

Designing Effective Synthesized Spatial Audio Cues For The Web

Tao Wang¹, Dr. Donald Patterson², Dr. David Redmiles¹

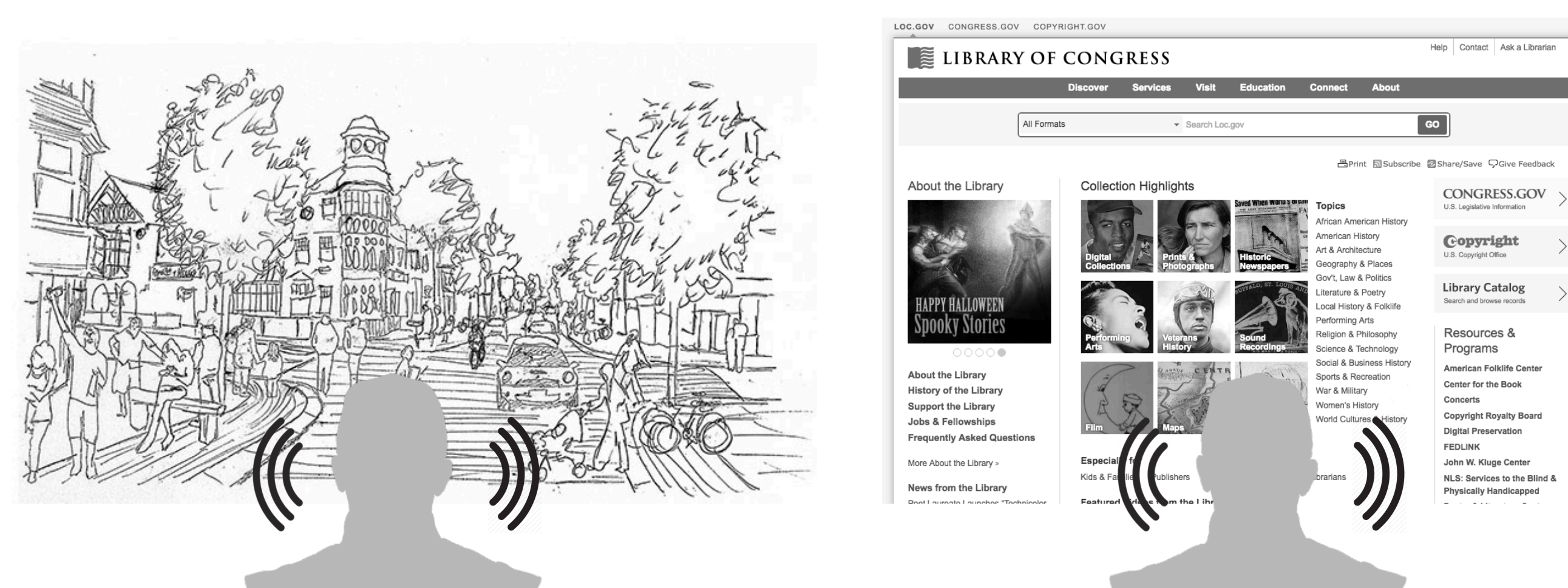
¹Department of Informatics, University of California, Irvine

²Department of Computer Science, Westmont College

INTRODUCTION

The human auditory system can localize objects and recognize movements based on sounds. It allows visually impaired people to perceive the surrounding physical environment effectively. However, this ability is not well utilized in another critical environment, the Web. Blind users often feel disoriented when navigating a web page, as screen reading software does not communicate sufficient layout information.

