Language Models," in *Adjunct Proceedings of the 27th International Conference on Mobile Human-Computer Interaction*, 2025, pp. 1–5.

[8] Y. Kaniwa et al., "Chitchatguide: Conversational interaction using large language models for assisting people with visual impairments to explore a shopping mall," *Proceedings of the ACM on Human-Computer Interaction*, vol. 8, no. MHCI, pp. 1–25, 2024.

[9] M. Matei, L. Alboaie, and A. Iftene, "Safety navigation using a conversational user interface for visually impaired people," *Procedia Computer Science*, vol. 207, pp. 1164–1173, 2022.

[10] L. Ran, S. Helal, and S. Moore, "Drishti: An integrated indoor/outdoor blind navigation system and service," in *Second IEEE annual conference on pervasive computing and communications, 2004. Proceedings of the*, IEEE, 2004, pp. 23–30.

[11] D. Sato et al., "Navcog3: An evaluation of a smartphone-based blind indoor navigation assistant with semantic features in a large-scale environment," in *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility*, 2017, pp. 270–279.

[12] D. Sato et al., "Navcog3 in the wild: Large-scale blind indoor navigation assistant with semantic features," *ACM Transactions on Accessible Computing (TACCESS)*, vol. 12, no. 3, pp. 1–30, 2019.

[13] A. Bokolo Jnr, "Examining the use of intelligent conversational voice-assistants for improved mobility behavior of older adults in smart cities," *International Journal of Human–Computer Interaction*, vol. 41, no. 7, pp. 3867–3888, 2025.

[14] H. Cuayáhuitl, N. Dethlefs, K.-F. Richter, T. Tenbrink, and J. A. Bateman, "A dialogue system for indoor wayfinding using text-based natural language.," *Int. J. Comput. Linguistics Appl.*, vol. 1, no. 1-2, pp. 285–304, 2010.

[15] P. Dang, J. Zhu, W. Li, and J. Lai, "A large language model-based agent for wayfinding: Simulation of spatial perception and memory," *Cartography and Geographic Information Science*, vol. 52, no. 4, pp. 350–369, 2025.

[16] K. Rahimi, M. W. Haque, S. Dasgupta, and M. Rahman, "Vision-based localization and llm-based navigation for indoor environments," *arXiv preprint arXiv:2508.08120*, 2025.

[17] M. Manzoni, S. Mascetti, D. Ahmetovic, R. Crabb, and J. M. Coughlan, "Mapio: Embodied interaction for the accessibility of tactile maps through augmented touch exploration and conversation," *arXiv preprint arXiv:2412.00946*, 2024.

[18] R. G. Long and E. Hill, "Establishing and maintaining orientation for mobility," *Foundations of orientation and mobility*, vol. 1, p. 45, 1997.

[19] J. Ducasse, A. M. Brock, and C. Jouffrais, "Accessible interactive maps for visually impaired users," in *Mobility of Visually Impaired People: Fundamentals and ICT Assistive Technologies*, Springer, 2017, pp. 537–584.

[20] J. Guerreiro et al., "Virtual navigation for blind people: Transferring route knowledge to the real-world," *International Journal of Human-Computer Studies*, vol. 135, p. 102 369, 2020.

[21] J. R. Blum, M. Bouchard, and J. R. Cooperstock, "What's around me? spatialized audio augmented reality for blind users with a smartphone," in *International Conference on Mobile and Ubiquitous Systems: Computing, Networking, and Services*, Springer, 2011, pp. 49–62.

[22] M. Bleau, C. van Acker, N. Martiniello, J. P. Nemargut, and M. Ptito, "Cognitive map formation in the blind is enhanced by three-dimensional tactile information," *Scientific Reports*, vol. 13, no. 1, p. 9736, 2023.

[23] A. Brock, "Interactive maps for visually impaired people: Design, usability and spatial cognition," Ph.D. dissertation, Université Toulouse 3 Paul Sabatier, 2013.

[24] K. Papadopoulos, E. Koustriava, and M. Barouti, "Cognitive maps of individuals with blindness for familiar and unfamiliar spaces: Construction through audio-tactile maps and walked experience," *Computers in Human Behavior*, vol. 75, pp. 376–384, 2017.

[25] D.-R. Chebat, S. Maidenbaum, and A. Amedi, "The transfer of non-visual spatial knowledge between real and virtual mazes via sensory substitution," in *2017 International Conference on Virtual Rehabilitation (ICVR)*, IEEE, 2017, pp. 1–7.

[26] C.-H. Chen, M.-F. Shiu, and S.-H. Chen, "Use learnable knowledge graph in dialogue system for visually impaired macro navigation," *Applied Sciences*, vol. 11, no. 13, p. 6057, 2021.

[27] B. Gamage, T.-T. Do, N. S. C. Price, A. Lowery, and K. Marriott, "What do blind and low-vision people really want from assistive smart devices? comparison of the literature with a focus study," in *Proceedings of the 25th international ACM SIGACCESS conference on computers and accessibility*, 2023, pp. 1–21.

[28] L. Clark et al., "What makes a good conversation? challenges in designing truly conversational agents," in *Proceedings of the 2019 CHI conference on human factors in computing systems*, 2019, pp. 1–12.

[29] G. Jain et al., "I want to figure things out: Supporting exploration in navigation for people with visual impairments," *Proceedings of the ACM on Human-Computer Interaction*, vol. 7, no. CSCW1, pp. 1–28, 2023.

[30] G. Jain, L. Findlater, and C. Gleason, "Scenescout: Towards ai agent-driven access to street view imagery for blind users," *arXiv preprint arXiv:2504.09227*, 2025.

[31] J.-E. Kim, G. Sahas, and M. Bessho, "Toward assisting blind individuals in exploring unfamiliar indoor environments using multimodal llm and smartphone lidar," in *2025 IEEE International Conference on Consumer Electronics (ICCE)*, IEEE, 2025, pp. 1–6.

[32] J. Balata, Z. Mikovec, and P. Slavik, "Conversational agents for physical world navigation," in *Studies in Conversational UX Design*, Springer, 2018, pp. 61–83.

[33] A. W. Siegel and S. H. White, "The development of spatial representations of large-scale environments," in *Advances in child development and behavior*, vol. 10, Elsevier, 1975, pp. 9–55.

[34] S. Werner, B. Krieg-Brückner, H. A. Mallot, K. Schweizer, and C. Freksa, "Spatial cognition: The role of landmark, route, and survey knowledge in human and robot navigation1," in *Informatik'97 Informatik als Innovationsmotor: 27. Jahrestagung der Gesellschaft für Informatik Aachen, 24.–26. September 1997*, Springer, 1997, pp. 41–50.

[35] B. Shneiderman, "The eyes have it: A task by data type taxonomy for information visualizations," in *The craft of information visualization*, Elsevier, 2003, pp. 364–371.

[36] B. Craft and P. Cairns, "Beyond guidelines: What can we learn from the visual information seeking mantra?" In *Ninth International Conference on Information Visualisation (IV'05)*, IEEE, 2005, pp. 110–118.

[37] D. I. Adelani et al., "Irokobench: A new benchmark for african languages in the age of large language models," in *Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, 2025, pp. 2732–2757.