METHODS FOR SPEECH AND AUDIO EVALUATION IN VEHICLES

Jan Reimes – Rapporteur Q4/12

Workshop on Voice Recognition Implementation

11 July 2023

Introduction

Introduction

- Your new rapporteur for Q4/12 (since January 2023)

Jan Reimes

Introduction

- Your new rapporteur for Q4/12 (since January 2023)
- Took over position from Hans Gierlich in Q4/12 (and Q6/12)

Jan Reimes

Introduction

- Your new rapporteur for Q4/12 (since January 2023)
- Took over position from Hans Gierlich in Q4/12 (and Q6/12)

- Focus on standardized acoustic testing and quality prediction models
- Drafted technical documents/specifications/reports in relevant committees:
  3GPP SA4, ITU-T SG12 (WP1), ETSI STQ, …

Jan Reimes

Introduction

- Your new rapporteur for Q4/12 (since January 2023)
- Took over position from Hans Gierlich in Q4/12 (and Q6/12)

Jan Reimes

- Focus on standardized acoustic testing and quality prediction models
- Drafted technical documents/specifications/reports in relevant committees:
  3GPP SA4, ITU-T SG12 (WP1), ETSI STQ, …

Objective:

- Q4/12 scope & core specifications
- New work on automatic speech recognition & overlap with H.VM-VMIA

Introduction to Q4

"Objective methods for speech and audio evaluation in vehicles"

Introduction to Q4

"Objective methods for speech and audio evaluation in vehicles"

- Test methodologies and performance requirements of vehicle-related speech communication terminals, in particular:
  - Hands-free telephony → P.1100/1110/1110 (NB/WB/SWB codecs)
  - Emergency call (eCall) systems → P.1140
  - In-car communication → P.1150

Introduction to Q4

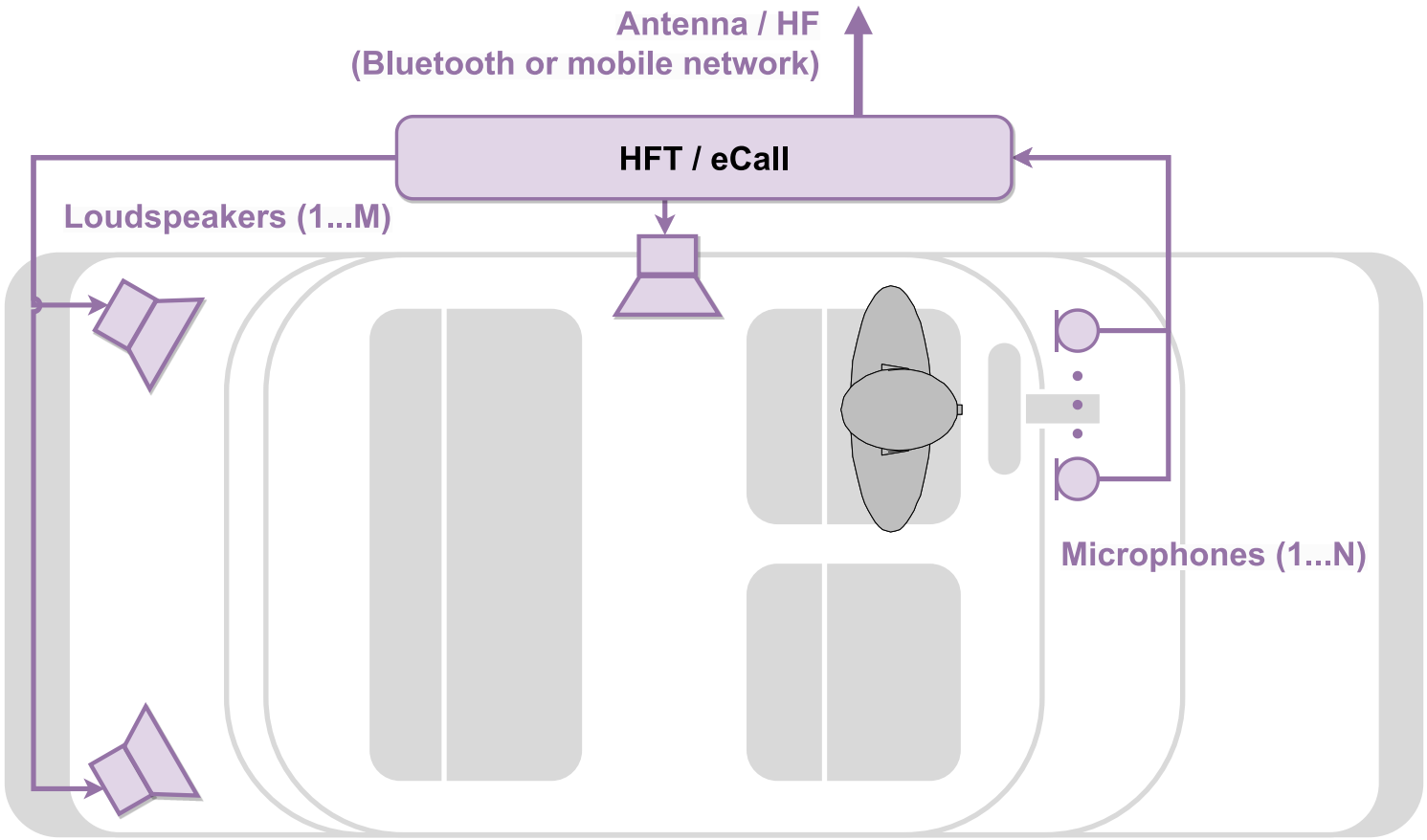
"Objective methods for speech and audio evaluation in vehicles"

- Test methodologies and performance requirements of vehicle-related speech communication terminals, in particular:
  - Hands-free telephony → P.1100/1110/1110 (NB/WB/SWB codecs)
  - Emergency call (eCall) systems → P.1140
  - In-car communication → P.1150

- Assumption for all Recommendations:
  - No requirements on components/implementations
  - No access to internals is required to perform testing ("black box")
  - Only use standard electrical and acoustical interfaces

