Challenges in the introduction of wideband hands-free in cars

H.W. Gierlich

HEAD acoustics GmbH, 52134 Herzogenrath, Germany, Email: H.W.Gierlich@head-acoustics.de

Introduction

The introduction of wideband speech services in the new generation networks will be one of the major steps in the near future. The support of such services by the different networks and services also forms the basis for the integration of wideband speech hands-free services in cars. Wideband speech quality would be mostly welcomed in cars since especially on the receiving side the poor (narrowband) speech quality is more obvious in cars than in other type of terminals. The narrowband speech is compared directly to the typically high quality audio presentation of all other audio equipment in the car and even less sensitive users observe the difference in quality immediately. To achieve a superior speech quality inside the car all elements involved in the transmission have to provide adequate speech quality. The challenges faced when connecting the hands-free car system via Bluetooth to the mobile terminal, via the mobile terminal to the network and connecting the mobile network to the wideband NGN type network is introduced, the users expectation with respect to the different quality dimensions of wideband speech and the wideband presentation of narrowband speech signals is discussed. An overview of the standardization work in the ITU-T Focus group CarCom is given.

The System Configuration

The implementation of car hands-free systems is rather complex involves the variety of different components. A typical example for car hands-free system is given in **figure 1**.

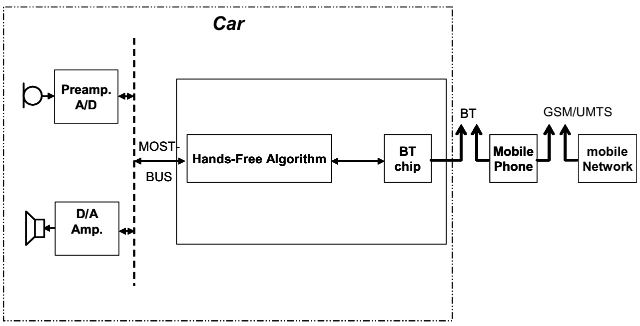

Figure 1: Components of a modern car hands-free implementation in wideband

The first part of a hands-free system is typically a microphone or a microphone arrangement which often is separated from the actual hands-free algorithm and is connected to the hands-free algorithm e.g. by a MOST-bus. The hands-free algorithm itself is typically implemented in a so called “head unit”, the construction in principal is probably the same as in narrowband hands-free systems. Also all signal processing elements such as echo cancellation, noise cancellation, microphone equalization and AGC functionality will be found. A quite innovative component for wideband car hands-free will be the Bluetooth link and its associated speech coding. Currently no wideband Bluetooth connection has been specified. There are quite a few quality requirements with respect to the Bluetooth link:

- low delay for transmission
- mostly transparent coding (ideally tandem free)
- Bluetooth coding shall not degrade audio quality
- More detailed specification for Bluetooth audio link required than currently in narrowband

The link between the integrated car hands-free system and the mobile network is typically the mobile phone which on the one hand side is using the Bluetooth connection to interface with the car hands-free system and which is on the other side is used to interface with a mobile network. The speech coding used by the mobile phone is most likely AMR-wideband.

In downlink again the Bluetooth link is used to interface with the car hands-free system. In addition to narrowband car hands-free systems wideband hands-free systems may offer techniques to enhance also the narrowband speech quality. Examples can be found in ([1], [2]). Artificial bandwidth extension may be used in order to achieve also higher audio quality in narrowband calls. This function may be useful to enhance the speech sound quality in narrowband calls and such reduce the perceived audio quality degradation compared to wideband calls.

For wideband car hands-free systems the integrated car audio system will play a major role. Due to the high quality speaker systems typically installed in a car and normally used for music reproduction a superior speech sound quality in downlink can be expected when introducing wideband in cars. However, care must be taken in order to deactivate all audio enhancement function in such systems which may be not desirable in car hands-free communication. Also any additional delay which might be inserted by these systems should be avoided.

System Performance Requirements

In order to successfully launch a wideband service a superior service quality has to be provided. This superior quality concerns all aspects of speech communication in wideband:

- speech quality in the talking situation
- speech quality in the listening situation
- speech quality in the conversational situation

If all components are implemented properly it can be expected that the wideband service will improve the perceived speech quality for the driver significantly. Superior speech quality which can be achieved by high quality speech coding in conjunction with a high quality audio playback system in a car will contribute to less driver's distraction and enhanced speech intelligibility. It is known (see [3]) that wideband communication will increase speech intelligibility by up to 15 %. Especially in noisy driving conditions which is the normal condition in the car typically wideband will help significantly to improve the listening situation.

