Evaluation Procedures for Voice Interfaces in Vehicles

Peter Burns, Joanne Harbluk, Julia Mitroi

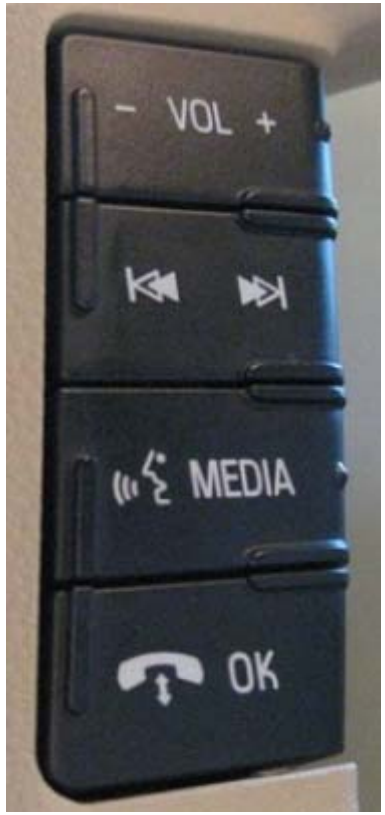
Ergonomics and Crash Avoidance

Road Safety and Motor Vehicle Regulation Directorate

Canada

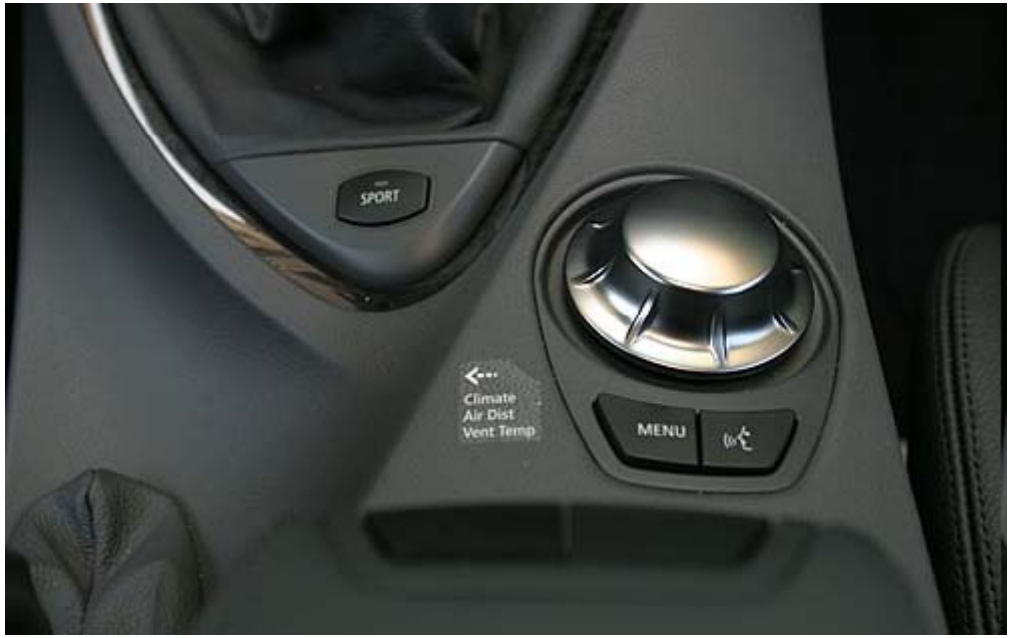
Outline

- Definition and characteristics of voice interfaces
- Pros and Cons
- Developing safer voice interfaces
- Tools for evaluating voice interfaces
- Research
- Summary and conclusions



