Applied Human Factors and Ergonomics (AHFE2025), Vol. 199, 2025, 978–986
https://doi.org/10.54941/ahfe1006910

Design and Structure of Sightseeing Verbal Maps: A Case Study in Shanghai, China

Tianyu Li[1] and Takashi Uchida[2]

[1]Graduate School of Engineering, Osaka City University, Osaka, Japan
[2]Graduate School of Engineering, Osaka Metropolitan University, Osaka, Japan

ABSTRACT

This study investigates a novel at-home simulation system designed to enhance independent mobility for visually impaired individuals. The system leverages sightseeing verbal maps, integrating spatially encoded audio guidance with contextual environmental sounds and synchronized dynamic video to generate highly immersive virtual reconstructions of navigation routes. A qualitative evaluation was conducted on a linguistically and culturally adapted prototype for Chinese users. Through semi-structured interviews with seven Chinese international students in Japan, thematic analysis revealed three central determinants of user experience: navigation accuracy, cultural engagement, and sound design. Findings underscore the necessity for precise, timely instructions, culturally resonant points of interest, and context-rich soundscapes to foster effective environmental awareness and emotional comfort. This work presents the first cross-cultural validation of sightseeing verbal maps beyond their original Japanese context, highlighting the critical importance of localized design in the scalability and adoption of assistive mobility technologies.

Keywords: Sightseeing verbal maps, Visually impaired, Universal design, Voice navigation

INTRODUCTION

Sightseeing verbal maps represent a paradigm shift in accessible tourism by enabling visually impaired individuals to engage in meaningful spatial exploration before physical visits. These auditory tools transform traditional navigation approaches by offering immersive virtual experiences that allow users to familiarize themselves with unfamiliar environments from the comfort of their homes. Through carefully crafted verbal descriptions of spatial layouts, architectural features, and cultural landmarks, the system facilitates cognitive mapping of destinations prior to actual travel. This preparatory function proves particularly valuable for non-routine environments where conventional navigation aids often fall short, as it provides opportunities for repeated mental rehearsals of routes and potential challenges. By bridging the gap between virtual preparation and real-world wayfinding, sightseeing verbal maps not only enhance spatial confidence but also reduce the anxiety typically associated with navigating complex tourist sites. The ability to preview and practice journeys in a risk-free

setting fundamentally redefines what independent mobility means for visually impaired travelers in the context of tourism.

Current assistive navigation technologies face significant limitations in infrastructure dependence and cross-cultural adaptability, particularly in unstructured environments like tourist destinations. While existing voice navigation systems (e.g., those developed for Japanese contexts) show some effectiveness for routine pathfinding, they frequently fail in regions with distinct urban geometries, pedestrian behaviour patterns, and soundscape characteristics (such as China). These cultural-geographical disparities lead to wayfinding logic errors and environmental cognition biases, severely restricting spatial autonomy for visually impaired individuals. More critically, the lack of effective pre-trip rehearsal mechanisms in current systems prevents users from safely practicing navigation strategies for unfamiliar spaces, thereby exacerbating real-world travel anxiety and risks.

This research develops sightseeing verbal maps with one core mission: to transform home environments into training grounds for real-world mobility. Our solution enables visually impaired users to: (1) Pre-emptively master routes through audio-based virtual rehearsal, building detailed cognitive maps of target locations before visitation. (2) Practice decision-making at critical waypoints (intersections, landmarks) via simulated environmental soundscapes and culturally adapted verbal cues. (3) Gain spatial confidence through repetitive, stress-free training—reducing reliance on physical infrastructure or companion guidance during actual travel.

By shifting navigation training from uncontrolled real-world settings to safe domestic rehearsal, the system addresses a fundamental limitation in assistive technology: the lack of deliberate practice opportunities for spatial learning. This pre-journey preparation paradigm represents a paradigm shift from reactive wayfinding to proactive spatial education for visually impaired travellers.

DEFINITION OF THE SIGHTSEEING VERBAL MAP

In this research, Sightseeing Verbal Maps are used to locate resources that can assist those who are visually impaired in working environments. This map combines a map of words with environmental sounds and sound effects for better spatial understanding (more details follow). Combining these elements, the system aims to increase realism and inspire people to explore their surroundings. This is accomplished using auditory information accessed at home. Voice messages within the system are categorized into three primary types: Route Guidance, Map of Words (Broad Sense), and Environmental Sounds and Sound Effects (see Figure 1).