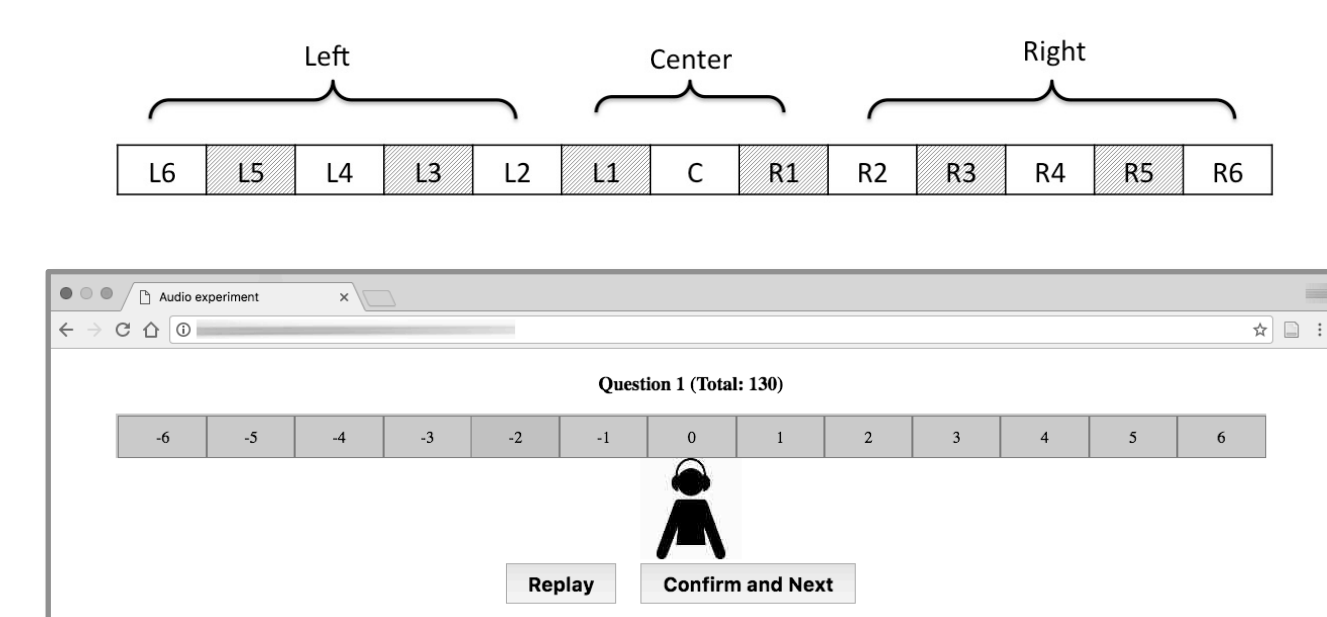
Our research attempts to improve web accessibility by conveying web page layout information to users via synthesized spatial audio cues. In order to produce effective audio cue designs, we have conducted experiments to establish baselines on how well regular users could recognize spatial audio synthesized with the Web Audio API in the horizontal audio plane. We have also explored how an audio cue's various properties affect its recognition.



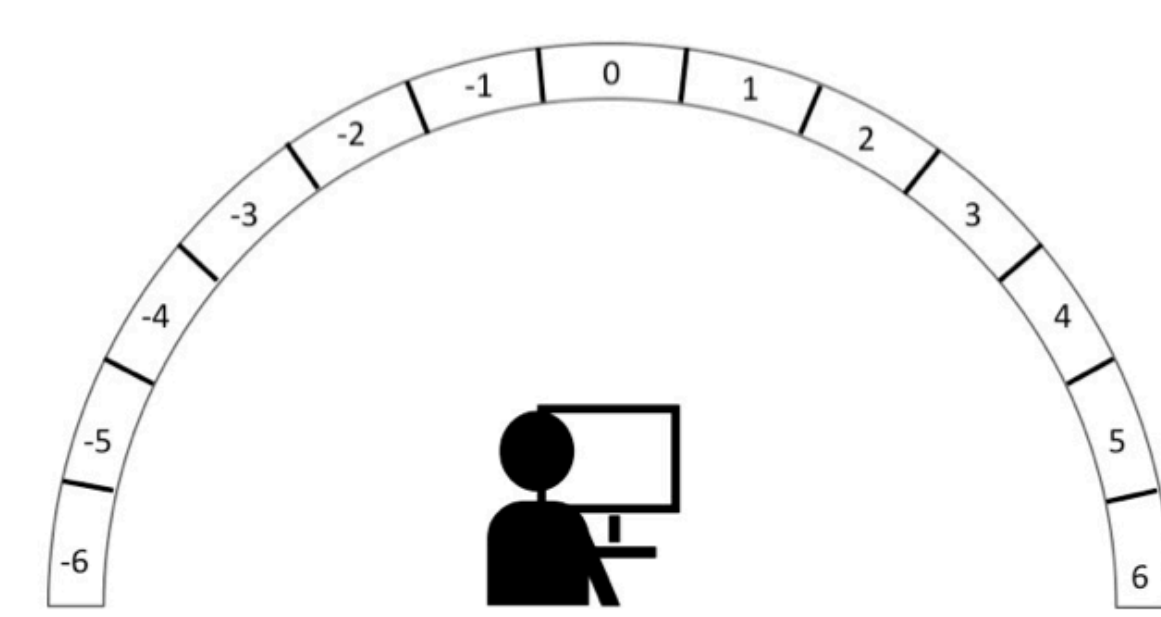
Can we hear a webpage's layout?