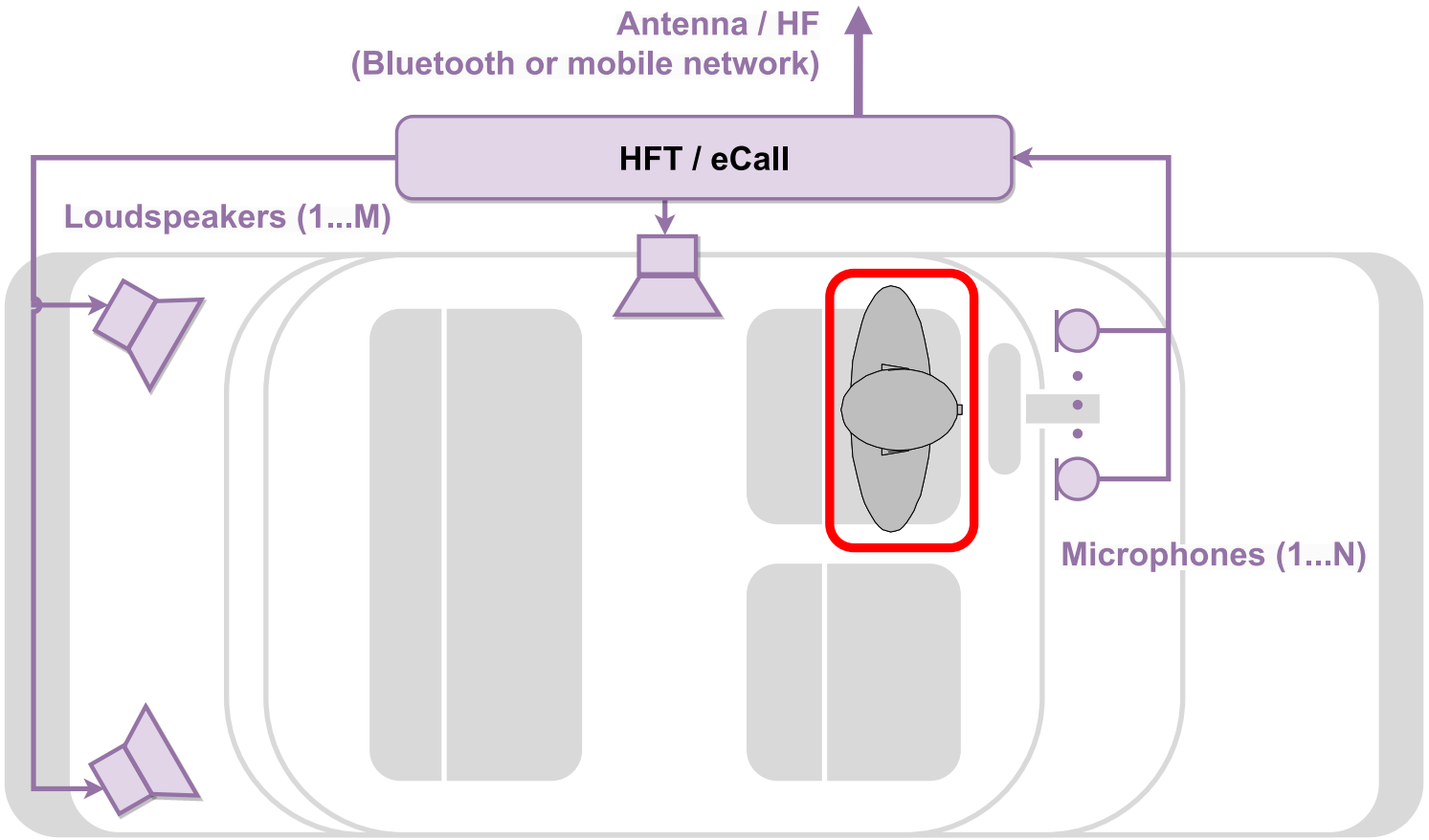
Testing principles – Example setup

Testing principles – Example setup



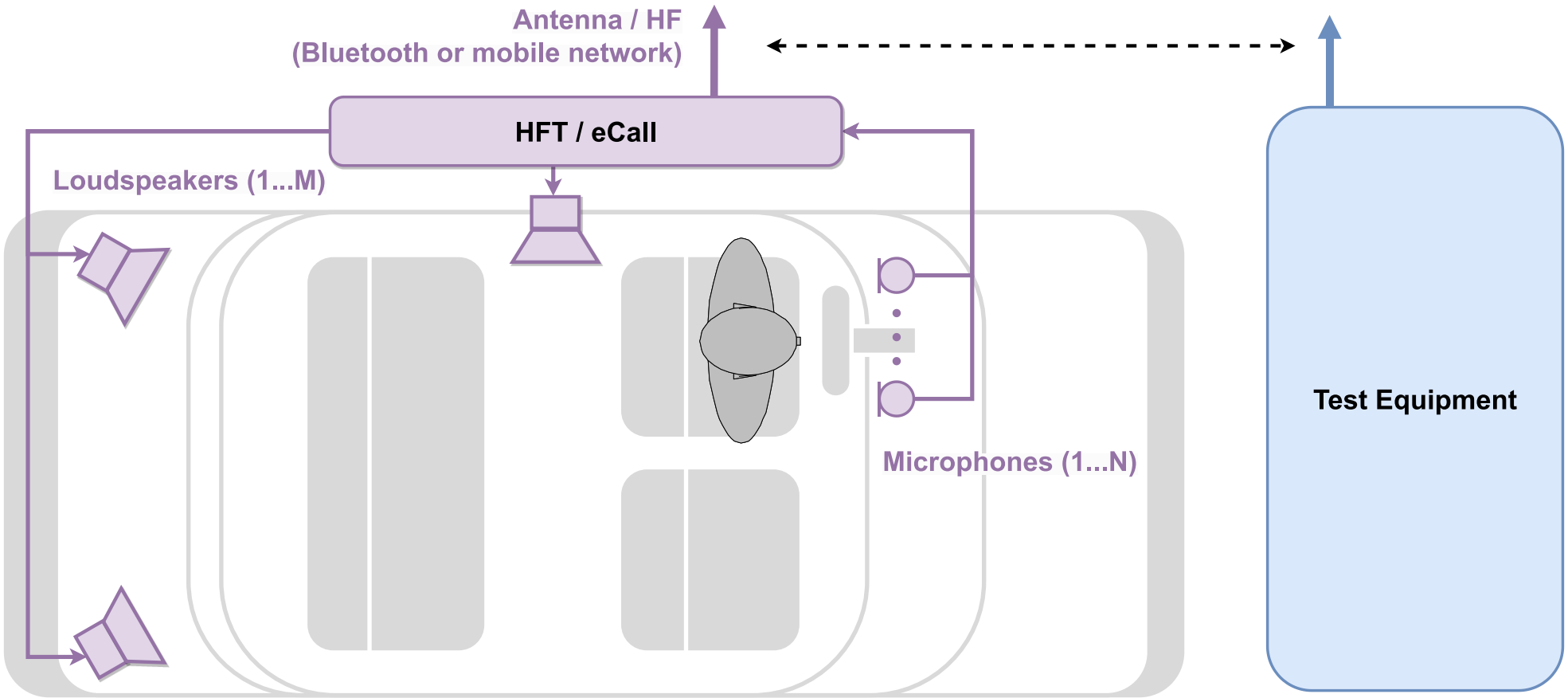
Testing principles – Example setup

- Acoustic test equipment: Head-and-torso simulator (HATS) (→ P.58)



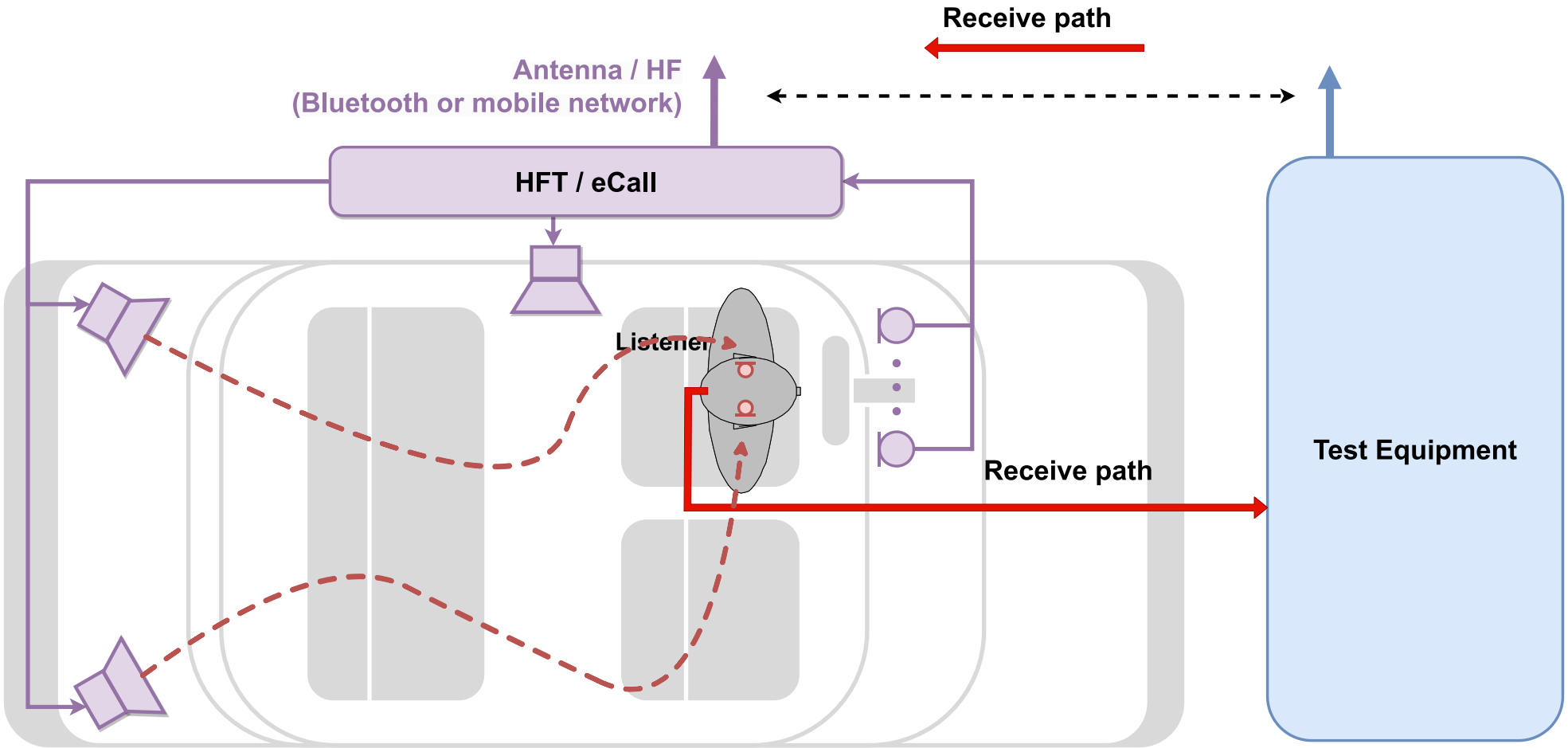
Testing principles – Example setup

- Acoustic test equipment: Head-and-torso simulator (HATS) (→ P.58)
- Electrical interface: Bluetooth, 3G/4G/5G radio tester, etc.



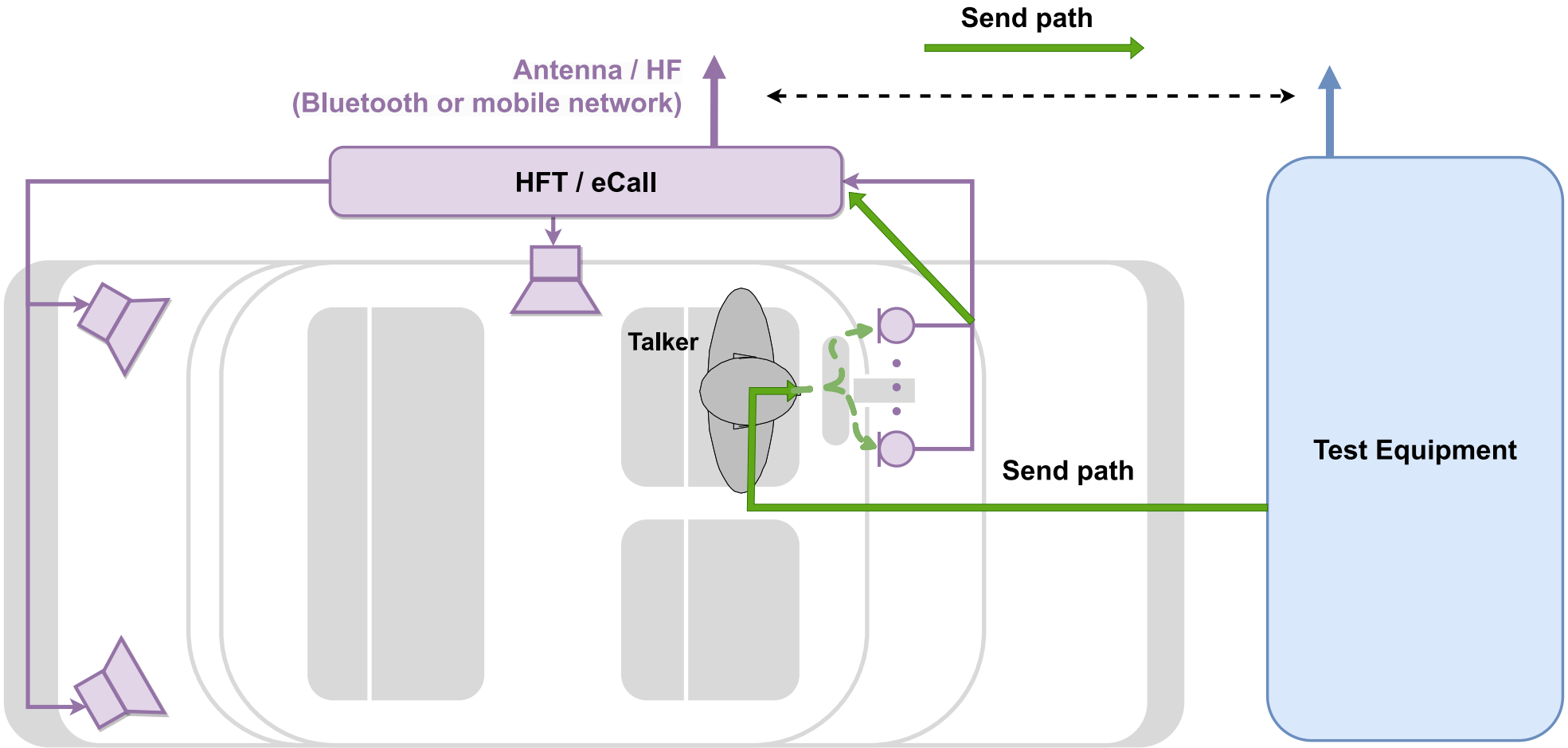
Testing principles – Example setup

- Acoustic test equipment: Head-and-torso simulator (HATS) (→ P.58)
- Electrical interface: Bluetooth, 3G/4G/5G radio tester, etc.



