# ACCESSING MULTI-MODAL INFORMATION ON CELL PHONES WHILE SITTING AND DRIVING

Susan R. Fussell[a], Delia Grenville[b], Sara Kiesler[a], Jodi Forlizzi[a], Anna M. Wichansky[b]

[a]Human Computer Interaction Institute
Carnegie Mellon University
Pittsburgh, PA 15213

[b]Advanced User Interfaces
Oracle Corporation
Redwood Shores, CA 94065

Multimodal interfaces have been identified as a possible solution for reducing the visual and motor demands of small devices such as cell phones. In a within-subjects factorial experiment, we explored where audio is useful in a cell phone interface that supports database applications. Participants sat at a desk and drove in a car simulator while choosing a hotel from a descriptive long list. We compared participants' performance with and without the option to listen to the information while it was presented in text. Participants rarely preferred or used the audio option while seated. A substantial number preferred and used the audio option while driving, especially when the hotel choice task was more difficult. Those who chose the audio option looked less at the phone, but increased their task time and did not improve their driving performance. We discuss implications of reading and listening for safety and design.

## INTRODUCTION

Computing devices are becoming more portable. As business software applications are ported to mobile computing platforms, people can use cellular telephones while traveling to perform tasks such as checking email, scheduling, and making reservations on the Web. In this study we aim to better understand how data should be presented in mobile contexts such as vehicle telematic systems and mobile office environments. We address issues of preference, usability, and safety in an experimental exploration of information retrieval in making hotel reservations on a cell phone.

Designing usable and safe interfaces for software running on cell phones has proven challenging. Research shows that the load on visual and motor resources while pressing buttons, scrolling, and reading a small display is comparatively high and potentially dangerous when people have other tasks such as driving (e.g., Alm & Nilsson, 1995; Briem & Hedman, 1995; Brookhuis, DeVries, & DeWaard, 1991; McKnight & McKnight, 1993; Radeborg, Briem, & Hedman, 1999; Redelmeier & Tibshirani, 1997).

Multimodal interfaces are a possible solution for reducing the visual and motor demands of small devices (e.g., Shneiderman, 2000). According to multiple resource theory (e.g.,Wickens, 1991), our visual and auditory modalities process perceptual information somewhat independently. Redundant cues from audio and text (e.g., hearing a caller's name while glancing at his name on the screen) might ease the job of retrieving information on a cell phone (cf. Spence & Driver, 1997).

Our study aimed to increase our understanding of where audio is most useful in a multi-modal mobile telephone user interface that supports database applications. We developed a prototype that had several user interface components commonly found in mobile interfaces, such as short forms, and short and long lists. In this study, we focused on one of these UI components, long lists (9 items or more). Long lists are of interest because database queries often result in a long list of items requiring attention and thought. We studied the effects of adding audio to text versions of long lists.

We presented participants with two usage contexts: sitting at a desk and driving a car. In each context, the participant was given the option to use

audio to listen to the long list while the same information was presented in text. We compared participants' performance with and without the audio option.

We addressed two sets of research questions:
(1) Do people prefer to hear or to view long lists of information? How do task and demographic variables influence these preferences?
(2) How does hearing versus viewing a long list of information affect task performance? We were interested in participants' ability to make a choice from the list and answer a question about it (in both contexts) and in their driving performance (in the driving context).

## METHOD

*Design.* Participants performed eight hotel reservation tasks using a cellular telephone with WAP (Wireless Application Protocol) service. Each of these 8 tasks entailed choosing a hotel and answering a question about its services. The experimental design was a 2x 2 x 2 within-subjects factorial design.

- *Text vs. text plus audio option interface*: Participants performed four tasks using a text-only display and four comparable tasks using the same display with an audio option present.
- *Stationary vs. driving*: Participants performed four tasks while sitting at a desk and four comparable tasks while driving in a car simulator.
- *Low vs. high task difficulty*: Participants performed four easy and four difficult tasks. In the easy task condition, participants had to find a particular hotel in the list such as "the Hyatt." In the difficult task condition, participants had to choose a hotel that satisfied two attributes (e.g., inexpensive motel).

