wayfinding without vision: Learning real and virtual environments using dynamically-updated verbal descriptions

Nicholas A. Giudice

Department of Psychology
University of California Santa Barbara
Santa Barbara, CA, U.S.A. 93101
Tel: +1 805-893-7274
Email: giudice@psych.ucsb.edu

Abstract.
This paper describes the details and rationale for a dynamically-updated, context-sensitive verbal interface, which has been developed to support learning and wayfinding of indoor environments. The purpose of this interface is to serve as a non-visual substitute for apprehending critical navigational information about the environment and facilitating efficient travel by blind navigators. The efficacy of this interface in supporting accurate environmental learning and wayfinding of real and virtual environments was demonstrated with blind and low-vision participants in two experiments. In both studies, participants were able to effectively use the verbal interface to freely explore novel environments and built up a cognitive map which supported subsequent wayfinding tasks. The results of these studies demonstrate that blind navigation is not solely based on route knowledge, as has been argued in the literature. The findings also help to guide future design specifications for speech-based navigation technology.

Keywords: blind navigation, dynamically-updated verbal descriptions, navigation technology

1. Background

The goal of this research was to develop techniques to enhance indoor navigation by blind individuals. This paper first discusses the development of a non-visual interface to describe layout geometry. Two experiments are then discussed which investigate the efficacy of this interface for supporting effective learning and wayfinding of complex, large-scale real and virtual indoor layouts by blind and low-vision navigators.

There are three major differences between the current research and most of the existing literature addressing blind spatial abilities. First, the emphasis of our speech-based interface is to provide orientation information which can support indoor navigation. There is a long history of electronic travel aids to help facilitate safe and efficient travel by blind individuals (see Giudice & Legge, in press for a review). However, most of this technology has focused on providing mobility information about obstacle detection and is best suited for outdoor usage. With the advent of accessible GPS-based navigation systems, it is now also possible to provide real-time, updated information about a person's position and orientation in the environment as they navigate (see NFB Access Technology, 2006 for a review of three commercially-available speech-based systems). Unfortunately, GPS-based devices only work outside and there is no comparable system for conveying updated orientation information during indoor navigation. The preliminary experiments discussed here are a subset of studies, not yet reported, on blind navigation (Giudice & Legge, In Preparation). This work is part of a larger project designed to develop an indoor navigation system that can provide precise spatial information about a person's position and orientation during indoor navigation and which seamlessly integrates with GPS and other localization/orientation technologies for outdoor travel (see The wayfinding group's website at www.wayfinding.org for details).

Second, the information provided by the verbal display discussed here, as well as the spatial tasks investigated, differ from previous work. The research addressing orientation technology for the blind, e.g. Talking Signs (Crandall *et al.*, 1999; Marston, 2002) and GPS-based navigation systems employing speech or spatialized sound displays (Loomis *et al.*, 1998; Loomis *et al.*, 1994; Petrie, 1996), convey information about the distance and direction of landmarks or decision points in the environment in order to support route following. By contrast, the current research employs verbal descriptions of layout geometry with the goal of facilitating free exploration (navigation without reference to routes) and wayfinding behavior. The notion of using verbal information to describe layout geometry rather than specific landmarks and to support wayfinding rather than route navigation has not previously been studied with blind individuals.

Third, although there is much debate in the literature regarding the spatial abilities of blind and low-vision people (see Thinus-Blanc, 1997 for an excellent review), differences in spatial performance are hypothesized here as stemming from insufficient access to reliable sources of environmental information rather than any inherent deficits in spatial ability (Fletcher, 1980). Blind performance is predicted to be highly accurate in these studies as the verbal descriptions provide access to navigation-critical information which may otherwise be difficult to obtain from non-visual sensing. This result would be particularly noteworthy given that the current design requires free exploration and wayfinding rather than route following. Wayfinding and a grasp of overall layout configuration is generally considered difficult for this population; thus, blind navigation is traditionally thought to primarily rely on specified routes (see Millar, 1994 for a discussion).

At the heart of these experiments is a verbal protocol, a set of formal instructions to describe layout geometry through discrete speech-based messages that update in registration with the movement of the navigator through the space. An example of this context-sensitive, first-person verbal description is: "You are facing North at a 3-way intersection, there are hallways to your left, right and behind." If the participant made a left rotation, the message would dynamically update the original verbal description to reflect that they are now facing west, with hallways extending ahead, left and behind.