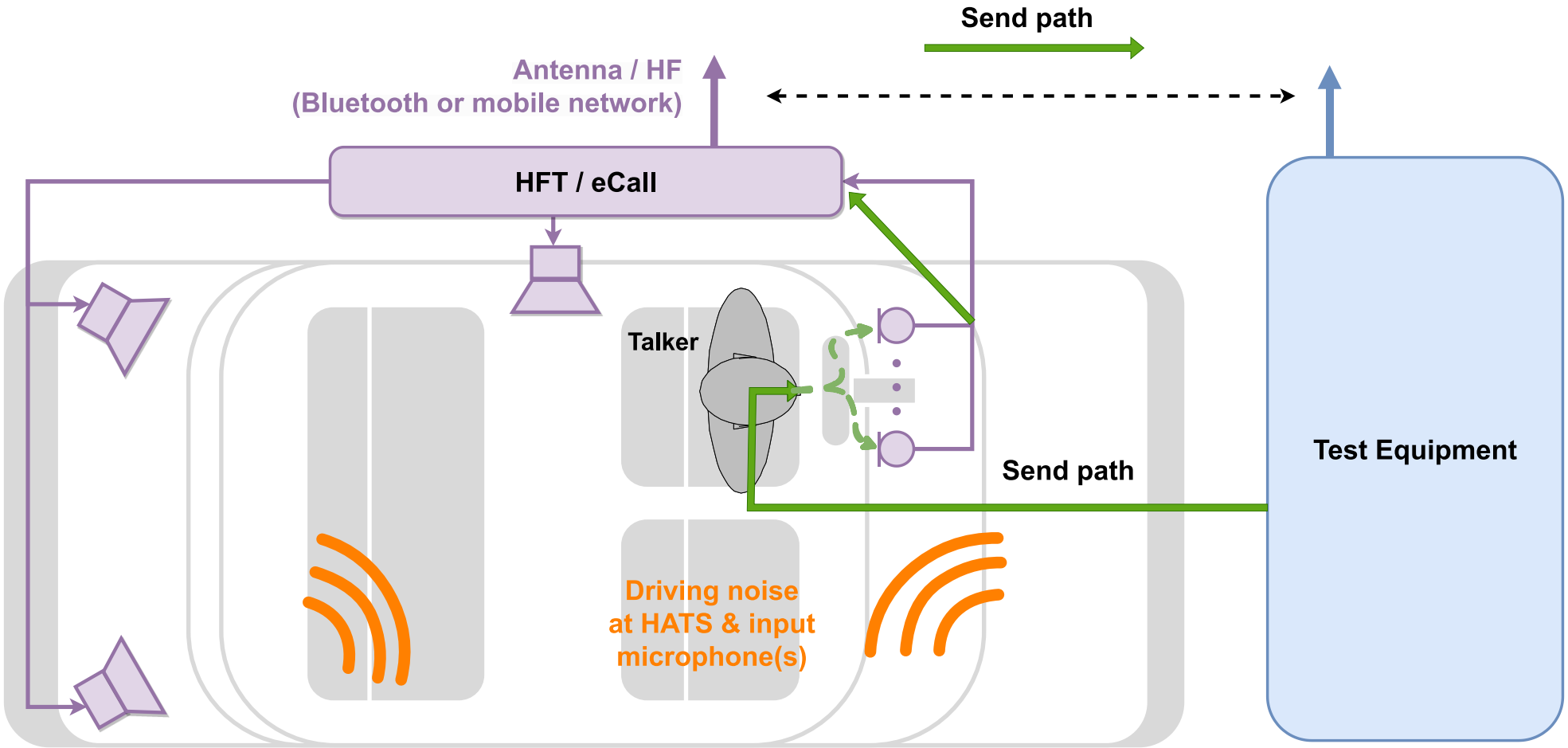
Testing principles – Example setup

- Acoustic test equipment: Head-and-torso simulator (HATS) (→ P.58)
- Electrical interface: Bluetooth, 3G/4G/5G radio tester, etc.



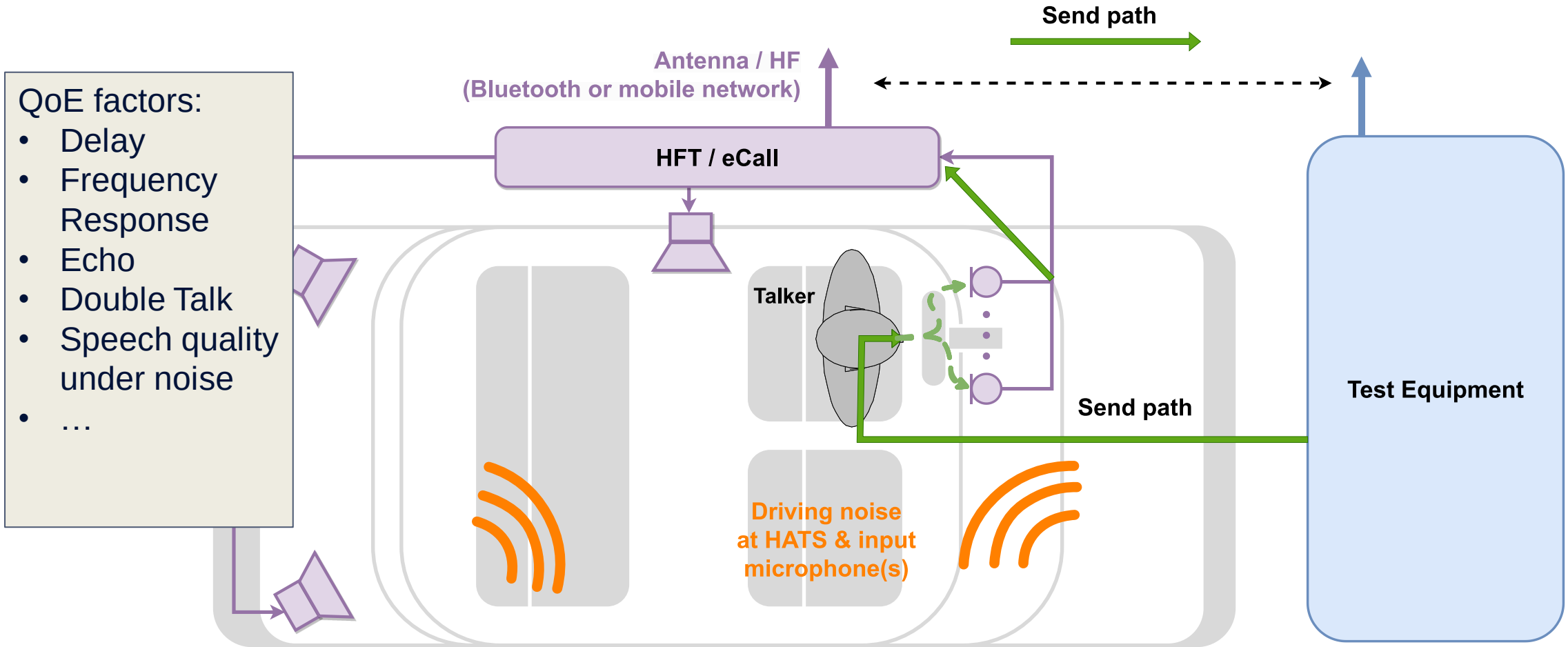
Testing principles – Example setup

- Acoustic test equipment: Head-and-torso simulator (HATS) (→ P.58)
- Electrical interface: Bluetooth, 3G/4G/5G radio tester, etc.



Testing principles – Example setup

- Acoustic test equipment: Head-and-torso simulator (HATS) (→ P.58)
- Electrical interface: Bluetooth, 3G/4G/5G radio tester, etc.



New work item P.ASR

New work item P.ASR

- Title of new Recommendation:
  "Specification of test methods, guidelines and performance metrics for automatic speech recognition (ASR)" (title is subject to change / WIP)

New work item P.ASR

- Title of new Recommendation:
  "Specification of test methods, guidelines and performance metrics for automatic speech recognition (ASR)" (title is subject to change / WIP)
- Different degrees of complexity in performance measure:
  - WER
  - Wake Word Detection
  - Commands
  - NLP / semantics

New work item P.ASR

- Title of new Recommendation:
  "Specification of test methods, guidelines and performance metrics for automatic speech recognition (ASR)" (title is subject to change / WIP)
- Different degrees of complexity in performance measure:
  - WER
  - Wake Word Detection
  - Commands
  - NLP / semantics

- Speech commands cannot be generic – recreate on a regular basis?