Participants performed in all 8 conditions of the experiment and thus served as their own controls. There was one correct response to each task. All tasks required the same level of navigation through the system, with similar workload, motor, and visual requirements. The order of the stationary and driving conditions was counterbalanced. The order of text-only versus text plus audio option interface was counterbalanced within the driving and stationary conditions. The easy tasks always preceded the difficult hotel choice tasks. To insure

each choice was different, the city, hotel names, and hotel attributes differed for each task. Examples of low and high cognitive complexity tasks are shown in Figure 1.

| Low Difficulty Task |
|---|
| You are stuck at the airport in Los Angeles in May after your flight has been cancelled. The airline has given you a voucher to stay at the Del Rancho motel by the airport. You have some work documents that you need to send, so you want to make sure the motel has a fax machine. Please find the Del Rancho motel in LA and tell me if it has a fax machine. |
| **High Difficulty Task** |
| You're in Houston in May for a conference, but you are unhappy with the low budget place you've been staying in. You decide to switch to a moderately priced motel in the area. You have a lot of suits you need pressed, so you want to make sure the motel has valet service. Please find a moderately priced motel and tell me if it has valet service. |

*Figure 1.* Sample scenarios for high and low difficulty tasks.

*Participants.* Participants consisted of 30 university staff members, graduate students, and undergraduates (20 male, 10 female). Mean age was 27 years old (range 19-49 years). Participants were paid $15 and given a small prize for good driving.

*Equipment.* Participants performed the reservation tasks using a Mitsubishi T250 telephone running digital Internet service. In the driving condition, participants used an I-Sim Alpha Series Driving Simulator. The simulator includes a realistic driver's seat and dashboard, simulated front and side mirrors, and other controls. The accelerator, brake and steering systems operate much like an ordinary car.

*Procedure.* Participants completed a pretest questionnaire. Then they were trained to use the cell phone and application, made a practice reservation with and without the audio option, and had a practice session in the simulator. Participants then performed their first set of four tasks in either the stationary or driving condition; they then switched to the second condition and performed a second set of four tasks. For each trial, the experimenter stated whether audio would be available or not for that task, read the scenario for the task, and then passed the phone to the participant. Participants could request audio at any point while the hotel list screen appeared on the telephone but not during other screens. After each trial, participants completed a

brief post-task questionnaire. After all eight trials, they completed a final questionnaire. Each experimental session lasted about 1.5 hours.

*Measures.* The pre-test questionnaire contained demographic items and questions about participants' driving, cellular telephone, video game, and hotel reservation experience.

Post-task surveys completed after each trial asked participants to rate their attention and their cognitive load using the NASA TLX workload questions (Hart & Staveland, 1988) and to assess the ease and usability of the telephone interface. Participants also rated the percentage of time their attention was focused on the task and the percentage of their visual attention focused on the telephone (on scales of 0, 25, 50, 75, and 100%). On the driving trials, they also rated their driving performance.

The final questionnaire asked participants to rate their preferences for text versus audio interfaces for a variety of information retrieval tasks and to make global assessments of the text interface.

A trained observer recorded total task times and hotel selection times. The observer also noted whether or not the participant chose to exercise the audio option when it was available and any keystroke errors made during the task.

All driving sessions in the simulator were videotaped. Because measurement software was unavailable for the simulator, we used the videotapes to code driving performance. Each 5-second interval of driving was rated on a scale of 1 (in lane) to 5 (crash). Because veering out of lane was comparatively rare, the percentage of time people were in lane was used in the analyses.

## RESULTS

*Do people prefer to hear or to view long lists of information?* Preferences for audio output were highly context-dependent (Figure 2). Overall, just 7% of participants in the easy task condition and 17% in the difficult task condition selected audio. When driving, participants selected audio 40% of the time it was available in the easy task condition and 63% of the time in the difficult task condition. A 2 (stationary/driving) X 2 (task difficulty) repeated measures ANOVA showed a marginally significant effect of mobility condition (F [1, 25] =

3.40, p = .08), indicating that people were more likely to choose audio in the driving condition. Participant attributes such as age, gender, driving experience, and cell phone experience were not correlated with modality choices.
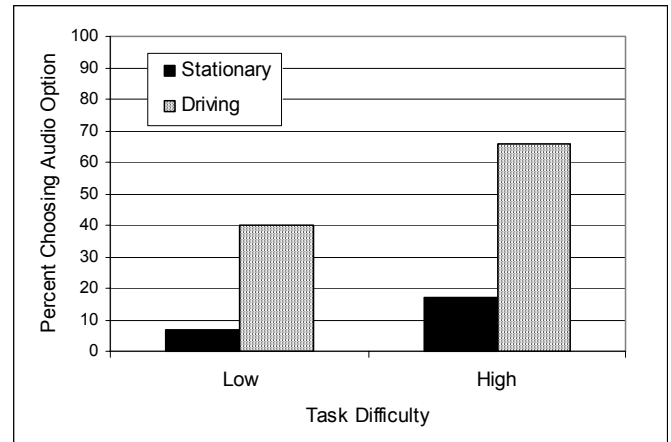


*Figure 2*. Use of the audio option across conditions.

