

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 **EXCLUSIVE Intel's voice recognition will blow Siri out of the water—because it doesn't use the cloud**

- So, we dunno 😞

Intel's offline voice recognition tech could be a Siri-killer



Siri as Personal Assistant



Voice Control



Oct 4, 2011

Personal Assistant



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

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



Thank You!