Important regional information is provided by route guidance, which serves as a basic southern support system to allow customers to arrive safely at predetermined locations. Important for effective wayfinding are primary orientation markers, terminal navigation parameters, movement continuation or directional alteration advisories, and return path specifications (Matsumoto et al., 2018). The recent analysis employs specific

marketing techniques to generate customized route guidance strategies that meet the unique tourism needs of those who are visually impaired.

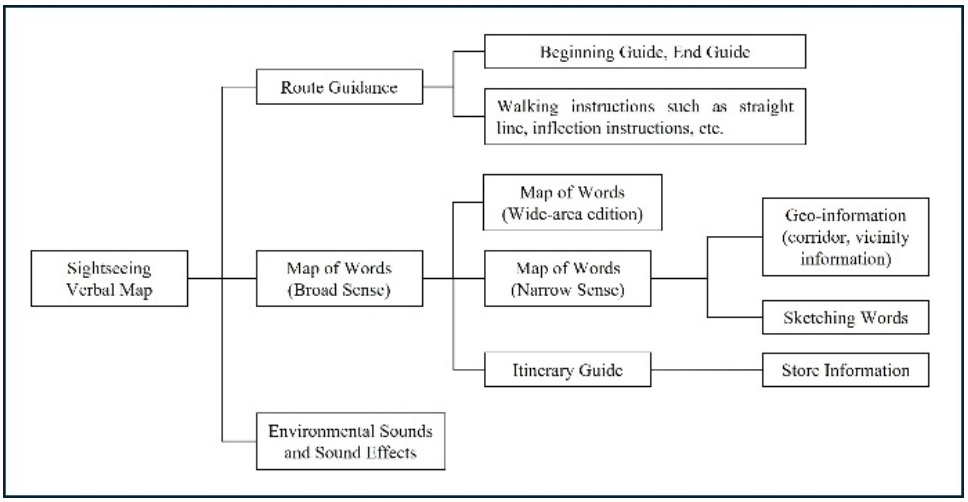

Figure 1: Overall structure of the sightseeing verbal map.

The main components of the map of words (Broad Sense) are the map of words (Wide-area edition), the map of words (Narrow Sense), and itinerary guide. The map of words (Wide-area Edition) contains two main components: geo-information and sketching words. Geo-information provides objective descriptors for tactilely discernible pathway characteristics, proximal infrastructure configurations, surface textures, and other environmental characteristics that can be verified by haptic interaction (such as white cane detection) or multisensory validation (auditory/factory cues). By establishing geographic thinking in bodily-verifiable referents, this database increases routing confidence. For phenomenological analyses of ambient sensory profiles, including auditory signatures, olfactory signatures, and taco-visual environmental characteristics, a map of words (Narrow sense) is used. Users can incorporate fleeting visual names (such as sunlight chorus sound and autumn foliage designs) into strong cognitive maps of locations by using dynamically and seasonally adjusted information to develop visual experiences in nautical settings.

The itinerary guide incorporates a sighted explorer's serendipitous discovery paradigm (Adachi et al., 2017). Outdoor means variance methods and micro-explorational vacations are simplified by creating intellectually stimulating regional versions of industrial areas. Integrated commercial establishment profiles extend beyond conventional retail descriptors to include accessibility parameters (e.g., entrance geolocation markers, tactile guidance systems) and multisensory product recommendations, prioritizing equitable environmental engagement.

When travelling, the sounds that typically occur in the area are called environmental noises. Individuals paused for 20 seconds while paying attention to the program of their vacation at bridges and other geographic

crossings. They could wait and examine the map like someone with visual impairment could do so with a standard picture. On the other hand, excellent results purposefully introduced appear to attract attention to a project's distinctive features and pique the speaker's mind. Sound effects specifically focused on auditory cues to enhance the sense of presence and the emotional connection with the environment (Takahashi et al., 2013).

METHODOLOGY

The process employed in this research was described in detail in Figure 2.