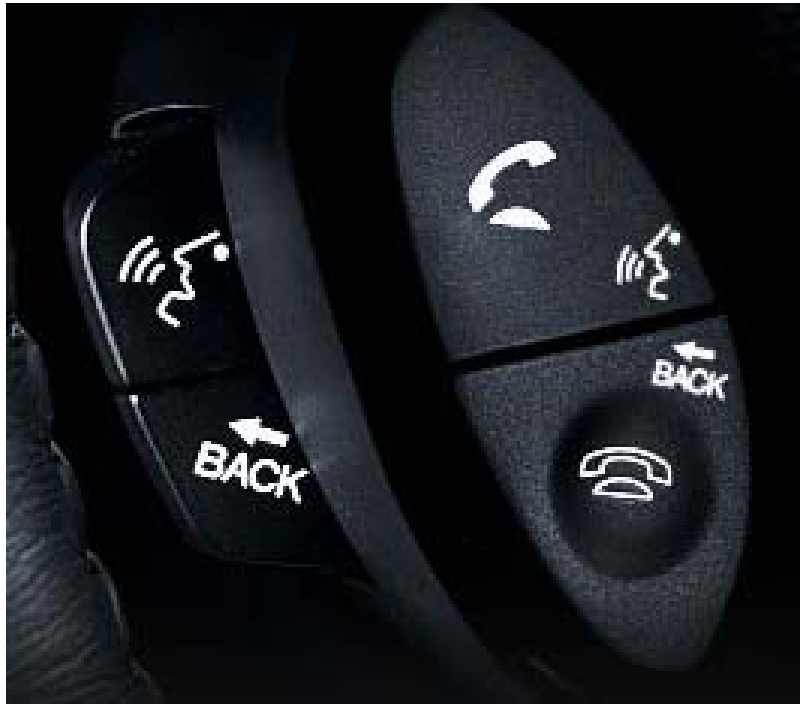
Definition of Voice Interfaces

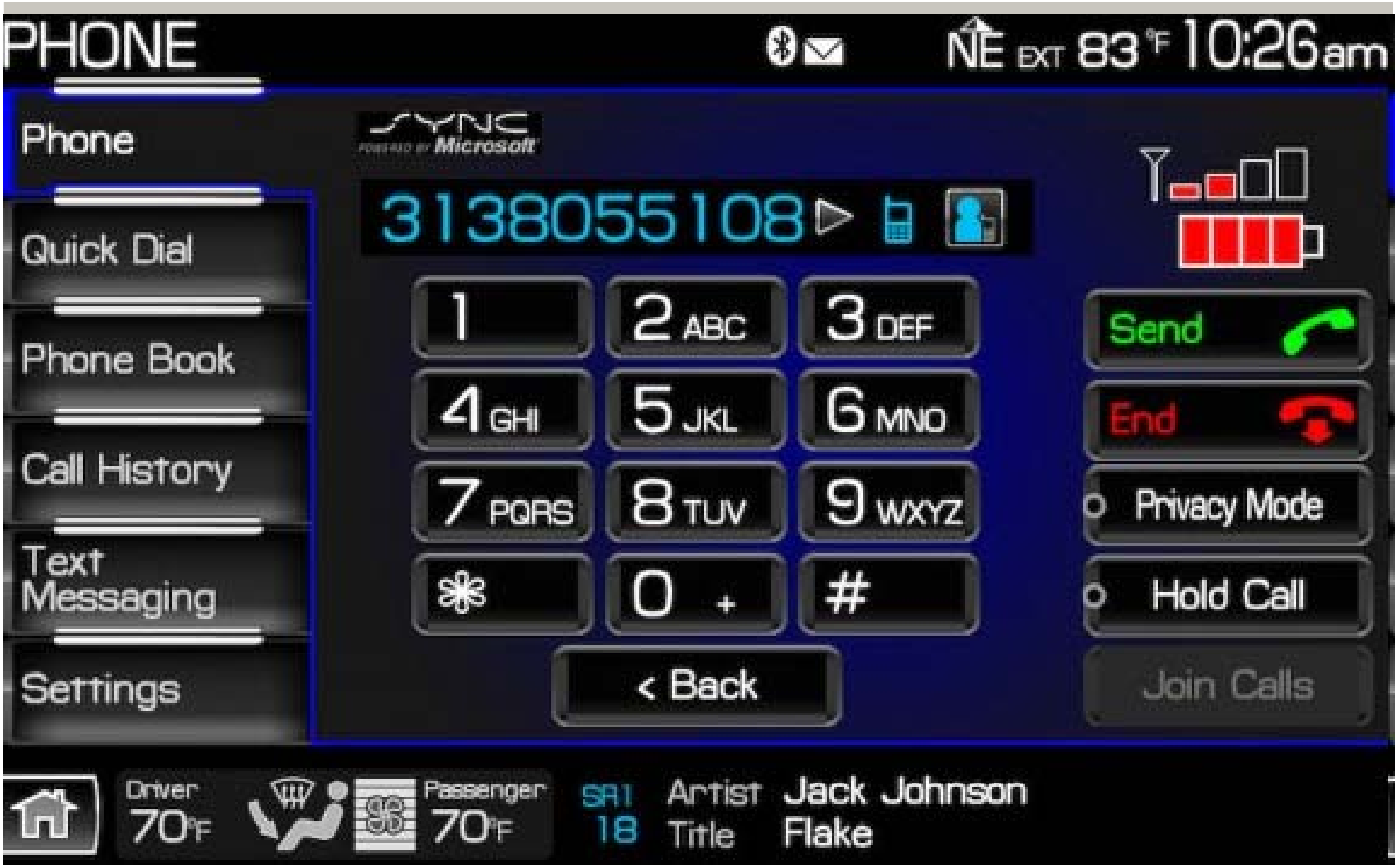
Voice interfaces are interactive media that allow speech input and provide speech feedback (e.g., Harris, 2005).

