Trends Toward Natural User Interfaces in the Car

Thomas Schalk
Agenda

- Multimodality in the car
- User satisfaction data for speech UI in the car
- Driver distraction study on speech and visual UI
- Why is the speech button so difficult to the user?
- Guidelines for multimodal UI design
Human Modalities for User Interfaces

**ENTRY**

- Gesture
- Speech
- Touch

**Communication to Device**

**PRESENTATION**

- Vision
- Sound
- Haptic

**Communication to User**
Speech in the Car: Are Users Satisfied?

2012 U.S. Navigation Usage And Satisfaction Study
-- J.D. Power and Associates (Press Release 1/11/13)

➤ Satisfaction study clearly shows that drivers want speech
  - 67% of the users without speech – want it
  - 80% of the users with speech – want it again

➤ Speech scored the lowest satisfaction score in the study
  - Significantly worse than second lowest score

➤ Speech was reported to be difficult to use
  - Third most frequently reported problem in the study
  - Recognition accuracy has been an issue in the car for years
Icons: Discovery and Task Selection

- Drivers are accustomed to touchscreen input
  - Mobile device usage continues to grow
  - Consumers expect touchscreen in luxury vehicles
  - Mass market vehicles are adopting touchscreen

- It can be safe to manage results with a visual / manual HMI
  - Glancing is quick and is interruptible
  - Selecting with touch is practical and quick (limited choices)
  - Information is best perceived through vision, not speech

- Supported by driver distraction data
  - Across all ages, touch to select from a short list was shown to be safe
  - Audio-only interfaces exhibit high task completion times
  - It takes time to play the items to choose from
Driver Distraction Variables of Interest

Task Duration
- Critical metric while driving

Task Accuracy
- Success rate at completing tasks

Speed Measures
- Mean Speed (mph)
- Speed Variance (mph)

Steering Measures
- Steering Variance (degrees)
- # of Lane Deviations (count)
- Time out of lane (seconds)

Eye Glances
- Glance Durations (seconds)
- Eyes Off Road Time (percentage)

Peripheral Event Detection
- %Detected/%Missed
- Latency to Detect

Workload Ratings
- Mental Demand
- Frustration Level
- Situational Awareness

Post Drive Questionnaire
- Overall Acceptability
Overview of VTTI Visual Speech Study

24 Participants
- Ages 18-30 (n=12)  Ages 65-75 (n=12)

3 User Interfaces

Destination Entry Tasks
- Address, POI, Category
Test Vehicle Center Stack Layout
TASK DURATIONS
Task Duration Comparison

Mean Observed Task Durations by Task Type

- Baseline: 89.11 seconds
- Aftermarket Nav: 37.56 seconds
- Speech-Only: 55.35 seconds
- Speech & Visual: 24.67 seconds
SPEED VARIANCE
Speed Variance

Mean Observed Vehicle Speed Variance by Task Type

- **Baseline**: 2.12
- **Aftermarket Nav**: 3.24
- **Speech-Only**: 2.36
- **Speech & Visual**: 1.43
STEERING VARIANCE
Steering Variance

Mean Observed Steering Angle Variance by Task Type

Steering Variance (degrees)

<table>
<thead>
<tr>
<th>Task Type</th>
<th>Mean Observed Steering Angle Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>5.17</td>
</tr>
<tr>
<td>Aftermarket Nav</td>
<td>6.88</td>
</tr>
<tr>
<td>Speech-Only</td>
<td>5.41</td>
</tr>
<tr>
<td>Speech &amp; Visual</td>
<td>4.35</td>
</tr>
</tbody>
</table>
LANE DEVIATION ANALYSIS
Lane Deviations by Age

Mean Observed Frequency of Lane Deviations by Task Type and Age Group

- Younger (18-30)
- Older (65-75)

<table>
<thead>
<tr>
<th>Task Type</th>
<th>Younger (18-30)</th>
<th>Older (65-75)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>0.00</td>
<td>0.29</td>
</tr>
<tr>
<td>Aftermarket Nav</td>
<td>0.17</td>
<td>1.10</td>
</tr>
<tr>
<td>Speech-Only</td>
<td>0.03</td>
<td>0.14</td>
</tr>
<tr>
<td>Speech &amp; Visual</td>
<td>0.04</td>
<td>0.11</td>
</tr>
</tbody>
</table>
EYE GLANCE ANALYSIS
Eye Glance Data

Mean Observed Percentage of Eyes Off Road Time by Task Type

- Baseline: 30.7%
- Aftermarket Nav: 64.5%
- Speech-Only: 28.4%
- Speech & Visual: 40.2%
WORKLOAD RATINGS
Workload Measurement: Mental Demand

Mean Ratings of Mental Demand by Task Type

- Baseline: 12.2
- Aftermarket Nav: 54.0
- Speech-Only: 26.7
- Speech & Visual: 22.4

Mental Demand 1 (Easy) - 100 (Hard)
Guidelines for the Best Automotive UI

- Maximize simplicity
- Minimize # of task steps
- Minimize number of menu layers
- Avoid voice menus
- Disallow typing
- Minimize incoming messages
- Maximize interruptibility
- Minimize verbosity
- Remove need for learning mode
- Minimize speech input
- Minimize glance duration
- Minimize glance frequency
- Minimize task completion time
- Maximize driving performance
Input Modalities
- Speech
- Touch
- Gesture

Output Modalities
- Sound
- Vision
- Haptic

Generic Subtask
- Select Task
- Present Results or Menu Choices
- Select Item From List of Results
- Prompt User
- Enter Text String
- Warn User of Danger
- Interrupt and Pause Task
- Resume Task
- Task Completed

〇 - Preferred
△ - Alternate
“Tap or say your selection”

**Scenario 1:**
User taps weather icon

**Scenario 2:**
User says “weather”

**Scenario 3:**
User says “What’s the forecast for tomorrow?”
Conclusion

- Multimodality in the car is key to usability and safety
- Speech UI in the car needs to be improved
- Driver distraction data supports mixing speech & vision
- The speech button needs to integrate with the touch screen
- More research is needed to validate proper UI design
Thank You

tschalk@agero.com