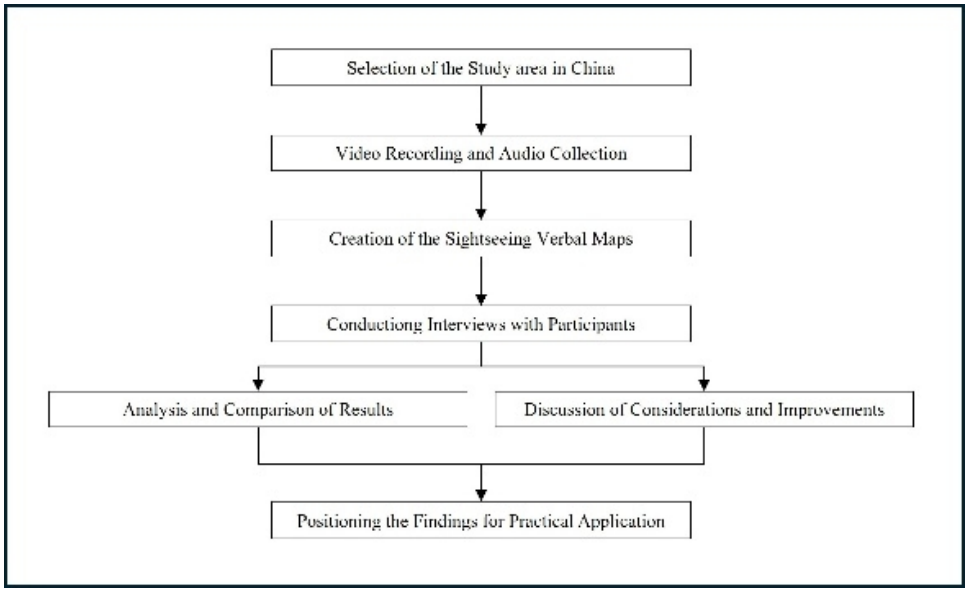

Figure 2: Research process.

This research begins by selecting a specific study area in China, aligning its location with the project's core objectives, and conducting a preliminary review of the site's cultural and aesthetic qualities using captured video and digital imagery. The natural and cultural components gathered are then synthesized to create sightseeing verbal maps. To evaluate these maps, interviews with participants are conducted, both in-person and via Zoom, to collect feedback on their usability and overall effectiveness. This user input directly informs subsequent revisions and refinements to the maps. Furthermore, a comparative analysis is undertaken to contrast the assessment results from China with those from Japan. The verbal maps are then field-tested in China, leading to major adjustments and informed discussions on potential improvements. Ultimately, the findings of this study are positioned to serve as a practical foundation for the beneficial application and further development of sightseeing verbal maps within China.

In contrast to Japan, Shanghai, a political area, is frequently the second stopover for some foreign visitors. Shanghai's Bund Place is a must-see for visitors because of its diversity of flowers, walking lines, and historical places.

Table 1 contains a succinct description of the empirical page in Shanghai, China.

Table 1: Study area information.

Location	Shanghai, China
Access	The People's Square Metro Station
Characteristic	Tourist attractions
Clientele	A broad range from youth to the elderly
Distinctive environmental sounds	The sound of the sightseeing trams

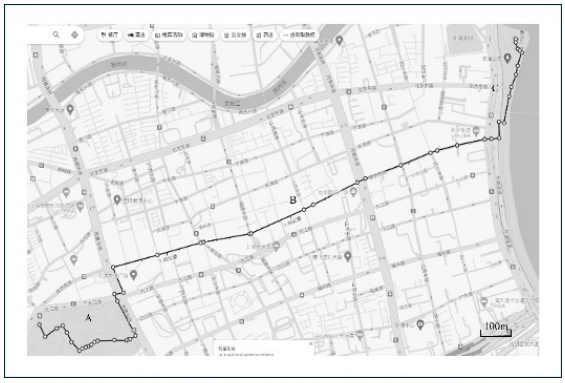

Figure 3: Walking course (Modified map of Shanghai based on google maps with added study area boundaries.).

Figure 3 shows the working course, which is 2. 8 kilometres long. Keep Exit 9 of the nearest Shanghai Metro People's Square Station as the starting place and the Bund Historical Memorial Hall (Shanghai People's Heroes Memorial Tower) as the endpoint. People's Park, Nanjing Road Pedestrian Street, and the historic Bund area are just a few feet apart. This vacation gives you access to People's Park's natural surroundings, Nanjing Road Pedestrian Street's energetic environment, the historic Bund area's historic buildings, and the stunning scenery along the Huangpu River.