While there are several models of verbal direction giving for route navigation (Allen, 1997; Couclelis, 1996), there are no formal guidelines for conveying dynamically-updated geometric descriptions for wayfinding tasks. Several factors guided our rationale for using information about layout geometry (corridor configuration) as the mode of environmental access for these experiments. Layout geometry is unlikely to change over time, can be described using relatively few spatial primitives and is readily extracted from existing sources, e.g. hardcopy maps or AutoCAD floor plans. By contrast, landmarks and other non-geometric features are often environmentally specific, subject to transients and less amenable to description using a lexicon of basic spatial terminology. By adopting a standard order and syntax for the presentation of verbal information, it is possible to ensure that consistent, unambiguous messages are conveyed to all participants and that the verbal information is isomorphic with the physical layout being described. Attention to such factors are known to be critical for effective use of spatial language (Ehrlich & Johnson-Laird, 1982; Levelt, 1996).

Our verbal protocol first provides the user with absolute information about their facing direction; e.g., "facing north, south, etc."), followed by the type of intersection at their location ("at a 2, 3, 4-way intersection" and finally an egocentric description of the intersection geometry ("to your left is a hallway, behind is a hallway").

Since the most accurate verbal descriptions in route navigation are those that employ both absolute and egocentric terminology (Tversky, 1996), the adoption of multiple reference frames in the current research was reasoned as also benefiting spatial updating and the development of an accurate spatial representation during free exploration and wayfinding.

A consideration of spatial scale is particularly important when using language to describe environmental relations. In contrast to the spatial senses, e.g. vision and touch, where the amount of the environment perceived from a given vantage point is based on the field and depth of view of the input modality, verbal descriptions have no inherent constraints on the breadth or depth of environmental information described. While the amount and precision of verbal information inevitably varies depending on the situation, e.g. describing the layout of a room would be at a different scale and level of detail than a description of New York City, the optimal amount of the environment that should be conveyed from a given position to support large-scale navigation has not been studied using language. At one end of the continuum, a verbal description could attempt to depict a map-like overview of the space; at the other end, it could describe only the layout geometry at the wayfinder's immediate position. In this paper, the amount of the environment that is accessible to the navigator from a given vantage point is termed "verbal view-depth". Given that increasing the verbal real estate of a spatial description comes at a high cognitive cost (Mani & Johnson-Laird, 1982), three verbal view-depth conditions were used in the current experiments in order to investigate the trade-off of message length and working memory demands.

Each view-depth condition conveyed a different amount of geometric information about the participant's current location in the environment. The three conditions included Local, Maplet, and Global verbal view-depth modes, representing a range of environmental detail from conveying a description of only nearby information to an overview description of layout configuration. See figure 2 for a depiction of what would be heard from each of the three view-depth conditions.

By adopting these three levels of verbal description, we can determine the minimal information requirements that facilitate efficient verbal learning and navigation. These findings are also important in defining design specifications for future research using dynamically-updated verbal displays and the development of non-visual navigation technology. Besides its application to real-time indoor navigation systems, this verbal interface could be used for virtual learning and exploration of computer-based environments before traveling to the physical environment, something which is known to be particularly beneficial for blind navigators (Holmes *et al.*, 1996).

The efficacy of the verbal protocol used in this paper has already been demonstrated with blindfolded sighted participants exploring real building layouts (Giudice *et al.*, in press) and virtual environments (Giudice *et al.*, submitted). It is predicted that a similar high level of learning will be observed with the blind and low-vision participants in the current studies.

2. General Methods

Eight participants ran in Experiment I and fourteen in Experiment II. Approximately half were totally blind and half low-vision in each experiment. Both experiments employed the same verbal view-depth conditions, environments and procedure. During the fixed-length training phase, participants were instructed to learn unfamiliar environments and find four hidden target locations by freely exploring the layout (floors of university buildings). No route information was given, participants were simply told to try to cover the entire floor and find the four target locations during their search. In Experiment I, training involved receiving verbal descriptions at each corridor intersection (by the experimenter) while walking around the physical building, whereas training in Experiment II involved virtual navigation of computer-based environments, using a synthetic-speech interface here dubbed a virtual verbal display (VVD). Participants explored the virtual environments using the keyboard arrow keys to navigate. Verbal descriptions were given automatically upon reaching an intersection (or on request at any point). See figure 1 for an illustration of a sample experimental layout.