Speech Quality in Receiving Direction

The components in a car hands-free systems which contribute to the speech quality typically are the loudspeaker system and it's associated equalization, any type of AGC or other speech enhancement algorithms in the HFT processing and the speech coding – in the mobile phone and possibly also transcoding in the Bluetooth link. The main system requirements in downlink needed to achieve a superior wideband quality are as follows:

- **loudness ratings**

Provisionally (until a new loudness calculation scheme is developed) the current loudness rating calculation is described in ITU-T P.79 Annex A [4] can be used. As in narrowband, sufficient headroom should be provided to increase the listening level in the car in such a way that under all driving condition a sufficient SLR can be obtained.

- **frequency response**

The frequency response measured in receive should be mostly flat between 100 Hz and 8 kHz. The measurement is based on HATS [5]. Currently the work in ITU-T Focus Group CarCom requires the use of both artificial ears which are voltage summed and averaged in order to obtain the frequency response characteristics. By averaging both ear signals, the measured response characteristic is more smoothened and reflects more the average characteristics than the use of one artificial ear only.

- **Listening speech quality**

Even more important than in narrowband is the verification of the speech sound quality. An appropriate test method is TOSQA2001 [6].

Speech Quality in Sending Direction

The system design in sending direction is much more critical than for narrowband hands-free systems. The extended bandwidth leads to the problem that the low frequency background noise produced by the car disturbs the transmitted send signal more heavily than in narrowband. On the other hand subjective tests carried out with different types and levels of background noises [7] indicate that subjects prefer mostly in all conditions the full band transmission without heavy highpass or lowpass filtering of the transmitted signals. As a consequence it is desirable in wideband systems to put more emphasis on the microphone arrangement and on the implementation of noise cancellation systems. Currently it seems that single microphone systems probably are not sufficient to achieve the desired noise cancellation for wideband. The most important performance parameters in sending direction are:

- **loudness ratings**

As in the receiving direction, the current wideband loudness ratings as defined in Annex A, ITU-T Recommendation P.79 [4] can be used.

- **frequency response**

The frequency response measured in sending should be mostly flat between 100 Hz and 8 kHz. A slight highpass-/lowpass- filtering might be applicable.

- **speech quality with background noise**

Speech quality in background noise situations is determined typically by two impairments: either the speech signal is impaired e.g. by speech sounding robotic, metallic or by musical tones which are produced by the background noise cancellation algorithm. Or the background noise itself is impaired e.g. by background noise modulations, interruptions, unnatural transmission. Both, the speech sound quality as well as the background noise transmission quality can be optimized in order to achieve good overall performance in the presence of different background noises. Therefore a new objective test method according to ETSI Standard EG 202 396-3 [8] can be used. This method was introduced in [9] and [10]. This objective test procedure allows to predict a speech quality in the presence of background noise by three parameters:

- S-MOS (quality of the transmitted speech)
- N-MOS (quality of the transmitted background noise)
- G-MOS (overall quality as perceived by the user)

This method which is also called 3QUEST is available for wideband terminals and shows a high correlation between subjective tests and objective prediction. S-MOS, N-MOS and G-MOS are predicted on a MOS scale ranging from 1 to 5 as defined in ITU-T Recommendation P.835 [11]. Figure 2 gives a block diagram of this method.

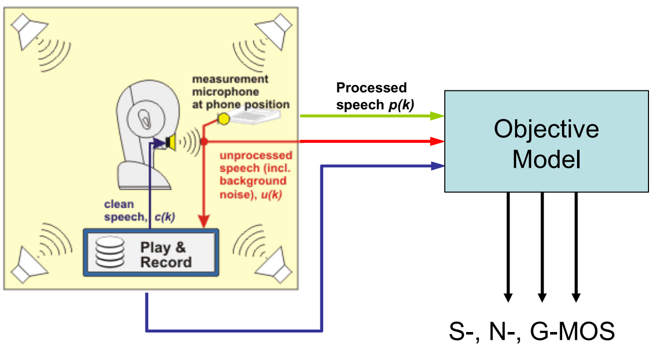

Figure 2: Block diagram of the objective test method for speech quality in background noise according to [8]

Echo Performance

Wideband systems typically require – besides the wideband transmission implemented in a mobile network – IP based transmission in the backbone. In general, IP based transmission will introduce higher delays than a classical PSTN based transmission. As a consequence the impact of any echo in wideband connections is more severe than in typical narrowband connections. Furthermore the sensitivity of humans on wideband echo is significantly higher than in narrowband systems. Recent investigations reported in [12] and [13] show that humans are especially sensitive for echoes in the frequency range between 3 and 6 kHz. This is due to the higher sensitivity of the hearing threshold in this frequency range as well as due to less self masking effects in this frequency range. This frequency range was not transmitted in narrowband system. As a consequence the echo loss requirements especially in a spectral domain has to be increased significantly. The requirement currently introduced in the ITU-T CarCom Specification is shown in figure 3.