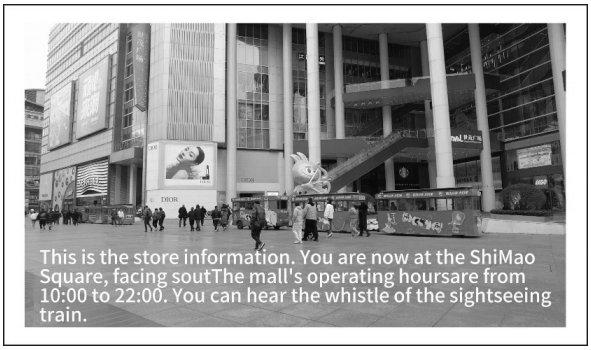

Figure 4: Screenshot of the sightseeing verbal map.

Figure 4 shows an animated screenshot of the sightseeing verbal map as viewed by study participants during the interview survey.

PARTICIPANT RECRUITMENT

After informing potential participants about the option to conduct the experiment either in person or remotely, it was determined that 7 individuals were able to participate remotely via Zoom. The participant group consisted of 4 males and 3 females, with an age distribution of 5 participants in their 20s and 2 participants in their 30s.

Given that the use of the remote conferencing tool Zoom has been confirmed as an effective means of conducting meaningful interviews, this experiment primarily utilized Zoom for data collection. This approach ensured flexibility and accessibility, allowing for a diverse range of participants to contribute to the study despite geographical constraints (Ishii et al., 2022). In this study, Zoom was primarily used to gather data. Despite physical restrictions, Zoom allowed diverse participants to contribute to the study while maintaining richness and convenience.

To mitigate potential biases and improve response quality, prior to commencing the interview, the following information was communicated: (1) the interview is anonymous; (2) data will be used solely for academic research; (3) there are no right or wrong answers; (4) participation is voluntary and randomized; (5) individual responses will not be used to identify any participants. This study adhered to the ethical guidelines of the institutional review board (Research Ethics Review Committee, Graduate School of Engineering, Osaka Metropolitan University). All interview survey procedures were in accordance with ethical standards. For participants joining the experiment remotely, the interview session was conducted over approximately 100 minutes per participant. The procedure was structured as follows: (1) Initial Setup and Explanation (Approx. 20 minutes); (2) Virtual Walking Experiment in the People's Park–the Historic Bund Area (Approx. 25 minutes); (3) Division of the Course and Interview (Approx. 40 minutes); (4) Resumption of the Virtual Walking Experiment and Interview (Approx. 15 minutes).

This structured approach allowed for a comprehensive evaluation of the sightseeing verbal maps and its impact on the participants' virtual experience, while also ensuring that their feedback was captured at key points during the experiment.

RESULTS

The feedback from the interview survey was categorized into three main areas: map of word, route guidance and store information, and environmental/sound effects. The responses were further divided into positive and negative comments. Below is a summary of the findings.

Map of Words

Positive Feedback

Participants appreciated the inclusion of obstacle information in the route guidance. For example, one comment highlighted, "It was helpful to be informed about obstacles along the way."

This feature was particularly valued in the context of China's public facilities, where obstacles such as trash bins and poles are commonly placed on walkways, differing from Japan's environment.

Negative Feedback

Some participants noted that the walking course was relatively long and suggested that more detailed descriptions of the surrounding scenery should be included in the route guidance.

Adding information about the landscape, historical context, and other points of interest during the journey could enhance the experience for visually impaired users, allowing them to better enjoy the tourist attractions.

Route Guidance and Store Information

Positive Feedback

Participants found the detailed descriptions of recommended foods at local stores to be clear and helpful. Comments such as, "The specific explanations of recommended dishes were easy to understand," were common.

Some participants also enjoyed the interactions with store staff, noting that these moments added a fun and engaging element to the experience.

For first-time visitors, information about local traditional foods and cultural practices is often unfamiliar, making it essential to actively include such details in the guidance.

Negative Feedback

No significant negative feedback was reported in this category, but the emphasis was on the importance of providing more comprehensive and engaging store-related information.

Environmental Sounds and Sound Effects

Positive Feedback

Participants appreciated the immersive quality of the environmental sounds, particularly the sound of the tram-like sightseeing vehicles on Nanjing Road Pedestrian Street. Comments such as, "The sound of the sightseeing tram added a sense of realism," were noted.

Negative Feedback

Some participants criticized the challenge of using their legs to learn. In Shanghai, where there are a lot of foot customers, increasing the level of economic appearance typically amplifies speech and overshadows different appearances.

To create a more traditional and fascinating good practice, it was suggested that more emphasis be placed on controlling the number of feet with several outside noises during recording.

