Characteristics of Voice-Based Interfaces for In-Vehicle Systems and Their Effects on Driving Performance

Thomas A. Ranney, TRC Inc.
Elizabeth N. Mazzae, NHTSA
G. H. Scott Baldwin, TRC Inc.
Background

Categories of Distraction

- Peripheral distraction
  - Visual interference
  - Manual interference
- Attentional distraction
  - Cognitive interference
Manufacturers use voice interfaces to reduce distraction. Characteristics include:
- Text-to-speech and voice recognition capabilities
- Hierarchical menu structures

Our previous research indicates that voice interfaces:
- Reduce peripheral (visual and manual) distraction
- Have little effect on cognitive (attentional) distraction
Program Objectives

- Develop protocol and metrics for assessing distraction potential of tasks performed using voice interfaces.
- Use protocol to assess selected attributes of secondary tasks and voice interfaces.
<table>
<thead>
<tr>
<th><strong>Traveler Information Systems</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>- Accessible via telephone by dialing ‘511’</td>
</tr>
<tr>
<td>- Implemented in parts of the United States since 2000, 511 service will be:</td>
</tr>
<tr>
<td>- Available in at least 25 states by 2005</td>
</tr>
<tr>
<td>- Operational throughout the United States by 2010</td>
</tr>
</tbody>
</table>
Traveler Information Systems

- Voice-activated navigation of hierarchical menu structures
- May be used by commuters and unfamiliar travelers
- Provide current information about:
  - Traffic conditions - accident and construction delays
  - Road conditions – slippery, snow-covered etc.
  - Public transit
Research Objectives

- To determine whether selected secondary tasks degrade driving performance
  - Simulated phone conversation
  - 511 information acquisition tasks
- To evaluate the effects of specific task interface attributes:
  - Required map use
  - Voice interface reliability
Research Approach

- We developed:
  - Simulated 511 system using a “Wizard of Oz” approach
    - Human replaces voice recognition component
  - Navigation questions related to a hypothetical system of interstate roads
511 Task: Hypothetical Roadway System
Example 511 Tasks

- Which route has better road conditions, I-34 or I-51? (No Map)
- Which segment has more incident-related time delay, segment 3 or segment 7 (Map required)
511 Task: Hypothetical Roadway System
511 Task: Hierarchical Menu Structure
Research Approach cont.

- **Primary (driving) task**
  - Car following
  - Peripheral detection task (PDT)

- **Secondary tasks (hands-free phone and voice interface)**
  - Simulated phone conversation task
  - Simulated navigation task (4 conditions)
    - Information acquisition mode (auditory vs. auditory + visual map)
    - System reliability (no voice recognition errors vs. 20% errors)
Car-Following Task
Peripheral Detection Task
Experimental Design

- **Independent variables:**
  - Driver age group
  - Secondary task
  - Lead vehicle speed signal

- **Dependent measures:**
  - Car following (coherence and delay)
  - Vehicle control (steering entropy)
  - Peripheral target detection (% correct, RT)
  - Subjective workload rating
Method

- Thirty-six drivers (18-25, 30-45, 50-60)
- Drivers drove a 1996 Honda Accord and followed a lead vehicle on TRC’s 7.5-mile test track
  - Data collected on both 2 mile straight segments
  - Speeds varied between 45 and 60 mph
## Categories of Driving Performance Measures

<table>
<thead>
<tr>
<th>Category</th>
<th>Interference</th>
<th>Primary Measures</th>
<th>Color Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vehicle Control</td>
<td>Peripheral</td>
<td>Steering entropy</td>
<td></td>
</tr>
<tr>
<td>Car Following</td>
<td>Attentional</td>
<td>Coherence, Delay</td>
<td></td>
</tr>
<tr>
<td>Target Detection</td>
<td>Peripheral</td>
<td>% Detected, Response Time</td>
<td></td>
</tr>
<tr>
<td>Subjective Rating</td>
<td>Overall</td>
<td>RSME Workload</td>
<td></td>
</tr>
</tbody>
</table>
Overview of Results

- Baseline comparisons
- Interface design hypotheses
  - Map vs. No Map
  - Error vs. No Error
- Differences between Phone and 511 tasks
## Baseline Comparisons

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Vehicle Control</th>
<th>Car Following</th>
<th>Target Detection</th>
<th>Subjective Rating</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Steering Entropy</td>
<td>CF Coherence</td>
<td>CF Delay</td>
<td>PDT Detect</td>
<td>PDT RT</td>
</tr>
<tr>
<td>No Error</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Error</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Map</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>No Map</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Phone</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
Summary of Results: Baseline Comparisons

- All Secondary tasks degraded driving performance.
  - 511 tasks degraded all categories of driving performance.
- Phone conversation task was least disruptive
  - Phone conversation did not affect car following performance.
## Interface Design Comparisons

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Vehicle Control</th>
<th>Car Following</th>
<th>Target Detection</th>
<th>Subjective Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Map vs. No Map</td>
<td>Steering Entropy</td>
<td>CF Coherence</td>
<td>CF Delay</td>
<td>PDT Detect</td>
</tr>
<tr>
<td></td>
<td>✅</td>
<td></td>
<td>✅</td>
<td></td>
</tr>
<tr>
<td>Error vs. No Error</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Summary of Results: Interface Design Issues

- Tasks requiring map were more disruptive than similar tasks not requiring map.
  - Impairment observed on 5 of 6 key performance measures
- Increased voice recognition errors did not affect driving performance.
  - No differences observed
# Phone vs. 511 Task

<table>
<thead>
<tr>
<th>Phone vs.</th>
<th>Vehicle Control</th>
<th>Car Following</th>
<th>Target Detection</th>
<th>Subjective Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phone vs.</td>
<td>Steering Entropy</td>
<td>CF Coherence</td>
<td>CF Delay</td>
<td>PDT Detect</td>
</tr>
<tr>
<td>Map</td>
<td></td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Error</td>
<td></td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>No Map</td>
<td></td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>No Error</td>
<td></td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
</tbody>
</table>
Summary of Results: Phone vs. 511 Tasks

- All categories of 511 tasks revealed greater degradation than phone conversation task on at least one measure
  - Map Tasks were most different from phone task
- Differences most apparent for car-following measures
  - Suggests 511 tasks were more cognitively demanding
- No differences in workload ratings
Other Results

- Effects of secondary tasks were consistent across all three age groups.
Conclusions

- Test track protocol is sensitive to distraction effects of secondary tasks performed with voice interface and hands-free phone.
- 511 tasks interfered with all aspects of driving performance.
- Phone conversation task impaired target detection and vehicle control but not car-following performance.
Safety Implications

- Complex secondary tasks, including those requiring active navigation of hierarchical menu structures, are likely to disrupt driving more than simple phone conversations.
Design Implications

- Avoid unnecessary or redundant visual displays
- Drivers can tolerate systems with some voice recognition errors
- Minimize complexity of hierarchical menu structures
Thanks for your attention!