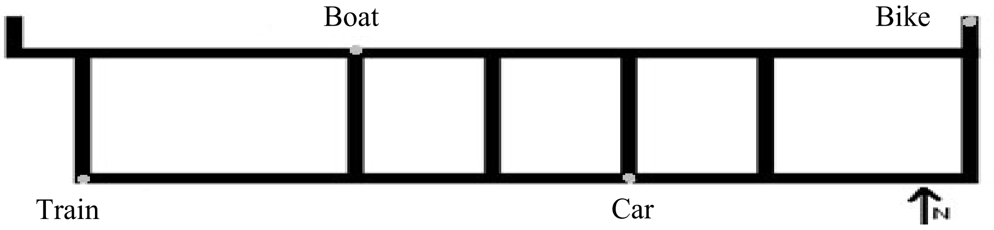

Figure 1, A sample experimental layout with targets denoted.

In the testing phase, participants had to plan and execute routes between requested target pairs. The test phase for both experiments was carried out in the real building. Thus, Experiment I testing was performed in the same environment as training but testing in Experiment II required transfer of knowledge gained from training with the VVD in virtual environments to physical navigation in the real environment.

2.1 Verbal modes.

Each verbal view-depth condition described a different level of geometric detail about corridor structure (see figure 2).

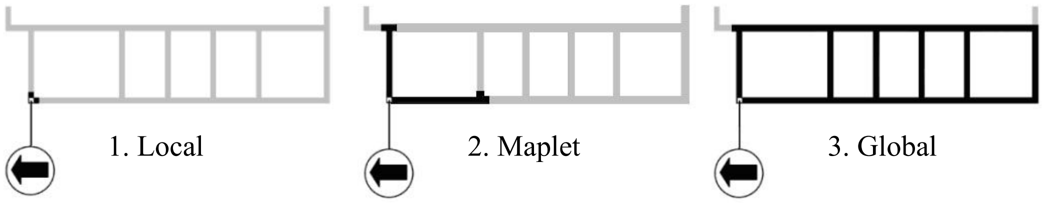

Figure 2, Verbal view-depth conditions. The circle and arrow represent the user's location and orientation in the layout. The black lines are the areas described by the verbal description and the gray lines symbolize the entire floor.

1) Local verbal mode: Describes layout geometry at the user's current position. The information provided is similar to what a blind person might receive when navigating the building with a cane or guide dog. "Facing west, at a two-way intersection, behind is a hallway, to the right is a hallway."

2) Maplet verbal mode: Includes the Local information and adds a description of the distance and geometry for all adjacent intersections. Note that the information about adjacent intersections is given as if you were walking down the hallway being described. "Facing west, at a two-way intersection, behind is a 60 foot hallway extending through a 3-way intersection to the left, to the right is a 45 foot hallway ending at a 3-way intersection."

3) Global verbal mode: Includes the Maplet information and adds a general description of the overall geometric structure of the layout. "This floor can be thought of as a 195 by 45 foot East-West rectangle intersected by four 45 foot North-south hallways…" The Global message was immediately followed by a Maplet description.

3. Results

Three key findings should be highlighted from these data. First, no statistically reliable differences were observed between totally blind and low-vision participants for any of the training or testing measures in either study, *ps* > 0.05. This finding supports the hypothesis that deficits in blind spatial performance can be mitigated by providing a reliable source of environmental information during navigation.

Second, training in real and virtual environments yielded a highly similar pattern of results. In Experiment I, participants, on average, covered 94% of the floors during their exploration, found 100% of the hidden target locations and traveled 9.71 shortest paths between targets. (The shortest path refers to the route with the minimum distance between target pairs). In Experiment II, participants covered an average of 93.3% of the floors, found 98.7% of the target locations and navigated 9.7 shortest paths between targets. Given the claims in the literature that blind navigation is based on route knowledge, the highly accurate wayfinding performance observed during training is particularly noteworthy as no routes were specified.