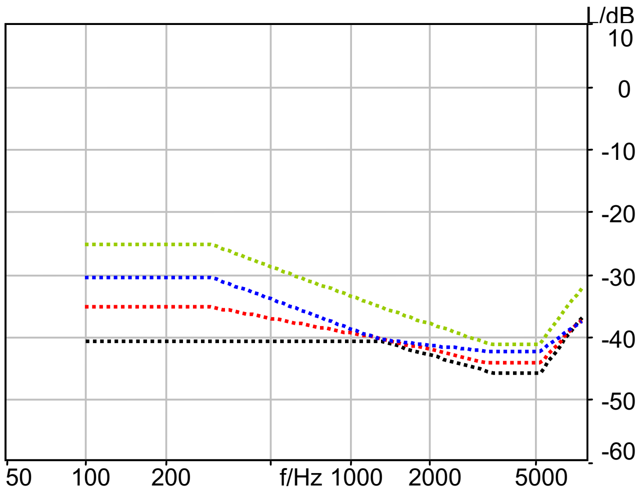

Fig. 3: Spectral Echo Loss Requirement in wideband hands-free systems, round trip delay:

green – 100ms, blue – 200ms, red – 300ms, black – 500ms

Delay and its impact on conversational quality

Even with perfect echo cancellation, perfect terminal characteristics and no other impairments delay will impair the conversational quality significantly. The relationship between delay and the perceived overall quality can be derived from ITU-T Recommendation G.107. [14] It can be seen, that for delays higher than appr. 150 to 200 ms user's satisfaction is rapidly reduced. Therefore it is highly desirable to minimize the delay in any component involved in a wideband transmission in order to achieve a superior service. For the car hands-free system this requires a minimum delay signal processing and a minimum delay link between the hands-free system and the mobile phone. A high delay (20 to 60 ms) on a Bluetooth link will **significantly** impair the conversational speech quality and therefore should be avoided. For the car hands-free system an overall delay of not more than 70 ms, distributed in sending and receiving direction is currently specified in the ITU-T FG CarCom Specification.

Double Talk Performance

Currently no new information exists on wideband double talk requirements. However, with some reasonable degree of confidence in a first approach the narrowband requirements can be used (see [15]). Also in wideband clearly the main impairing factors are switching and echo during double talk. It is likely that echo impairments in double talk as in single talk are more annoying but further subjective testing is required.

Conclusions

The introduction of wideband in cars certainly has the potential to significantly increase the speech quality especially for the user in the car. However, only a proper system design including all components involved in transmission will guarantee a superior service which is prerequisite for the acceptance of wideband services in car hands-free.

References

[1] E. Hänsler, G. Schmidt (Editors): Speech and Audio Processing in Adverse Environments, Springer ISBN978-3-540-70601-4, 2008

[2] P.Vary, P. Jax: On artificial bandwidth extension of telephone speech, Signal Processing Vol.83, Aug. 2003, ISSN 0165-1684

[3] H. Fletcher, R.H. Galt.: The perception of speech an it's relationship to telephony, JASA, 1950, Vol.22, No.2

[4] ITU-T Recommendation P.79: Calculation of Loudness Ratings for telephone sets, ITU-T, 1999

[5] ITU-T Recommendation P. 58: Head and torso simulator for telephonometry, ITU-T, 1996

[6] J. Berger: Instrumentelle Verfahren zur Sprachqualitäts-schätzung - Modelle auditiver Tests, PHD thesis, Kiel, 1998

[7] H. W. Gierlich, F. Kettler, S. Poschen, A. Raake, S. Spors: Optimum Frequency Responses for Wideband Terminals, ITU-T Workshop "From Speech to Audio: bandwidth extension, binaural perception", Lannion, 11/08

[8] ETSI EG 202 396-3: Speech Quality performance in the presence of background noise Part 3: Background noise transmission - Objective test methods, ETSI,11/2008

[9] H.W. Gierlich, F. Kettler, S. Poschen, J. Reimes: Objective Prediction of Speech Quality for Wideband Communication Scenarios including background noise, DAGA 07, Proceedings

[10] H.W. Gierlich, F. Kettler, S. Poschen, J. Reimes: Objectve Speech Squality for Wide-and Narrowband Communication Scenarios including Background Noise, EUSIPCO 2007, Lousanne, Proceedings

[11] ITU-T Recommendation P.835: Subjective test methodology for evaluating speech communication systems that include noise suppression algorithms, ITU-T, 11/2003

[12] S. Poschen, F. Kettler, A. Raake, S. Spors: Wideband Echo Perception, IWAENC, 09/2008, Seattle, WA, USA

[13] F. Kettler, M. Lepage, M. Pawig: Evaluation of aurally-adequate analyses for echo assessment, DAGA 2009, Proceedings

[14] ITU-T Recommendation G.107: The E-model, a computational model for use in transmission planning, ITU-T, 08/2008

[15] ITU-T Recommendation P.1100: Narrowband hands-free communication in motor vehicles, ITU-T, August 2008 (pre-published)