EXPERIMENT

A web page is divided into 13 sections horizontally. A script generates spatial audio cues dynamically during experiment trials. We tested the recognition of both stationary audio cues that are positioned in fixed locations and moving audio cues that move from one location to another. 18 sighted subjects (9F/9M, no self-reported hearing issues) participated in the experiment.



Experiment Interface



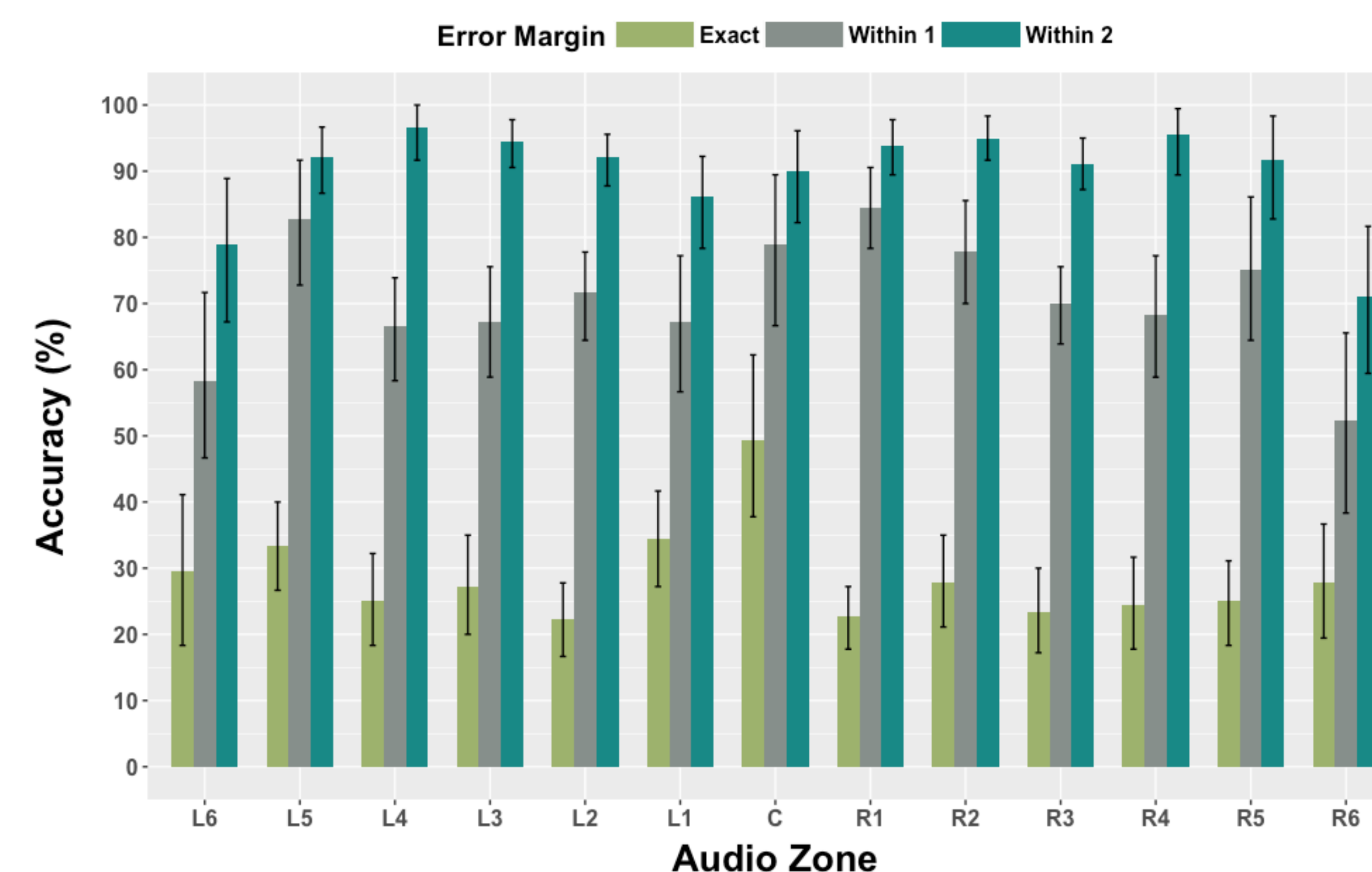
Virtual Audio Space From Above

RESULTS

Stationary Audio Cue Localization

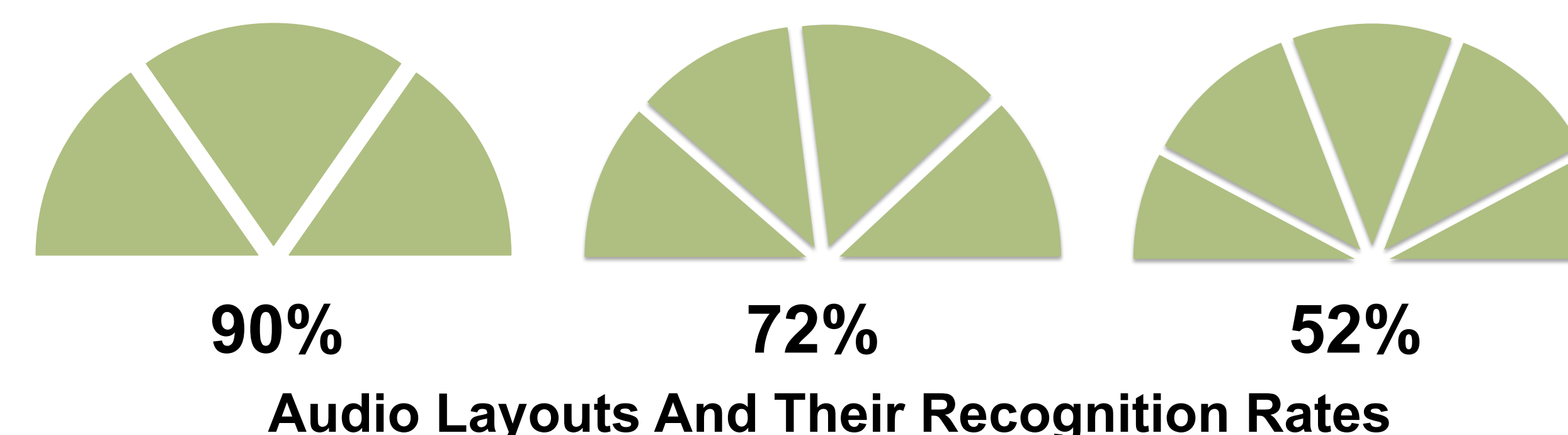
- 13 locations
- Each location was tested 10 times per subject.

Recognition Rates



- Exact location: Mean is 28% (SD=7%)
- ± 1 location: Mean is 71% (SD=9%)
- ± 2 locations: Mean is 90% (SD=7%)