New work item P.ASR

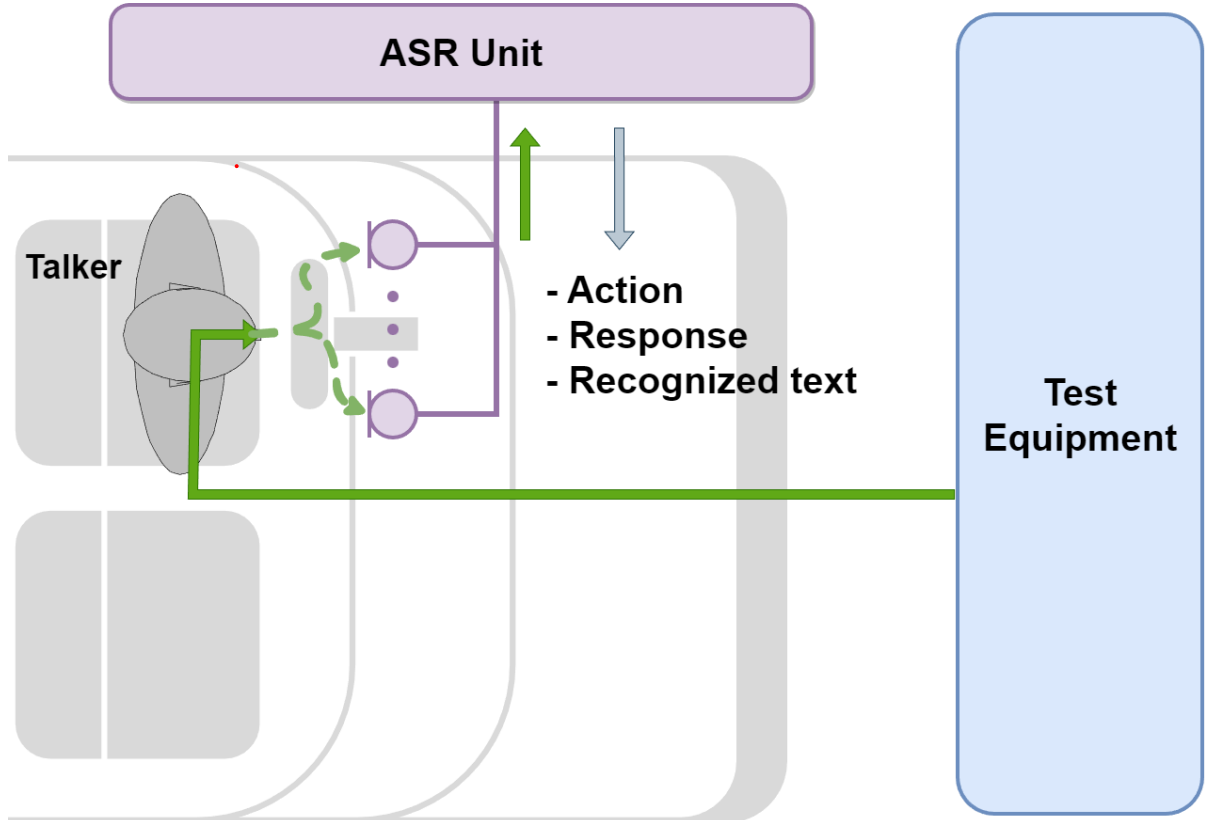
- Title of new Recommendation:
  "Specification of test methods, guidelines and performance metrics for automatic speech recognition (ASR)" (title is subject to change / WIP)
- Different degrees of complexity in performance measure:
  - WER
  - Wake Word Detection
  - Commands
  - NLP / semantics

- Speech commands cannot be generic – recreate on a regular basis?
- Consideration of different ASR systems / classes

New work item P.ASR

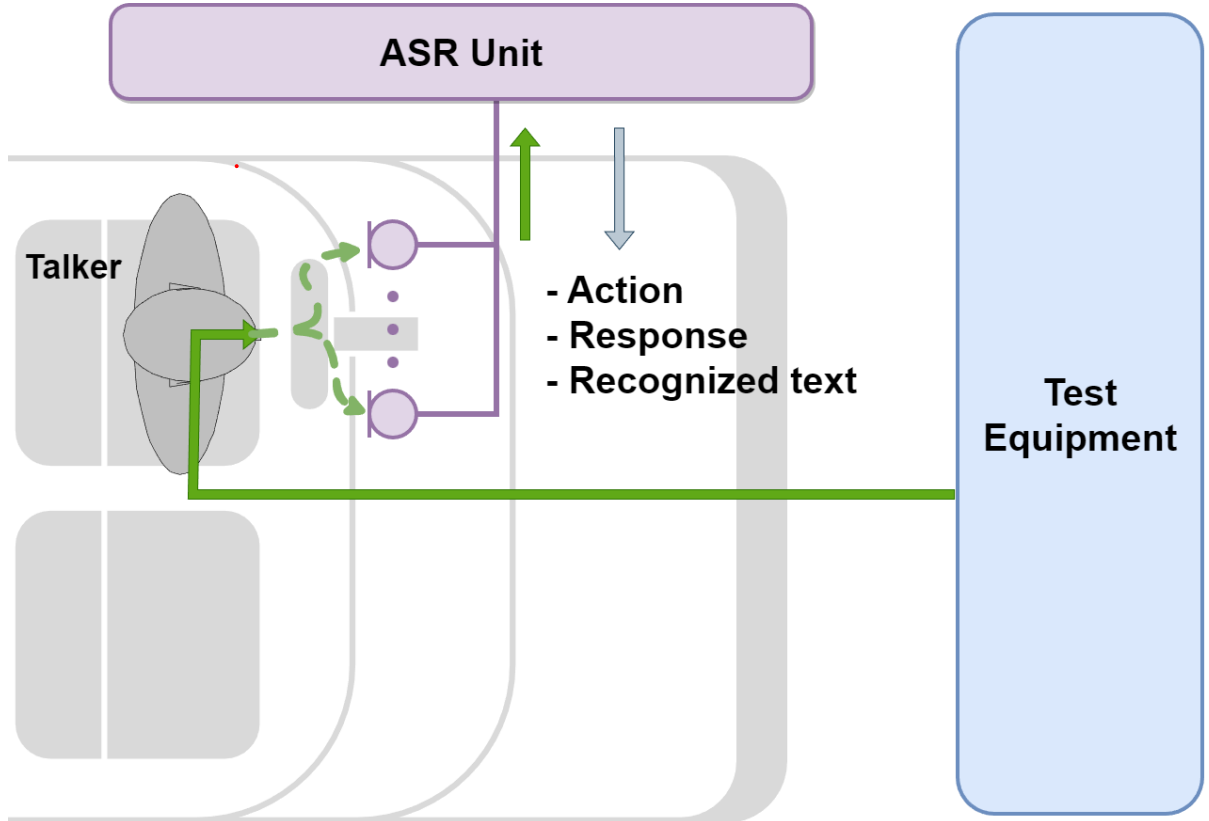
New work item P.ASR

- Testing should be generically possible (no custom internal interfaces), …



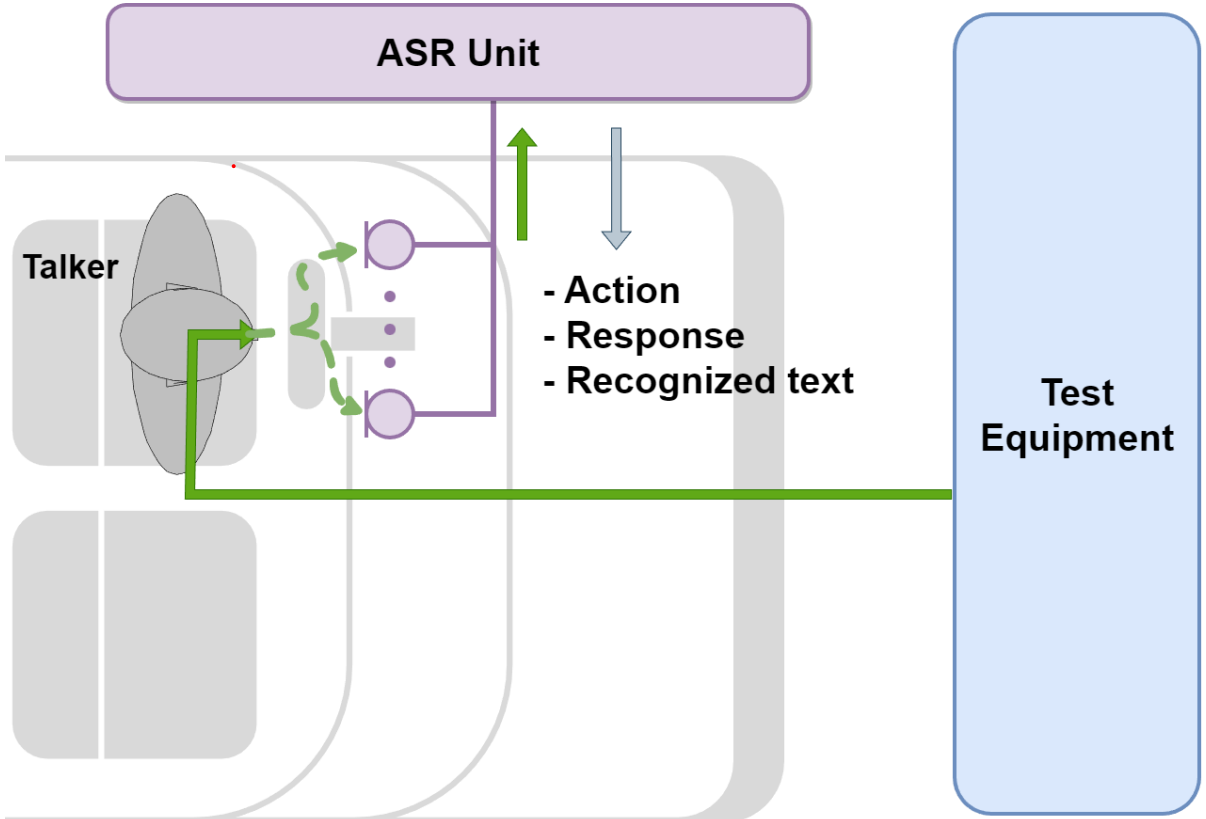
New work item P.ASR

- Testing should be generically possible (no custom internal interfaces), …
  - … but that will not work as expected.