Results from the post-experimental questionnaire showed a similar pattern. Participants said they preferred the text interface for hypothetical information retrieval tasks when stationary but audio when driving (Figure 3). Ratings were analyzed in a 5 (reservation task) by 2 (stationary/mobile) repeated measures ANOVA. Results indicated a highly significant effect of mobility condition (F [1, 26] = 60.68, p < .0001).
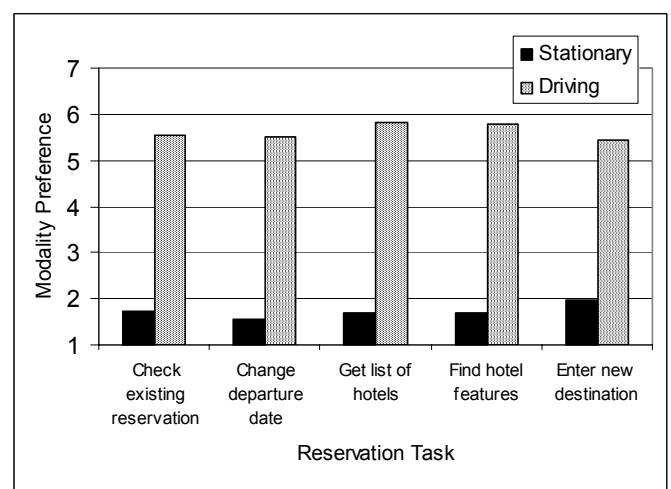


*Figure 3*. Preferences for text versus audio interfaces for five hotel reservation tasks by mobility condition (1 = strongly prefer text; 7 = strongly prefer audio).

*How does hearing vs. viewing a long list of information affect performance?* Hotel reservations took longer when people heard rather than viewed the list of information. Figure 4 shows that participants completed the hotel choice task somewhat faster in the stationary condition (F [1, 168] = 2.73, p = .10), and significantly faster in the easy task condition (F [1, 168] = 8.00, p = .005). Again, response times were significantly longer when participants used the audio option (F [1, 168] = 6.35, p = .01).
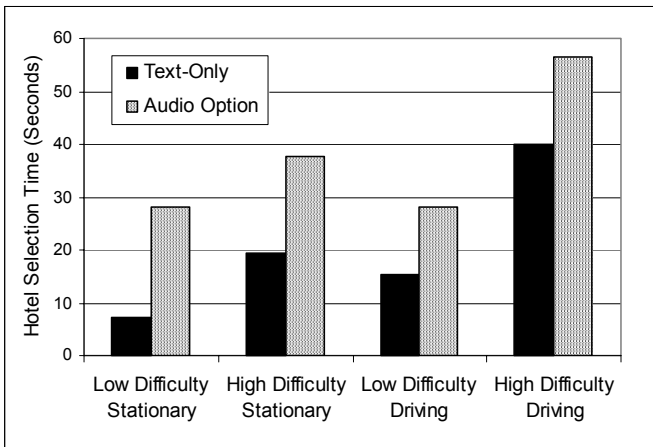


*Figure 4.* Mean hotel selection time as a function of mobility condition, task difficulty, and use of the audio option.

Driving performance was unaffected by whether the information was presented by audio or text (r < .02, ns). Overall, participants drove fairly well as they were performing the reservation tasks; 88% of the time they stayed in lane. However, they veered out of lane as much as 11% of the time when driving and doing the difficult hotel choice task. Participants reported significantly higher workload on the NASA TLX scale in the driving condition (F [1, 190] = 76.99, p < .0001), reflecting the fact that they had to complete two tasks at the same time. However, there were no effects of audio use on these reports.

As one would expect, participants indicated that they focused significantly less visual attention toward the telephone while driving (F [1, 188] = 141.63, p < .0001). There were no significant effects of task complexity or use of the speech option; however there was a borderline significant interaction between mobility condition and use of speech (F [1, 188] = 2.85, p = .09): When participants used the speech interface while driving, they directed less attention toward the telephone (Figure 5).
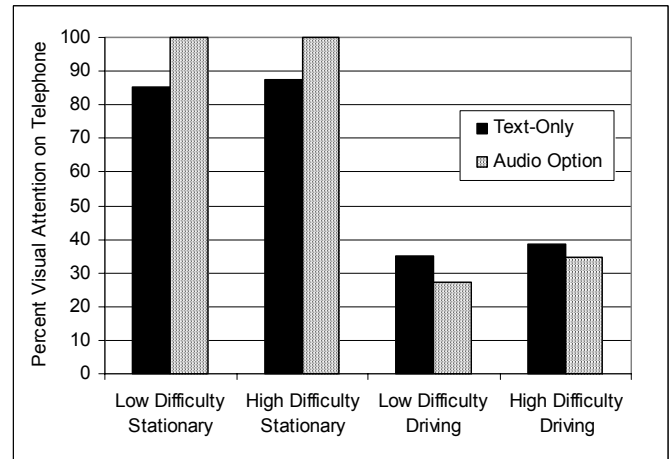


*Figure 5.* Mean attention directed toward the telephone as a function of mobility condition, task difficulty, and use of the audio option.