Third, in contrast to training, test performance was not equivalent between experiments. Although target accuracy did not differ significantly between the studies, there was a strong trend for lower performance after virtual learning for both the Maplet and Global conditions. After real building learning, participants found targets with 76% localization accuracy, whereas participants who learned in the virtual environments were only 66% accurate at finding targets at test. The poorer transfer performance after learning with the VVD suggests that physical body movement may play an important role in converting the verbal descriptions into an accurate, and accessible, spatial representation.

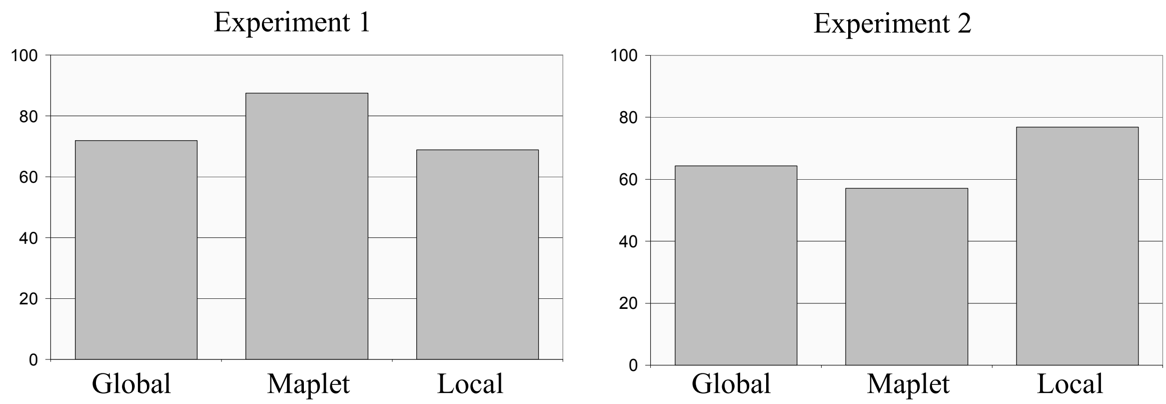

Figure 3, Graphs depicting Target Localization Accuracy by View depth condition.

Fourth, one-way repeated measures ANOVAs comparing the three verbal view-depth conditions (Global, Maplet, and Local) with each of the training and test measures in both experiments revealed no reliable differences in training or test performance by view-depth in either experiment, *ps* > 0.05. These results suggest that the minimal geometric information provided by the Local view-depth condition is sufficient to describe an environment and supports accurate spatial learning and wayfinding behavior.

4. Conclusions

These experiments represent the first known work to study whether dynamically-updated verbal descriptions support wayfinding behavior by blind individuals in large-scale layouts and to investigate whether learning with a virtual verbal display in simulated environments transfers to accurate navigation performance in the corresponding real environment. The results support the efficacy of dynamically-updated verbal descriptions as an effective non-visual mode of environmental access for blind navigators. The data from both studies provide strong evidence that access to these descriptions facilitates free exploration and supports accurate spatial learning. Specifically, performance for the Local condition proved as good as the longer messages of the Maplet and Global conditions, demonstrating that increasing access to geometric detail does not result in better learning or wayfinding performance. The current results highlight the importance of future development of speech-based navigation technology and virtual verbal displays.

Although more research is necessary to understand why transfer to real navigation is worse after learning in virtual environments vs. real environments, the data clearly demonstrate that dynamically-updated verbal descriptions are a viable foundation for a real-time system for indoor navigation.

For an indoor navigation system to be effective as a product for real-time use, we need to solve the indoor wayfinding challenge. That is, there needs to be a way to track the navigator's position and orientation in the environment as they move (much like the task of the human experimenter in Experiment 1). Determining an inexpensive method for updating a pedestrian's location and heading during indoor navigation is nontrivial. To be effective, a system would also necessitate a base map of the building environment, much like the GIS databases used with GPS outdoors. This building database would provide context-sensitive verbal descriptions, facilitate route guidance and indicate points of interest in the building. The author of this paper is part of a group investigating one solution to this indoor navigation challenge. The preliminary system uses an infrared digital sign system (DSS) to pick up passive tags placed in the environment. Once the user picks up a sign, by means of a hand-held device, the information is captured and sent to a building database. The building database works in conjunction with "Building Navigator," which gives verbal descriptions to the user about their surroundings, routing instructions or other information pertinent to their current location. For a preliminary discussion of the DSS, see (Tjan *et al.*, 2005) and for more detail on the proposed integrated system, see (Giudice & Legge, in press).