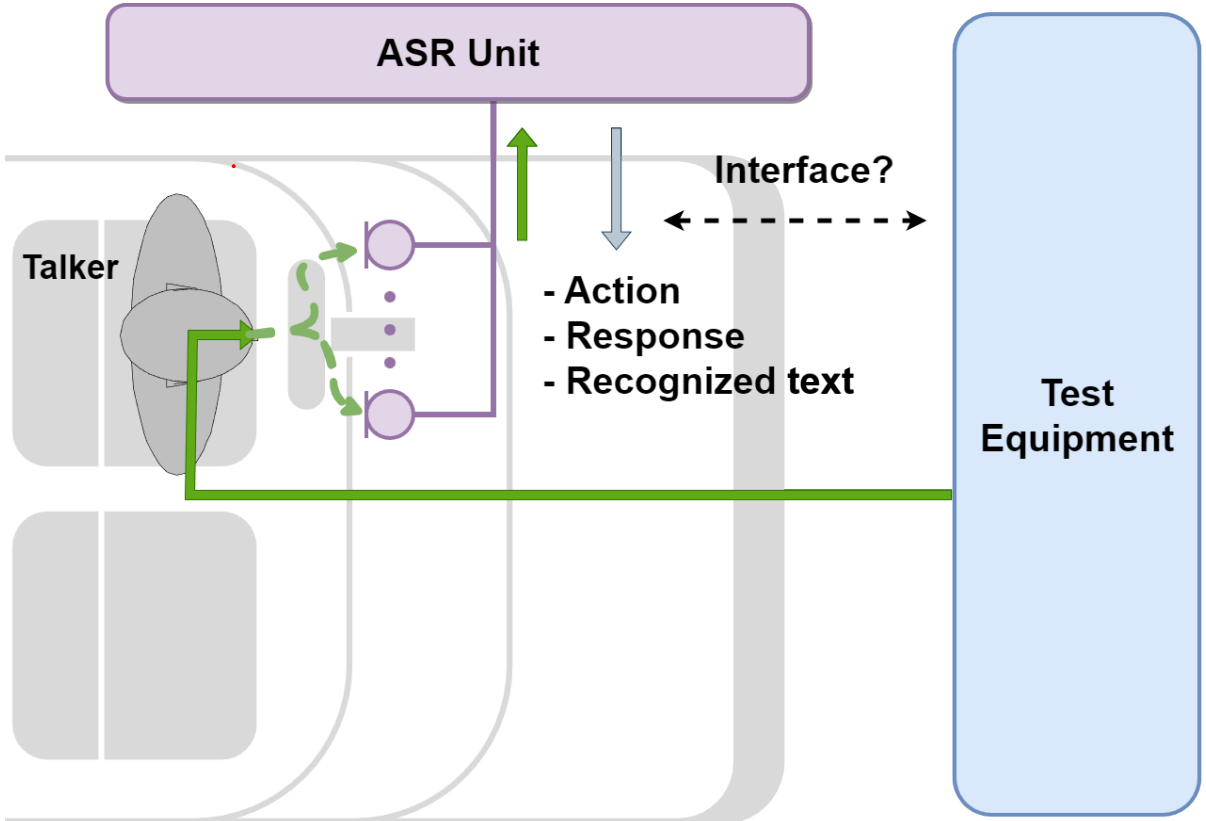
New work item P.ASR

- Testing should be generically possible (no custom internal interfaces), …
  - … but that will not work as expected.
- Existing acoustic interface can be used ✓



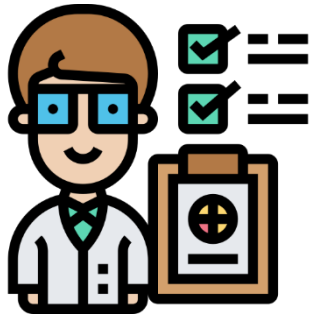
New work item P.ASR

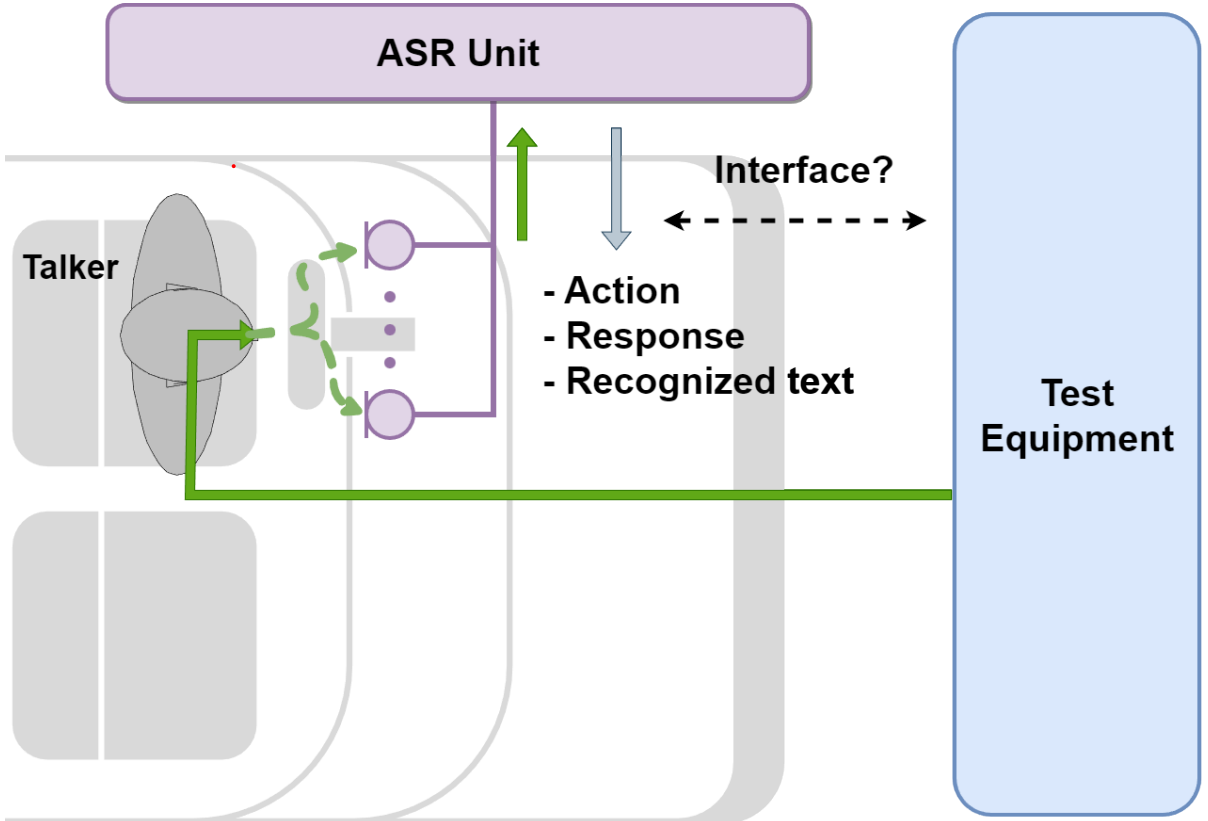
- Testing should be generically possible (no custom internal interfaces), …
  - … but that will not work as expected.
- Existing acoustic interface can be used ✓
- No standardized interface for …
  - Recognized words/commands
  - Action triggered
  - etc.



New work item P.ASR

- Testing should be generically possible (no custom internal interfaces), …
  - … but that will not work as expected.
- Existing acoustic interface can be used ✓
- No standardized interface for …
  - Recognized words/commands
  - Action triggered
  - etc.
- … use pen & paper?





Recent work in FG-VM & Q27/16

Recent work in FG-VM & Q27/16

- Focus Group on Vehicular Multimedia (FG-VM) published FGVM-03: "Implementation aspects of vehicular multimedia"

Recent work in FG-VM & Q27/16

- Focus Group on Vehicular Multimedia (FG-VM) published FGVM-03: "Implementation aspects of vehicular multimedia"

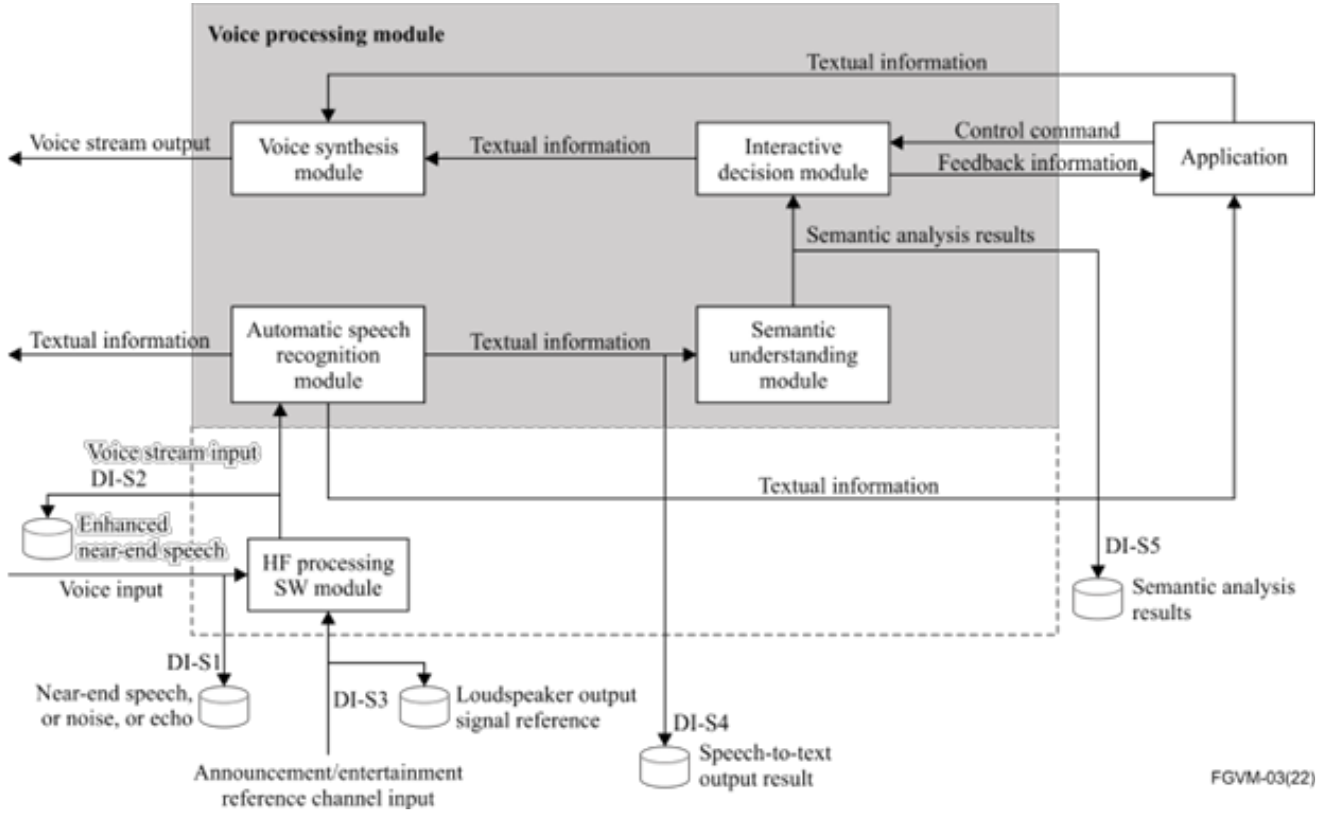


Fig. 6 of FGVM-03

Recent work in FG-VM & Q27/16

- Focus Group on Vehicular Multimedia (FG-VM) published FGVM-03: "Implementation aspects of vehicular multimedia"

- Diagnostic Interfaces (DI) are defined
  ➔ Part of H.VM-VMIA (?)