**DISCUSSION**

Our results suggest that adding audio to the presentation of a long list on a mobile device is not a simple design alternative. Context matters. Participants' stated preferences for cellular telephone interfaces and their actual use of the audio option were strongly dependent on the situation in which they were using the device. They overwhelmingly preferred text-based interfaces when they were stationary. A sizable proportion preferred audio-enhanced interfaces when they were mobile. These findings suggest that audio, even if it is available, will not be widely used in settings in which people can focus full attention on the phone (e.g., while seated at a desk, in a parked car, on a train, or in an airport). We speculate that people will choose to use audio for working with long lists only when they are in a visually-demanding task context (such as driving) and when their use of the list demands sustained visual attention.

Using audio in the presentation of a long list added time to the easy and difficult tasks across the experiment, and did not benefit driving performance. The additional performance time for hotel selection can be explained by the sequential nature of speech. With the text interface, list entries

could be viewed simultaneously on the screen. People who commute or have mobile offices such as salespeople, field service engineers, and warehouse workers do not spend all of their time on the move. Our results suggest that while their vehicles are stationary, most people will read long lists instead of listen to them.

Our study had several limitations that may temper these conclusions. First, our system provided information via audio but did not allow for audio input from the listener. We also did not compare different scroll and listening rates in the audio interface. Users might have a different evaluation of audio when they have the ability to manipulate the list (e.g., to barge in on the list using speech once they have identified their selection) or can set scroll and speech rates to meet their needs. Further research is required to systematically evaluate these dimensions of audio interfaces.

If many people prefer text when in a car, the question arises as to how best to display a long list via text. This question becomes critical if these readers are driving. We observed that participants who read the long list on the cell phone were able to do so while driving by holding the phone against the steering wheel at eye level. We believe a significant number of people will chose to read even long lists on a cell phone while driving if this option is offered to them, a plausible threat to safety (see Just et al., 2001; Redelmeier & Tibshiran, 1997). The design and regulatory alternatives to this possibility should be explored. Design alternatives for multi-modal interfaces include darkening the screen while the car is in motion, providing for sequential audio search and sorting routines, and breaking lists into smaller chunks. Multi-modal vehicle telematic systems offer much potential convenience but require more investigation into designs that are both safe and usable.

## ACKNOWLEDGMENTS

## REFERENCES

Alm, H., & Nilsson, L. (1995). The effects of a mobile telephone task on driver behaviour in a car following situation. *Accident Analysis & Prevention, 27,* 707-715.

Briem, V., & Hedman, L. (1995). Behavioural effects of mobile telephone use during simulated driving. *Ergonomics, 38,* 2536-2562.

Brookhuis, K. A., DeVries, G., De Waard, D. (1991). The effects of mobile telephoning on driving performance. *Accident Analysis and Prevention, 23,* 309-316.

Hart, S. G., & Staveland., L. E. (1988). Development of a multi-dimensional workload rating scale: Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload.* Amsterdam. The Netherlands: Elsevier.

Just, M. A., Carpenter, P. A., Keller, T. A., Emergy, L., Zajac, H., & Thulborn, K. R. (2001). Interdependence of nonoverlapping cortical systems in dual cognitive tasks. *Neuroimage, 14*, 417-426.

McKnight, A., & McKnight, A (1993). The effect of cellular phone use upon driver attention. *Accident Analysis and Prevention, 25,* 259-265.

Radeborg, K., Briem, V., & Hedman, L. R. (1999). The effect of concurrent task difficulty on working memory during simulated driving. *Ergonomics, 42,* 767-777.

Redelmeier, D. A., & Tibshirani, R. J. (1997). Association between cellular-telephone calls and motor vehicle collisions. *New England Journal of Medicine, 336*, 453-458.

Shneiderman, B. (2000) The limits of speech recognition. *Communications of the ACM, 43*, 63-65.

Spence, C. & Driver, J. (1997). Cross-modal links in attention between audition, vision, and touch: Implications for interface design. *International Journal of Cognitive Ergonomics*, 1, 351-373.

Strayer, D. L., & Johnston, W. A. (2001). Driven to distraction: Dual-task studies of simulated driving and conversing on a cellular phone. *Psychological Science, 12*, 462-466.

Wickens, C. D. (1991). Processing resources and attention. In D. Damos (Ed.), *Multiple task performance* (pp. 3-34). Washington D. C.: Taylor and Francis.