References

Allen, G. L. (1997). From knowledge to words to wayfinding: Issues in the production and comprehension of route directions. In S. C. Hirtle & A. U. Frank (Eds.), *Lecture notes in computer science* (Vol. 1329, pp. 363-372): Springer.

Couclelis, H. (1996). Verbal directions for way finding: Space, cognition and language. In J. Portugali (Ed.), *The construction of cognitive maps* (pp. 133-153). Dordrecht: Kluwer Academic Publishers.

Crandall, W., Brabyn, J., Bentzen, B. L., & Myers, L. (1999). Remote infrared signage evaluation for transit stations and intersections. *Journal of Rehabilitation Research and Development, 36*(4), 341-355.

Ehrlich, K., & Johnson-Laird, P. N. (1982). Spatial descriptions and referential continuity. *Journal of Verbal Learning & Verbal Behavior, 21*, 296-306.

Fletcher, J. E. (1980). Spatial representation in blind children. 1: Development compared to sighted children. *Journal of Visual Impairment & Blindness, 74*, 381-385.

Giudice, N. A., Bakdash, J., Z., Roy, R., & Legge, G. E. (submitted). Spatial learning and navigation using a virtual verbal display.

Giudice, N. A., Bakdash, J. Z., & Legge, G. E. (in press). Wayfinding with words: Spatial learning and navigation using dynamically-updated verbal descriptions. *Psychological Research*.

Giudice, N. A., & Legge, G. E. (In Preparation). Verbal learning and wayfinding in real and virtual environments: Extension to blind and low-vision navigators.

Giudice, N. A., & Legge, G. E. (in press). *Blind navigation and the role of technology*: to be published by John Wiley & Sons.

Holmes, E., Jansson, G., & Jansson, A. (1996). Exploring auditorily enhanced tactile maps for travel in new environments. *New Technologies in the Education of the Visually Handicapped, 237*, 191-196.

Levelt, W. J. M. (1996). Perspective taking and ellipsis in spatial descriptions. In P. Bloom, Peterson, M. A., Garrett, M. F., & Nadel, L. (Ed.), *Language and space* (pp. 77-107). Cambridge, MA: MIT Press.

Loomis, J. M., Golledge, R. G., & Klatzky, R. L. (1998). Navigation system for the blind: Auditory display modes and guidance. *Presence, 7*, 193-203.

Loomis, J. M., Golledge, R. G., Klatzky, R. L., Speigle, J. M., & Tietz, J. (1994). *Personal guidance system for the visually impaired.* Paper presented at the Proceedings of the First Annual ACM/SIGGAPH Conference on Assistive Technologies, Marina Del Ray.

Mani, K., & Johnson-Laird, P. N. (1982). The mental representation of spatial descriptions. *Memory & Cognition, 10*(2), 181-187.

Marston, J. R. (2002). *Towards an accessible city: Empirical measurement and modeling of access to urban opportunities for those with vision impairments, using remote infrared audible signage*. Unpublished doctoral Dissertation: Department of Geography, University of California, Santa Barbara.

Millar, S. (1994). *Understanding and representing space. Theory and evidence from studies with blind and sighted children.* Oxford, England: Clarendon Press.

NFB Access Technology Staff. (2006). Gps technology for the blind, a product evaluation. *Braille Monitor, 49(2)*.

Petrie, H., Johnson, V., Strothotte, T., Raab, A., Fritz, S., & Michel, R. (1996). Mobic: Designing a travel aid for blind and elderly people. *Journal of Navigation, 49*, 45-52.

The wayfinding group. 2006, Retrieved from http://www.wayfinding.org

Thinus-Blanc, C., & Gaunet, F. (1997). Representation of space in blind persons: Vision as a spatial sense? Psychological Bulletin, 121(1), 20-42.

Tjan, B. S., Beckmann, P. j., Roy, R., Giudice, N. A., & Legge, G. E. (2005). *Digital sign system for indoor wayfinding for the visually impaired.* Paper presented at the Proceedings of the First IEEE Workshop on Computer Vision Applications for the Visually Impaired, in conjunction with CVPR 2005, San Diego.

Tversky, B. (1996). Spatial perspective in descriptions. In P. Bloom, M. A. Peterson, L. Nadel & M. F. Garrett (Eds.).