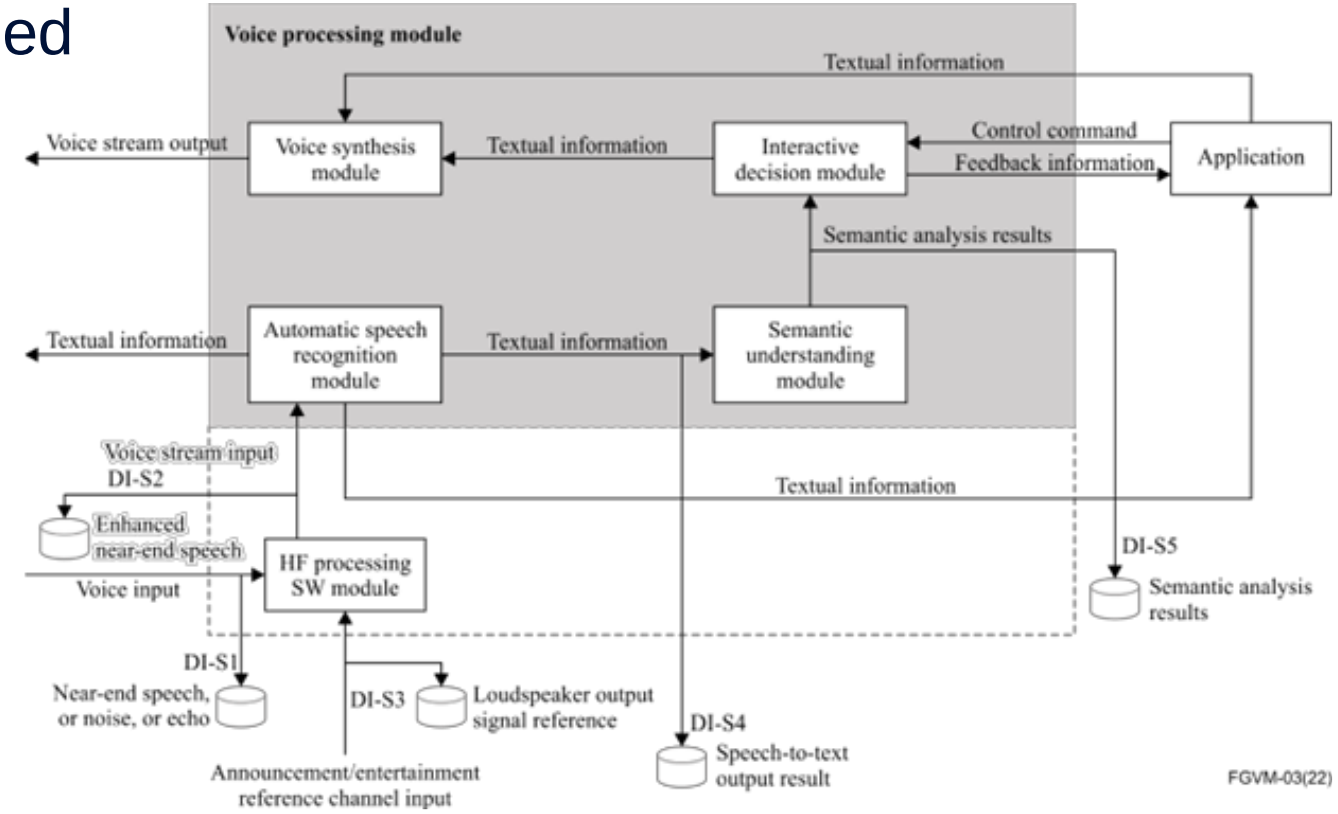


Fig. 6 of FGVM-03

Recent work in FG-VM & Q27/16

- Focus Group on Vehicular Multimedia (FG-VM) published FGVM-03: "Implementation aspects of vehicular multimedia"

- Diagnostic Interfaces (DI) are defined
  ➔ Part of H.VM-VMIA (?)

- DI-S4 & DI-S5 are highly relevant for performance testing in P.ASR!

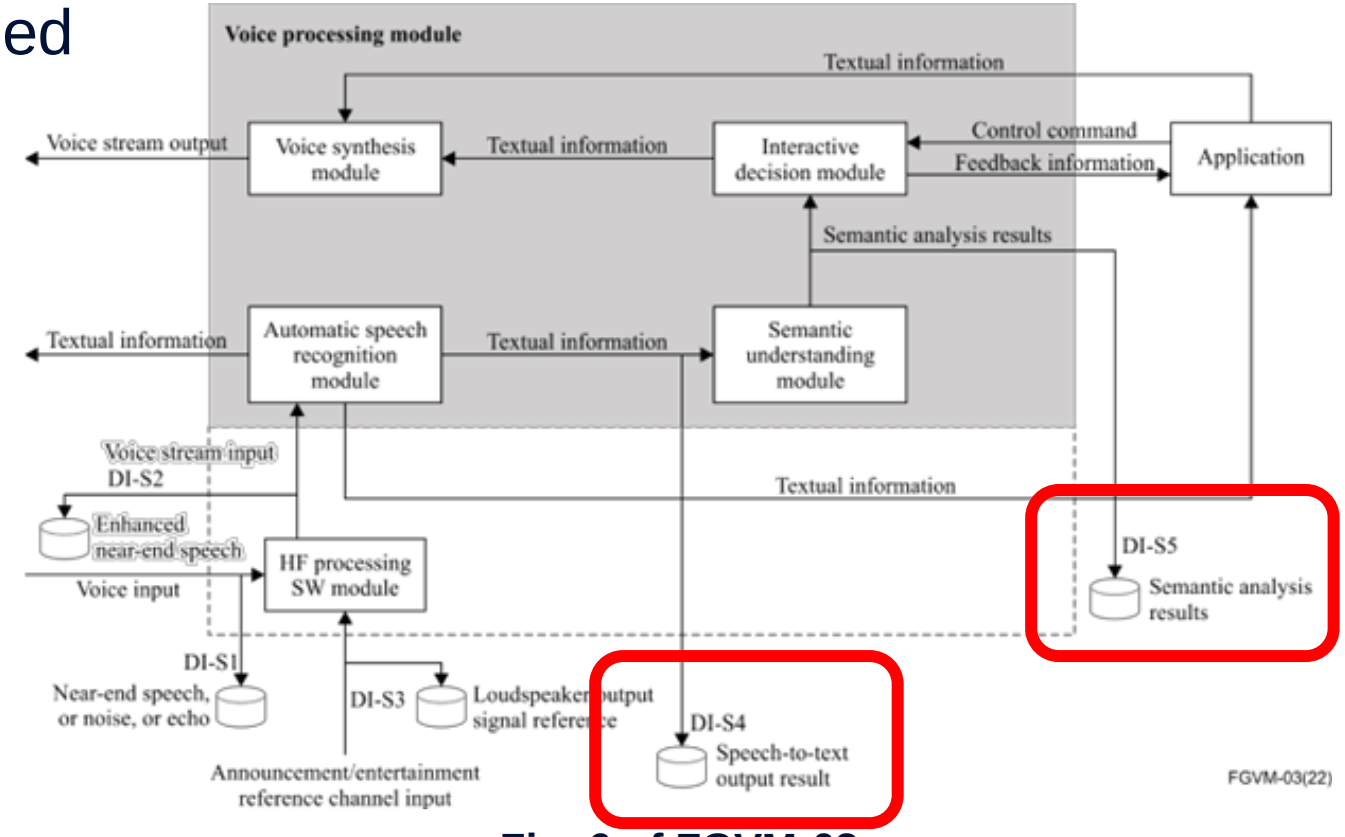


Fig. 6 of FGVM-03

Recent work in FG-VM & Q27/16

- Focus Gro
  "Implemen
- Diagnostic
  ➔ Part of I
- DI-S4 & DI
  for perform

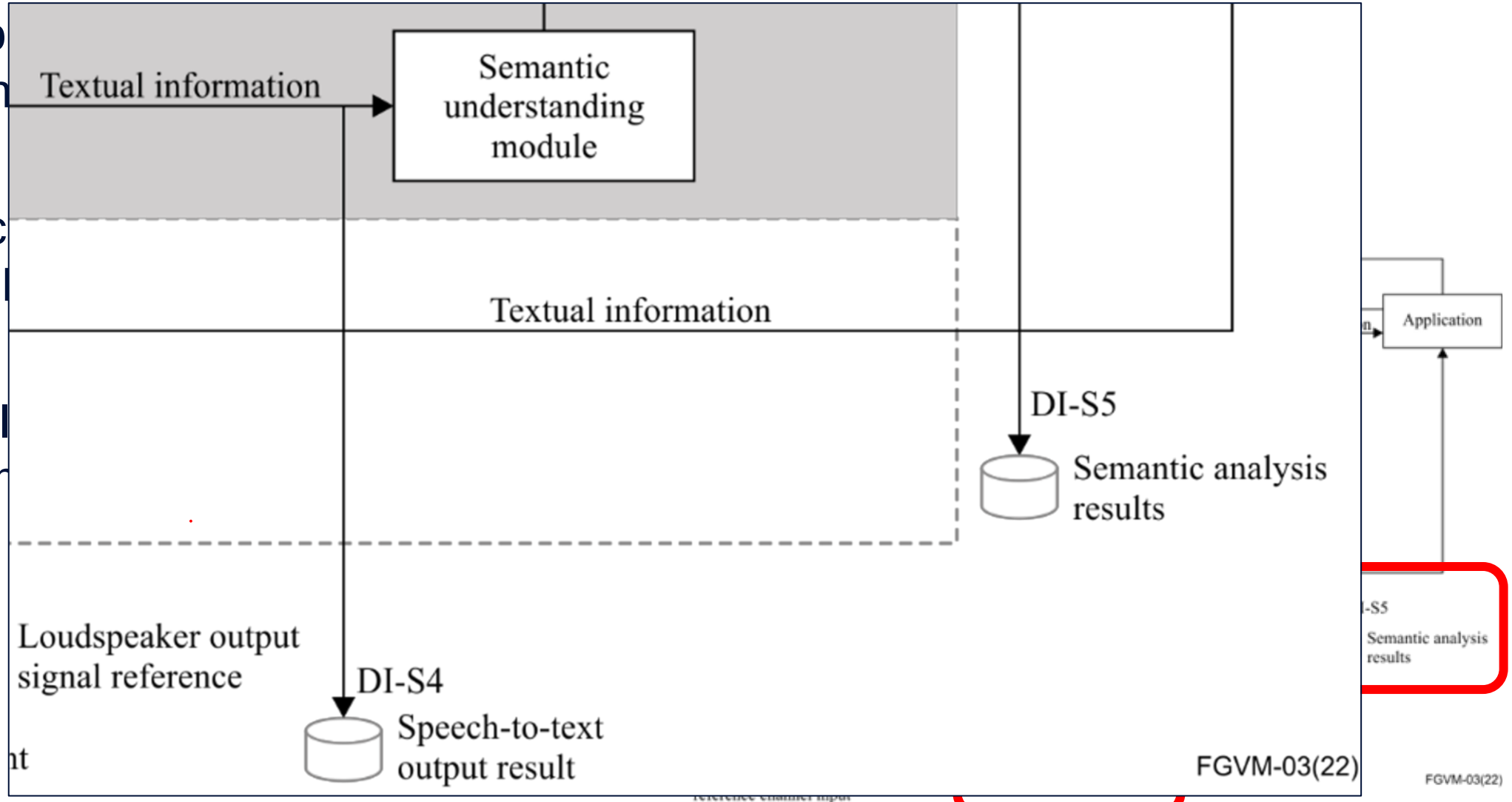


Fig. 6 of FGVM-03

Recent work in FG-VM & Q27/16

- Focus Group on Vehicular Multimedia (FG-VM) published FGVM-03: "Implementation aspects of vehicular multimedia"

- Diagnostic Interfaces (DI) are defined
  ➔ Part of H.VM-VMIA (?)

