Question 4/12 – Objective methods for speech and audio evaluation in vehicles

(Continuation of Question 4/12)

1 Motivation

Car infotainment systems, telematic services and all types of mobile communication services are used increasingly in vehicles; an increasing number of modern cars are equipped with integrated infotainment, communication systems and connection possibilities to personal devices such as smartphones. In order to provide a good user experience, low driver distraction, satisfying communication quality and optimum dialog quality for all speech based services under all driving conditions, a variety of user interfaces and technologies have to seamlessly interact and to be optimized for the car environment. All services and technologies deployed in the car should not distract the driver from his main task. Advanced hands-free devices are required which require sophisticated signal processing adapted to the individual car to provide superior speech quality for the driver as well as for far end conversational partner. The special needs for emergency calls need to be addressed. Sophisticated speech recognition and dialog systems are needed to use speech based services in the car. In-car communications systems need to be optimized to provide a mostly natural speech enhancement for all types of in-car communications. Zoning concepts allowing the use of different audio-/ speech-based services in different zones within vehicles need to be considered.

The use of headsets or other hands-free devices, is becoming mandated in an increasing number of countries and states throughout the world. A large percentage of the target market for these vehicles will own headsets prior to purchasing a vehicle equipped with infotainment systems. They will expect to continue to use them in the vehicle, and thus will expect the vehicle to exploit the headset. The introduction of wireless headsets (e.g. Bluetooth, 802.11, DECT) requires the definition of standard behaviour and interactions with the vehicle.

So far Recommendations were developed describing the transmission requirements and test methods for narrowband, wideband, super-wideband speakerphones, for subsystems in cars, for emergency call communication and In-Car Communication (ICC).

The study within the Question is based on the existing Recommendations P.340, P.313, P.501, P.502, P.583, P.1100, P.1110, P.1120, P.1130, P.1140, P.1150. The main focus of the Question will be updated tests and requirements for hands-free systems including emergency call systems, subsystem requirements in cars, in car communication systems, speech recognition and speech dialog systems and requirements on the design of user interfaces in the car. A special focus needs to be given on the requirements for autonomous driving in the context of speech and audio in cars.

The following Recommendations, in force at the time of approval of this Question, fall under its responsibility:

P.1100, P.1110, P.1120, P.1130, P.1140, P.1150.

2 Question

The following items are to be considered within the study of the Question:
– How can the driving situation be simulated while covering the most relevant parameters influencing driver distraction and the speech quality within a laboratory environment?

– What requirements and design guidelines are needed for user interfaces in the car?
– Are there communicational speech quality parameters in the driving situation not yet covered by the existing Recommendations?
– What additional aspects need to be taken into account in emergency call communications?
– Which additional parameters determine the quality of in-car communication systems and how can they be assessed?
– What are the most influential parameters for speech recognition systems in the driving situation?
– How can we assess and quantify the dialog quality of human-machine interfaces in cars?
– Which of the newly developed methodologies known in ITU can be used and/or adapted to the car hands-free situation?
– Do different mobile networks and network configurations or OTT solutions require individual setups for specific parameters?
– What is the appropriate behaviour of a wireless or wired headset or a hearing aid in the environment of a telematics enabled motor vehicle?
– What are the desirable features to be presented by the vehicle, and what is their behaviour when operating with a smartphone connected to the car or when connecting services directly to the car's head unit?
– What enhancements of the Recommendations P.1100, P.1110, P.1120, P.1130, P.1140 and P.1150 are needed to be developed ensuring seamless support for users of hands-free devices and ICC systems?
– Which applications and requirements in the context of speech and audio needs to be addressed for autonomous driving?

3   Tasks

Tasks include, but are not limited to:

– define the typical operating conditions to be simulated covering the most relevant parameters influencing the speech quality within a laboratory environment;
– define the typical operating conditions to be simulated covering the most relevant parameters influencing the quality of in-car communication systems within a laboratory environment;
– define the typical operating conditions to be simulated covering the most relevant parameters influencing automated speech recognition performance within a laboratory environment;
– define the typical operating conditions to be simulated covering the most relevant parameters influencing dialog systems performance within a laboratory environment;
– definition of the environmental conditions for testing the car hands-free terminal and verifying its acoustical performance characteristics under typical operating conditions;
– definition of the environmental conditions for testing the car hands-free subsystems and verifying their performance characteristics under typical operating conditions including the definition of QoS classes for such (sub-)systems;
– specification of all relevant transmission characteristics;

- definition of test signals and testing techniques for emergency call systems with special focus on speech intelligibility/listening effort;
- definition of test procedures for evaluating automated speech recognition;
- definition of test procedures for dialog systems in cars;
- define requirements for ICT systems that interact with drivers of vehicles;
- identify the needs in the area of speech and audio for autonomous driving and derive relevant test scenarios and requirements.

An up-to-date status of work under this Question is contained in the SG12 work programme http://www.itu.int/ITU-T/workprog/wp_search.aspx?q=4/12.

4   Relationships

WSIS Action Lines

- C2

Sustainable Development Goals

- 9

Recommendations

- P.340, P.313, P.381, P.382, P.501, P.502, P.581, P.582, P.TBN, P.DHIP

Questions

- 5/12, 6/12, 9/12

Study Groups

- ITU-T SG16

Other bodies

- ITU-R, 3GPP SA4, ETSI TC STQ, ETSI TC ITS, Bluetooth SIG, ISO TC22, ISO TC204