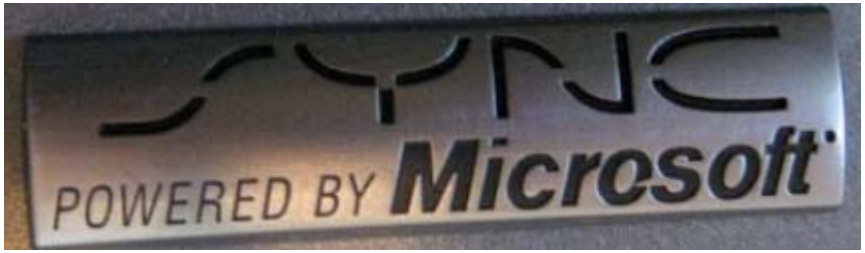


Characteristics of Voice Interfaces

- Speech output type and clarity
- Press-to-talk input
- Recognition vocabulary
- Recognition accuracy
- Text-to-speech
- Speech-to-text
- Control logic and menu structure




PHONE
NE EXT 83°F 10:26am
Phone
Quick Dial
Phone Book
Call History
Text Messaging
Settings
SYNC
POWERED BY Microsoft
3138055108
1
2 ABC
3 DEF
4 GHI
5 JKL
6 MNO
7 PQRS
8 TUV
9 WXYZ
0 +
#
< Back
Send
End
Privacy Mode
Hold Call
Join Calls
Driver
70°F
Passenger
70°F
SR1
18
Artist Jack Johnson
Title Flake


SYNC
POWERED BY Microsoft

Advantages of Voice

- Hands-free control of certain in-vehicle functions
- Facilitates multi-tasking
- Require less space than manual controls

Experiments comparing driving with a voice and visual-manual interfaces found that voice had:

- better driving performance (fewer lane departures, steadier speed),
- less subjective workload, and
- less time spent looking away from the road
- using a voice interface was worse than just driving in most cases.