Audio Layout Designs



Optimized audio locations

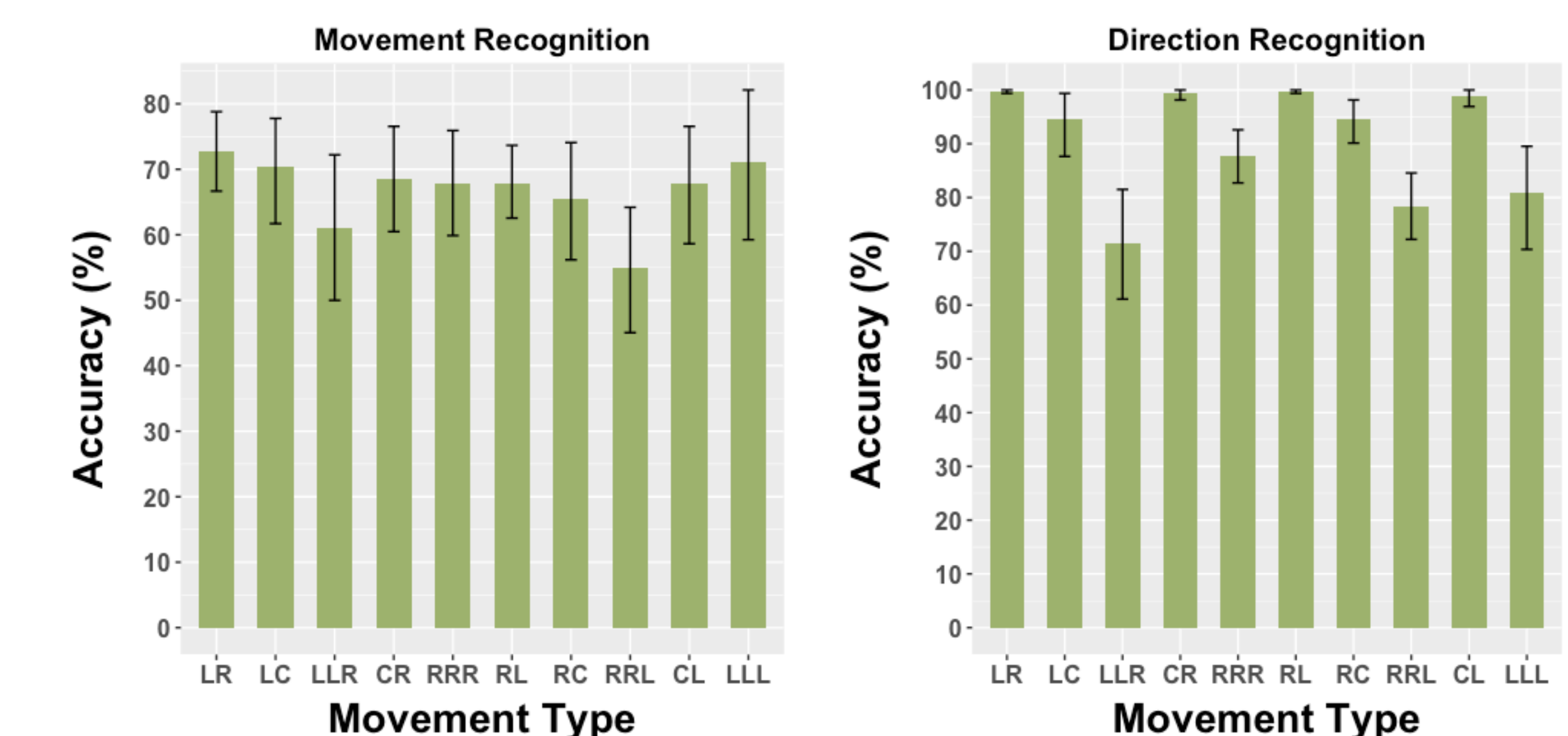
- Three audio zones: 21°, 90°, 160°
- Four audio zones: 21°, 76°, 118°, 160°
- Five audio zones: 7°, 49°, 90°, 132°, 173°

- Recognition errors are influenced by which regions audio cues are in. If an audio cue is in the left region, users tend to think it is further left; if an audio cue is in the right region, users tend to think it is further right.

- Users tend to play audio cues in the central area more times than those in the left or right region.

Moving Audio Cue Movement Recognition

- 7 locations
- 42 starting/ending location combinations
- 10 movement types, e.g. left region to right region (LR), center region to left region (CL), right region to right region moving left (RRL), etc
- Each combination was tested 3 times per subject.



Movement Type Recognition

- Recognition of the starting and ending regions
- Mean is 67.8% (SD=9.22%).
- The recognition rates of audio cues crossing the center are significantly lower than those of other moving types.
- The odds of being recognized correctly increase 1.9 times when the distance is increased by 1 unit.
- Direction has no significant effect.

Direction Recognition

- Recognition of direction only
- Mean is 93.2% (SD=3.38%).
- The recognition rates of audio cues moving within the left or the right region are significantly lower than those of other movement types.
- Movements starting from the center or crossing the center are stronger predictors of correct recognition than other movement types.
- It took significantly less time for users to recognize audio cues starting from the center.
- The odds of being recognized correctly increase 3 times when the distance is increased by 1 unit.
- Direction has no significant effect.