The findings highlight both the limitations and advantages of the current system, particularly in terms of cultural adaptability, obstacle recognition, and interactive features. Enhanced spatial cognition, accurate environmental information, and balanced feedback mechanisms could significantly improve the user experience for individuals with visual impairments. These improvements would facilitate more effective navigation and greater independence in unfamiliar environments.

DISCUSSION

The sightseeing verbal maps have some problems, but they have the potential to help visually impaired people locate historic places. It struggles to make powerful system changes when a man needs to change direction or movement around moving items. It does a fantastic job of describing moving objects like lines and garbage cans. In a loud environment, excellent quality likewise becomes problematic. Important audio signals are often drowned out by background noises, such as speaking or customers, making it more difficult for users to adhere to guidelines.

Social speech is another area that needs improvement. The sightseeing verbal maps need more translation to fully satisfy the buyers' needs in different areas. The system considers these disparities to be successful, for instance, because various people's walking, talking, or facing obstacles in different places can be very different.

To improve sound control, expand cross-cultural tests, and maintain the system working well in different settings in the future, attention should be drawn to improving music balancing and expanding cross-cultural testing. By solving these concerns, the plan might become more dependable and adaptable, enabling visually impaired people to explore their surroundings more easily and by their needs.

CONCLUSION

This research validates China's sightseeing verbal maps' success by demonstrating its potential for greater availability and immersion in earlier significant, non-traditional locations, such as tourist destinations. The chart provides visually impaired individuals with a comprehensive and interesting experience by incorporating in-depth issue information, local cultural narratives, and good fiscal sounds. Participants especially appreciated incorporating historical details, such as details of conventional foods, traditional websites, and native customs, which improved their understanding of the way of life and their knowledge of the sites' historical significance.

China's distinctive qualities in commercial settings must be addressed through social and economic adjustments. For instance, lateral logic changes were necessary to maintain densely populated areas and complex roads. Environmental soundscapes needed cautious modification to counteract the concomitant effects of polite company sounds and pedestrian traffic. These

changes help maintain the state's suitability for various options. Also, it has been demonstrated that distant monitoring systems like Zoom can carry discussions and tests very efficiently, making it possible to research in various settings or configurations.

Based on the study, China's persuasive travelling image could be used for online movie distribution. The next step will be the release of the enhanced sightseeing verbal maps video on well-known Chinese online platforms, the enhancement of the financial looks used, and the introduction of customer feedback. Researchers will be able to gather feedback from audiences to determine the site's ability to help users build mental pictures of certain channels and locations. The feedback will also affect more changes, such as improving audio clarity and enhancing cultural relevance.

By allowing online adventures from home, this cutting-edge strategy aims to bridge the gap between outside investigation and interior style and promote the development of visually impaired individuals. By offering a healthy, interesting way to identify oneself with unfamiliar roads and problems, the sightseeing verbal maps help consumers become more comfortable and independent. It also establishes a fresh standard for similar vacations, ensuring all historical and cultural sites are accessible. This study advances assistive technologies and creates a society where people with visual impairments may completely participate in and enjoy the beauty of their environment.

REFERENCES

Adachi Y., Uchida T., Suga Y., and Sawada Y. (2017). Expansion of Path Guidance and Walking Support Navigation Using NFC to Assist the Willingness of Visually Impaired People to Go Out. Proceedings of the 37th Traffic Engineering Research Conference, pp. 367–372 (in Japanese), https://ndlsearch.ndl.go.jp/books/R000000004-I028613573.

Ishii M., Matsumoto H., and Uchida T. (2022). Remote Zoom Experiment for Practical Deployment of "Word Tourist Map". Journal of Traffic Engineering, Vol. 8, No. 2 (Special Issue A), pp. A_310–A_318 (in Japanese). https://doi.org/10.14954/jste.8.2_A_310.

Matsumoto H., Uchida T., and Sakaguchi R., (2018). Expansion of Support Messages for Visually Impaired People's City Walks and Examination of Virtual Walk Experiment Methods. Proceedings of the 38th Traffic Engineering Research Conference, pp. 243–248 (in Japanese). http://id.ndl.go.jp/bib/029905653.

Takahashi S., Negi K., and Uchida T. (2013). Guidelines for Describing Landmarks in Audio AR Apps for Visually Impaired People. Proceedings of the 33rd Traffic Engineering Research Conference, pp. 509–514 (in Japanese). http://id.ndl.go.jp/bib/025569003.