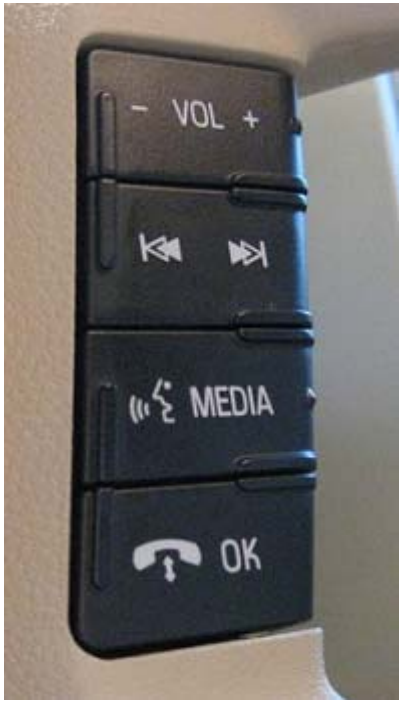
Baron and Green (2006)

Disadvantages of Voice

- Voice only interaction is not truly hands and eyes free
- Requires driver attention
- Facilitates multi-tasking
- Can take more time to perform tasks than conventional modes of interaction
- Usability problems (e.g., recognition accuracy, non-intuitive, cumbersome…)
  - e.g.,interfaces that "… respond to us with repetitions of their own, parroting our words back to us incessantly for confirmation, like a four-year-old kid trying to drive her brother crazy" (Harris, 2005).
  - Nielsen (2003) argues that visual interfaces are inherently superior to auditory interfaces for many tasks. "The Star Trek fantasy of speaking to your computer is not the most fruitful path to usable systems. Voice is a one-dimensional medium with zero persistence" .

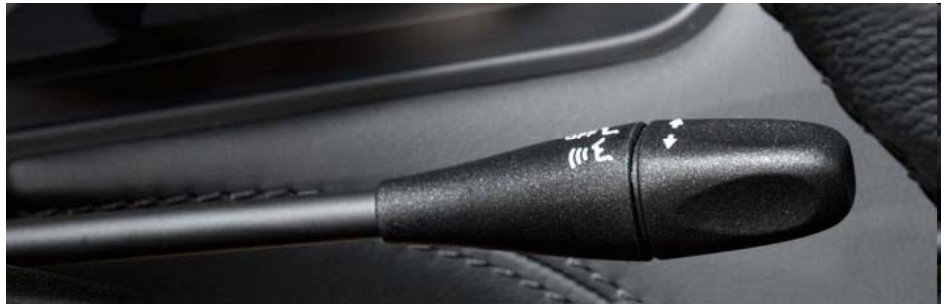
Destination input by voice

- Enter city destination - “Niagara Falls”
- Enter point of interest - Hospital



Concerns

- Having a voice interfaces is no guarantee that a task is safe to do while driving
- Some voice interfaces are better than others
- Safety should not be compromised for entertainment and convenience.
- Need tools to support the development of safer interfaces
- Need methods identify unsafe interfaces and tasks



Designing Safer Voice Interfaces: Guidelines

- Do not use a voice interface without proper reason.
- Simple one-word commands such as "next," "back," or "select," are more efficient and satisfactory than conversational-style dialogues.
- Allow for easy error recovery
- Focus is on reducing cognitive load;
- Provide user-centered feedback in response to the speech input
- good recognition accuracy, which can affect user acceptance and system performance
- speech recognition flexibility (grammar, and pronunciation)
- Voice interfaces should be evaluated to determine their usability and safety. Do not assume that voice makes it safe

(e.g., Schmidt-Nielsen et al., 2008)

Designing Safer Voice Interfaces: Evaluation Methods

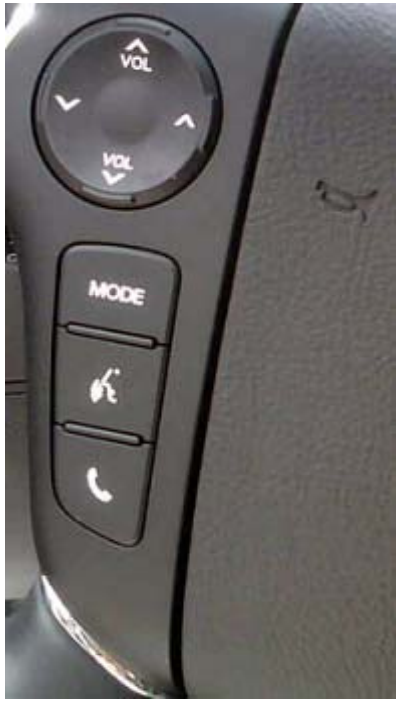
- Standard methodology that specifies the test equipment, procedures, participants, and data treatment.