- DI-S4 & DI-S5 are highly relevant for performance testing in P.ASR!

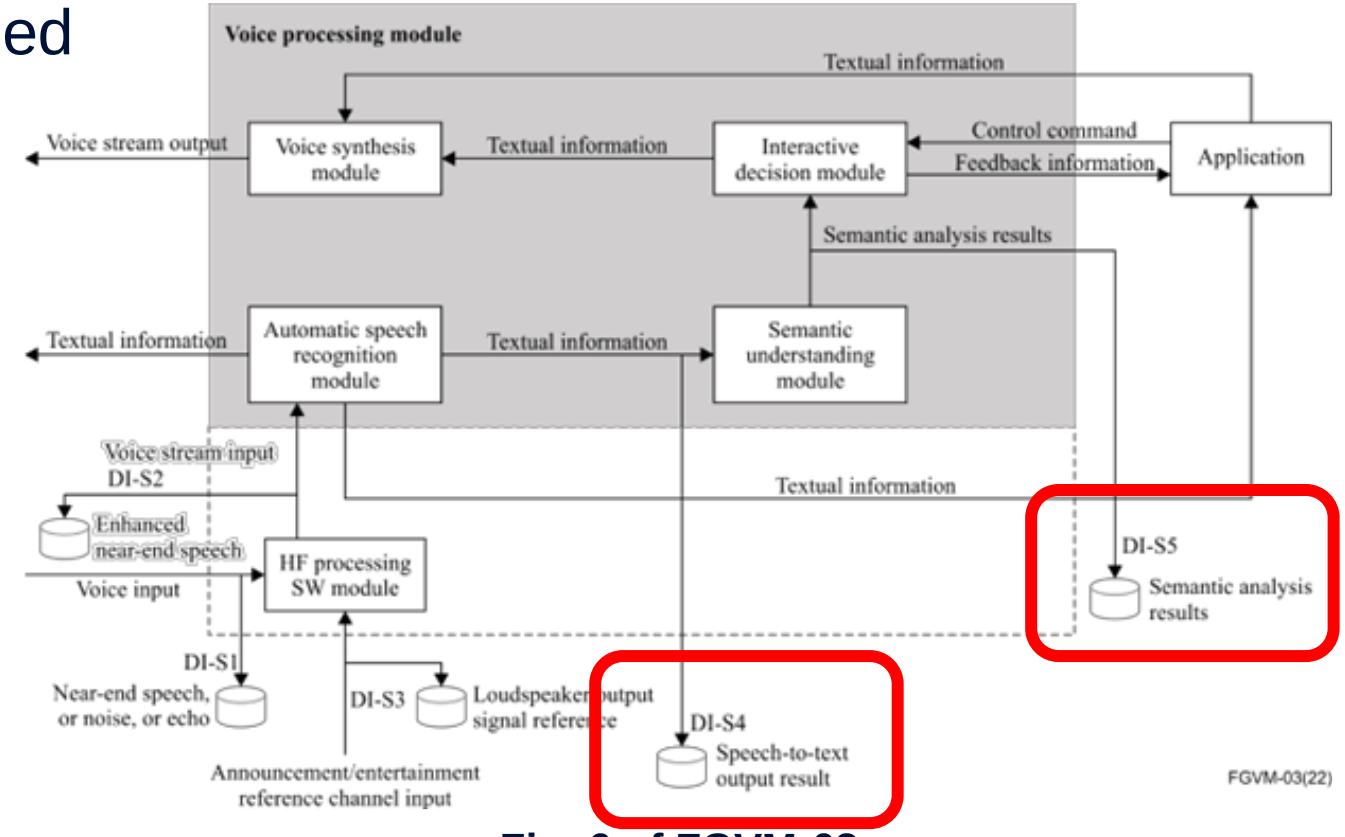


Fig. 6 of FGVM-03

Recent work in FG-VM & Q27/16

- Focus Group on Vehicular Multimedia (FG-VM) published FGVM-03: "Implementation aspects of vehicular multimedia"

- Diagnostic Interfaces (DI) are defined
  ➔ Part of H.VM-VMIA (?)

- DI-S4 & DI-S5 are highly relevant for performance testing in P.ASR!

