Automotive Speech Applications in the Post-Siri Era

Benjamin Ao
Chief Engineer
Alpine Electronics Research of America
Disclaimers

• Siri is not the only star – just the first in line.

• Siri is well and alive and will live long.

• Still, we see headlines such as
  Intel’s voice recognition will blow Siri out of the water—because it
  doesn’t use the cloud

• So, we dunno 😞
# Siri as Personal Assistant

<table>
<thead>
<tr>
<th>Voice Control</th>
<th>Personal Assistant</th>
</tr>
</thead>
<tbody>
<tr>
<td>User must memorize commands</td>
<td>User can use natural language expressions</td>
</tr>
<tr>
<td>User must follow command structure</td>
<td>User can say anything in any context</td>
</tr>
<tr>
<td>USA nationwide one-shot VDE difficult</td>
<td>USA nationwide one-shot VDE easy</td>
</tr>
<tr>
<td>SMS/email dictation is not possible</td>
<td>SMS/email dictation is possible</td>
</tr>
<tr>
<td>Generally poor ASR performance</td>
<td>Generally good ASR performance</td>
</tr>
<tr>
<td>Slow and expensive integration process</td>
<td>Fast and inexpensive integration process</td>
</tr>
<tr>
<td>Each implementation is new and needs validation</td>
<td>Server implementation is more mature and reliable</td>
</tr>
<tr>
<td>Can be 100% embedded without Internet access</td>
<td>Cannot be 100% embedded without Internet access</td>
</tr>
</tbody>
</table>
Automakers Facing Challenge
Options for Automakers

**Option 1: Cuss and ignore**
- **Pros:** No worries, no pains.
- **Cons:** Risk of fading into oblivion.

**Option 2: If ya can't beat 'em, join 'em**
- **Pros:** Fast and cheap ride on the bandwagon.
- **Cons:** Loss of control over own destiny.

**Option 3: Do it yourself, stupid!**
- **Pros:** New opportunities, bigger market share, more $$$$.
- **Cons:** Limitations in hardware and infrastructure.
Options for Do-it-Yourselfers

Option 3.1: Off-Board System
• Pros: Superior server-based ASR/NLU performance.
• Cons: Spotty Internet connectivity + network traffic jam.

Option 3.2: On-Board System
• Pros: Independence from Internet connectivity.
• Cons: Inferior ASR/NLU performance due to HW limitation.

Option 3.3: Hybrid System
• Pros: Best of both worlds.
• Cons: Difference between on-b. vs off-b. user-experience.
What Needs To Be Done

TBD 1: New Approach to HMI Design
• Do away with rigid hierarchical command structure and adopt flexible flat multimodal interface structure.

TBD 2: More Powerful Hardware
• Enable Siri-like performance in the embedded space requires hardware as powerful as today’s speech servers.

TBD 3: Embedded Dictation and NLU
• Enable Siri-like performance in the embedded space also requires porting dictation and NLU into embedded space.
Players to Watch For

- Intel
- Nuance
- Apple
- Google
- Microsoft
- Hyundai
- BMW
- Audi
- Mercedes-Benz
- Toyota
- Cadillac
- Ford
Thank You!