- Metrics should also be:
  - Sensitive
  - Objective
  - Reliable/ repeatable
  - Valid
  - Practical (cost, time, ease-of-use)
  - Diagnostic

- Furthermore, the metrics should have established analysis procedures, criteria and decision rules.

Tools for evaluating voice interfaces

- Object and event detection (PDT)
- Visual behaviour (eyes off road time)
- Vehicle control (lane keeping)
- Secondary task performance (task time, errors)
- Lane change task (LCT)
- Sternberg memory task
- Subjective workload



Transport Canada Studies

1. Visual-manual versus voice tasks on Auto-PC – test track
2. Hands-free phone tasks – on road study
3. Text-to-speech email readers – simulator study
4. Speech guidance directions - simulator study (HASTE)
5. Easy vs hard destination input by voice – Lane Change Task (LCT)
6. Easy vs hard destination input by voice – simulator study

Ratings of Methods

Measure	Look/ Touch	Talk/Listen
Object and event detection	✓	✓
Visual behaviour	✓	✓
Vehicle control		
Secondary task performance	✓	✓
Lane change task		
Subjective ratings	✓	✓

The Crash Avoidance Metrics Partnership (CAMP)

- CAMP investigated driving performance and driver workload.
- Aimed to develop a toolkit of performance metrics and test procedures to assess how using an in-vehicle system effects driving performance.
- For auditory-vocal and mixed-mode tasks, CAMP recommended a toolkit that includes:

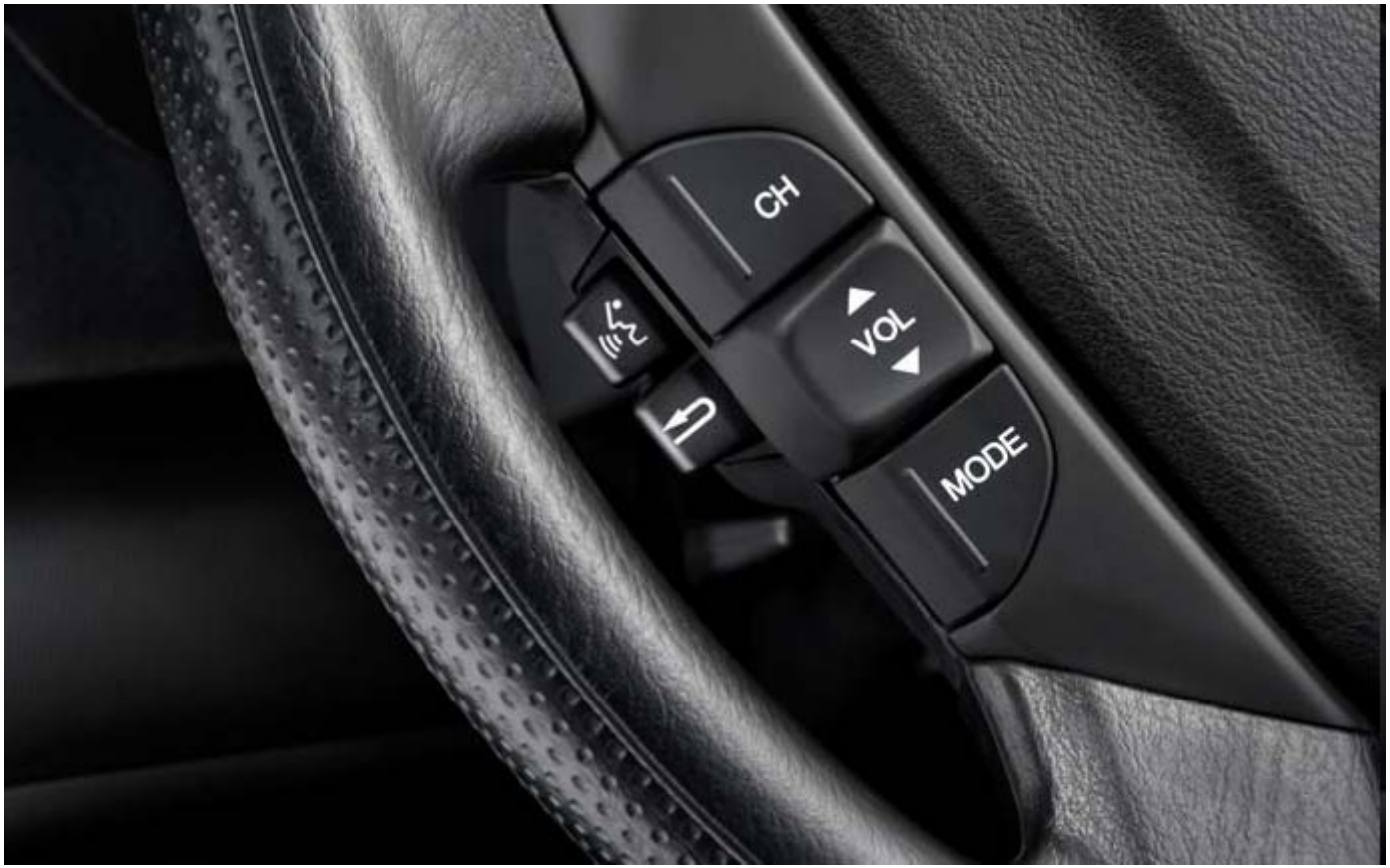
1. total activity time,
2. PDT and
3. Sternberg memory task.

(Angell et al., 2006)

Summary and Conclusions

- voice interface ≠ safe
- Having better performance than the visual-manual equivalent does not mean a task is safe
- A metric that is effective at evaluating both 1) voice and 2) visual-manual interfaces is needed if they are being compared
- Voice interfaces need to be designed in accordance with HF guidelines and their performance and usability need to be tested
- Effective metrics exist (PDT, dynamic task time) but more work is needed
- Criteria for safe and usable interfaces are needed

Thank you



Transport Canada Studies

1. Harbluk, J.L., Burns, P.C., Trbovich, P.L., McCurdie, T., Bleichman, D., Brown, M, Gamble, M. & Herdman, C. (2006). Using the Perceptual Detection Task to assess IVIS safety: Results from the HASTE project. Presented at Vision In Vehicles 11, Dublin, 29 July 2006.

2. Harbluk, J.L., Ranney, T.A., Noy, Y.I., Trbovich, P., & Eizenman, M. (2003). Using speech-based and manual interfaces while driving: The impact on visual behavior and driving performance. Presentation at VIV 10, Granada Spain, September 2003.

3. Harbluk, J.L., Noy, Y.I., Trbovich, P.L. & Eizenman, M. (2007). An on-road assessment of cognitive distraction: Impacts on drivers' visual behavior and braking performance. Accident Analysis & Prevention, 39, 372-379. Text-to-speech

4. Harbluk, J.L., Burns, P.C., Lochner, M., & Trbovich, P.L. (2007). Using the lane-change test (LCT) to assess distraction: Tests of visual-manual and speech-based operation of navigation system interfaces. 4th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design. Stevenson, Washington: July 2007.

5.Harbluk, J.L. & Lalande, S. (2005) Performing Email Tasks While Driving: The Impact of Speech-Based Tasks on Visual Detection. Driving Assessment 2005. Rockport, Maine: June 2005.