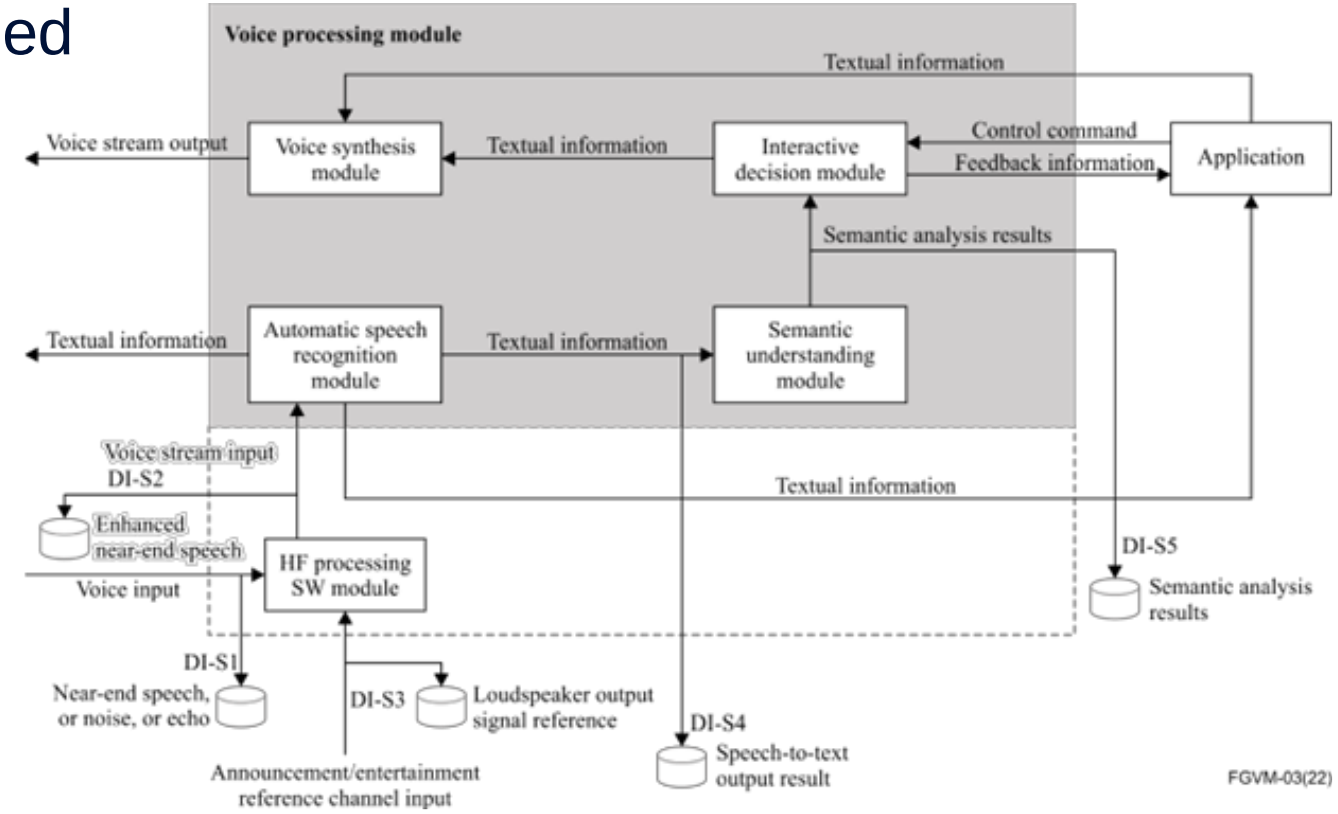


Fig. 6 of FGVM-03

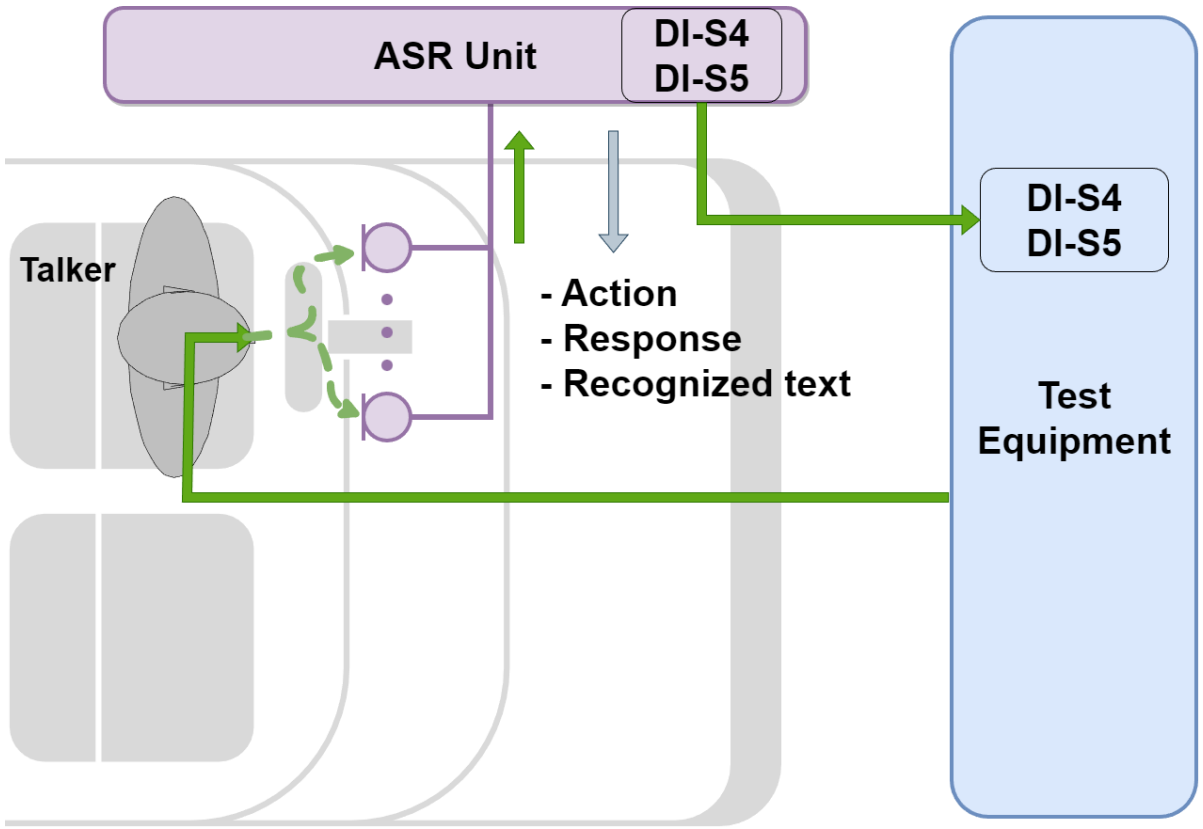
Recent work in FG-VM & Q27/16

Recent work in FG-VM & Q27/16

- DI-S4 (word recognition) and DI-S5 (semantic result) are "the missing piece" in the test setup for P.ASR

Recent work in FG-VM & Q27/16

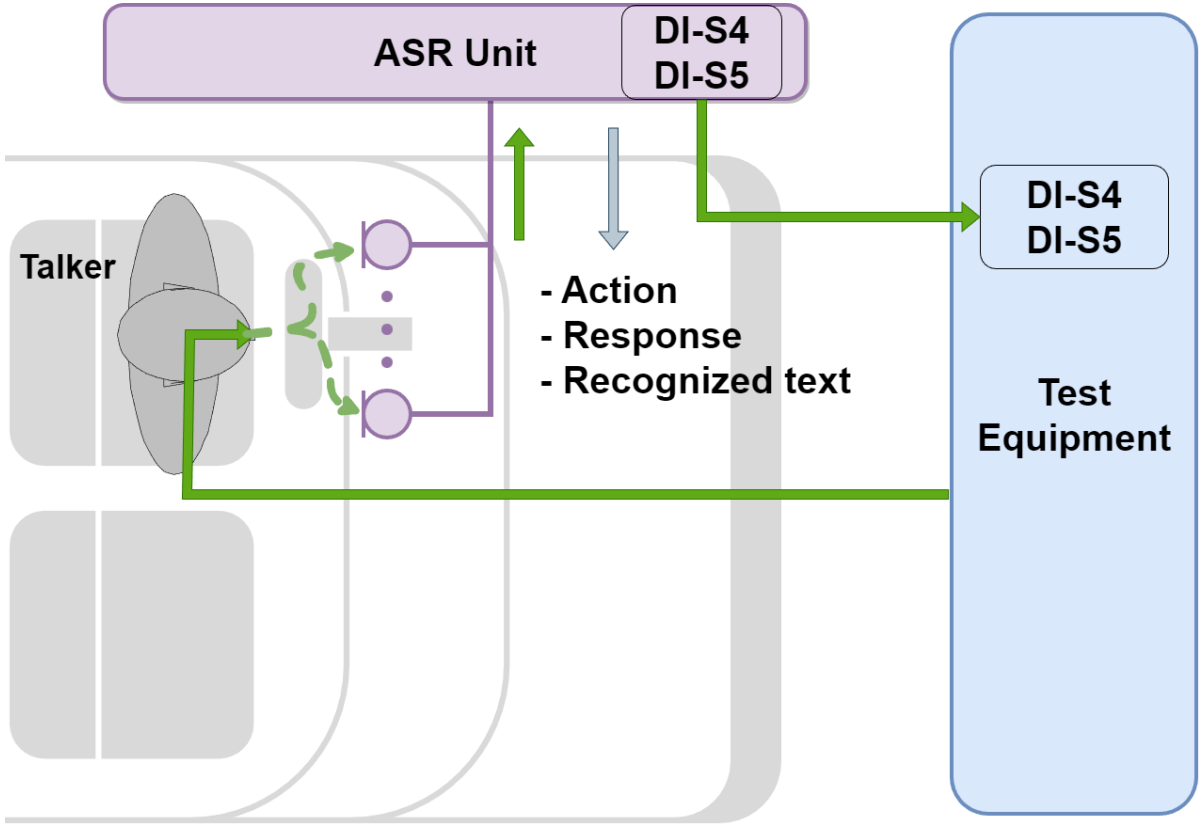
- DI-S4 (word recognition) and DI-S5 (semantic result) are "the missing piece" in the test setup for P.ASR



Recent work in FG-VM & Q27/16

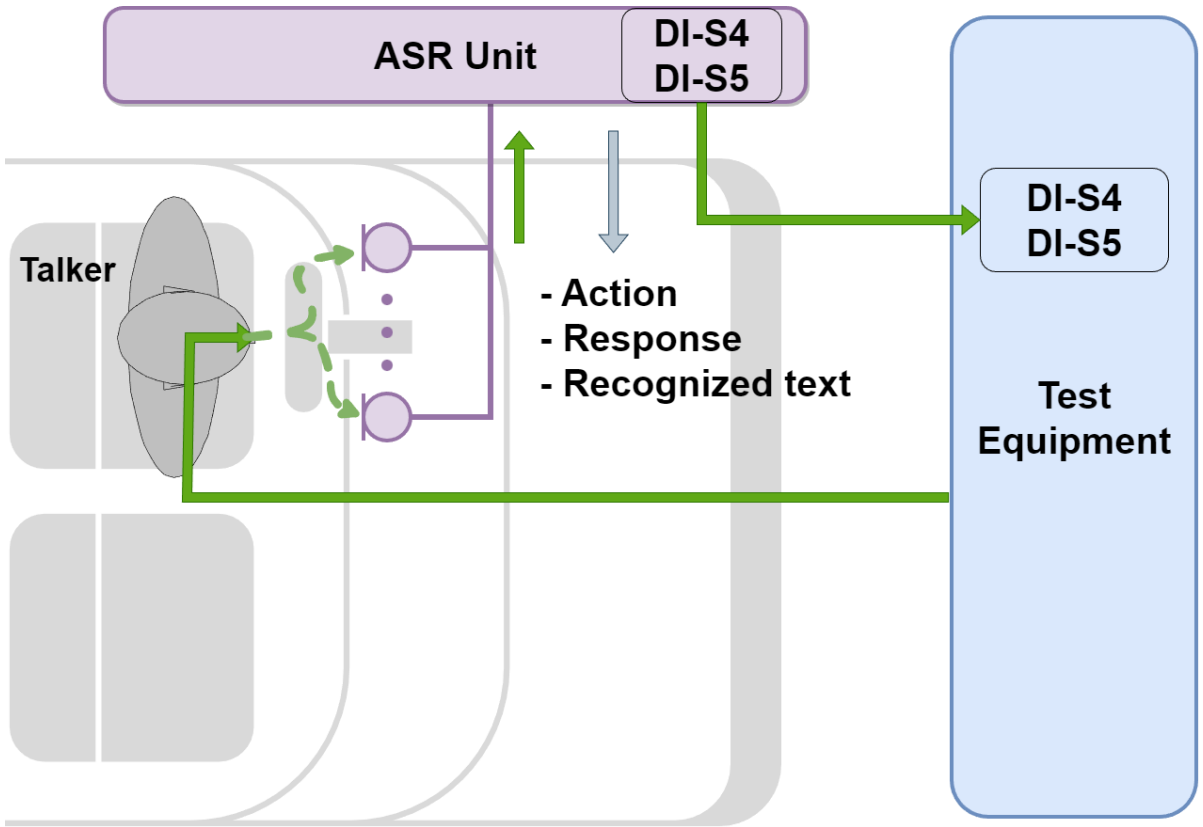
- DI-S4 (word recognition) and DI-S5 (semantic result) are "the missing piece" in the test setup for P.ASR

- Allows automated testing of voice assistant functionality



Recent work in FG-VM & Q27/16

- DI-S4 (word recognition) and DI-S5 (semantic result) are "the missing piece" in the test setup for P.ASR
- Allows automated testing of voice assistant functionality
- Signal-based DIs:
  - Simulated acoustical insertion
  - Offline testing/qualification of systems



Conclusion

Conclusion

- Q4/12 works on test methods and performance requirements for vehicle-based speech communication terminals

Conclusion

- Q4/12 works on test methods and performance requirements for vehicle-based speech communication terminals
- "Black box approach" for testing – no access to internals required
  - Worked fine in the past decades …
  - … but measurement of speech recognition is quite different!

Conclusion

- Q4/12 works on test methods and performance requirements for vehicle-based speech communication terminals
- "Black box approach" for testing – no access to internals required
  - Worked fine in the past decades …
  - … but measurement of speech recognition is quite different!

- Diagnostic Interfaces (DI) as introduced in FG-VM03 are essential for P.ASR

Conclusion

- Q4/12 works on test methods and performance requirements for vehicle-based speech communication terminals
- "Black box approach" for testing – no access to internals required
  - Worked fine in the past decades …
  - … but measurement of speech recognition is quite different!

- Diagnostic Interfaces (DI) as introduced in FG-VM03 are essential for P.ASR
- Q4/12 would appreciate to receive updates on the further normative work on H.VM-VMIA in Q27/16, in particular on specification of DIs.

Conclusion

- Q4/12 works on test methods and performance requirements for vehicle-based speech communication terminals
- "Black box approach" for testing – no access to internals required
  - Worked fine in the past decades …
  - … but measurement of speech recognition is quite different!

- Diagnostic Interfaces (DI) as introduced in FG-VM03 are essential for P.ASR
- Q4/12 would appreciate to receive updates on the further normative work on H.VM-VMIA in Q27/16, in particular on specification of DIs.

- H.VM-VMIA should complement existing Recommendations P.1100-1150, cross-check on possible redundancy/contradictions

HEAD acoustics GmbH
Ebertstraße 30a
52134 Herzogenrath
Germany

info@head-acoustics.de
www.head-acoustics.com

Follow us on