Is NLU the future of speech interaction in the vehicle?

Yvonne Gloria
Voice Recognition Engineer
Ford Motor Company
BUT I SAID, “BUY BIRTHDAY GIFT”. NOT “DRIVE OVER CLIFF”!

Driverless Car Mishap #13
Speech Recognition and Driver Distraction

* What is driver distraction?
  * 3 main types: visual, manual, cognitive
  * Manual-visual tasks vs. speech tasks
    * Manual tasks easily measured by glances away from the road or hand off the wheel
    * Speech tasks are harder to measure
Speech in the Vehicle

Strategy Analytics, 2014
Cognitive Impact of Speech Tasks

* Secondary tasks can increase cognitive load, which may lead to an increase in cognitive distraction
  * Speaking to the system
  * Listening to the system feedback
  * Glancing at the display for confirmation/further feedback
* An Increase in cognitive distraction can lead to
  * Increased braking
  * Increased lane departure
  * Following too closely
  * Decreased attention to what’s going on outside the vehicle
Natural Language Understanding (NLU) has many definitions and interpretations depending on where it is being used.

* In mobile devices, Apple Siri and Google Now are considered, by most, to be an example of an NLU-enabled device.
  * What does this mean to the consumer?
* Would this functionality serve the consumer as well in the vehicle?
NLU in the Vehicle

* Structured interactions (usually menu-based)
  * Great for novice users; frustrating for power users
* Less structured interactions (may be menu-based with more globally available commands)
  * There is less guidance for the novice user
  * A power user might experience more confirmation steps after giving a command due to the increase in the vocabulary the system is listening to
* Open-ended interaction
  * Little to no guidance on what can be said
  * The same request may have different results depending on confidence scores
  * Possible increase in processing delay
  * Possible increase in misrecognition
Sync 3 uses a combination of structured menus and a flattening of some grammar commands to cater to both novice and experienced users.

- Media “play” commands are accessible from the main menu to provide instant access to music on a USB device
- Audio source switching takes only one command and FM/AM/Sirius stations can be accessed by one command.
- More complex tasks, like POI/Address search, are in a more structured menu
  - SYNC 3 offers the ability to say partial POI name, which improves the accuracy of the search
  - The system also allows the user to say the address all at one time, without entering the pieces separately
A user won’t know what functions are available through the voice system unless the system tells them.

How do we provide that information in a way that will be helpful but not increase driver distraction or annoy the user?

Does the accuracy of the speech recognizer add to cognitive distraction?

Does the way the Text-to-speech in the vehicle sound lead to increased distraction?

The answer to this question will help us improve our prompt feedback.
Yvonne Gloria
ygloria1@ford.com