Comparative Speech Quality Evaluation of Mobile Phones Using Advanced Testing Methods

Frank Kettler, Frank Rosenberger, Hans Wilhelm Gierlich

HEAD acoustics GmbH, 52134 Herzogenrath, Germany, Email: frank.kettler@head-acoustics.de

Motivation

The complexity of signal processing implemented in mobile phones has grown continuously over the past years. Echo cancellation, comfort noise injection, noise reduction or dynamic level controlling typically found in hands-free devices are now implemented in mobiles. Therefore the traditional testing methods are no longer sufficient in order to reflect conversational speech quality. More complex testing is necessary. Conventional measurements like frequency responses can be made more realistic by using an artificial head equipped with a more realistic artificial ear which shows the pressure force dependent acoustical leakage comparable to the human ear. This measurement also guarantees more realistic results when determining objective listening quality scores (MOS-LQO according to ITU-T-Recommendation P.800.1 [2]). Furthermore the implemented signal processing significantly influences the double talk behaviour and the quality of background noise transmission. Consequently appropriate tests and analyses are necessary.

Advanced Testing Methods

One-way transmission

Conventional types of artificial ears give a very inaccurate reproduction of the human ear. This leads for instance to good results when measuring frequency responses although the sound quality of the mobile phone would be judged worse in subjective listening tests. Consequently an artificial ear has to be used which represents the subjective listening situation as close as possible. This includes the pressure force dependent acoustical leakage between the handset and the human ear which is not covered by most of the standards used today [1]. Figure 1 shows a comparison of two frequency response measurements carried out with the same mobile but with two different types of artificial ears.

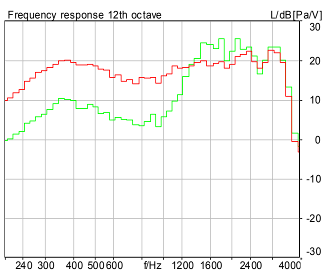

Fig. 1: Frequency responses measured with Type 1 (red) and Type 3.4 (green) artificial ear

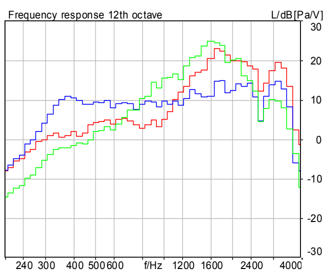

Fig. 2: Frequency responses of different mobile phones (green: mobile A, red: mobile B, blue: mobile C)

Figure 2 shows the frequency responses measured for three different mobiles. The according MOS-LQO values [2] are determined to 2.4 (A), 2.8 (B) and 3.4 (C). These values are calculated by TOSQA2001 [3],

Figure 3 gives an overview about MOS-LQO values measured for different mobiles (8 N pressure force between mobile an artificial ear). Different colours indicate different manufactures (A, B, C, D, E).

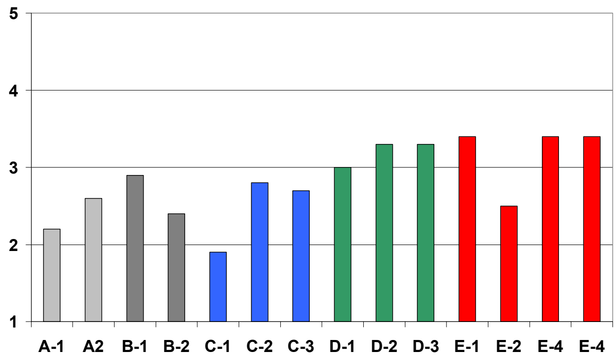

Fig. 3: MOS-LQO values for different mobiles

The results differ significantly between MOS-LQO 1.9 and 3.4. Moreover further measurements on additional mobiles also demonstrate that manufacturer specific differences are found.

Double talk performance

Besides measuring the acoustical parameters for one-way transmission it is indispensable to determine the performance during double talk and in noisy conditions due to the complex signal processing which is implemented in today's mobile phones. The size of the phones and the scenarios in which mobile phones are used cause several problems (in comparison to conventional handsets in quiet conditions):

- The coupling of the handset to the ear is impaired, at the same time the outdoor noise level is high. Therefore the user has to increase the volume of the phone.
- This leads to an amplification of the phone's echo path. This effect is even more enhanced due to the smaller size of the mobile phone which leads to a higher coupling between the phone's loudspeaker and its microphone. Echo cancellation has to be implemented; the residual echo has to be cancelled out by further echo suppression algorithms.
- If the user is speaking in this 'echo situation' the active signal processing which is affecting the sending direction of the transmission is also falsifying the speech transmission due to modulations, level variations etc. This so-called double talk situation also occurs when a background noise from the near end is transmitted. Due to limitations of the algorithms available today this

transmission of near end speech and noise underlies audible deteriorations (level variations, modulations etc.).

These problems are similar to the problems occurring with hands-free systems. The double talk tests for hands-free systems described in [4] can also be applied to mobile phones. Figure 4 and figure 5 show two examples of this double talk test determining the performance of the echo suppression. The curve shown is a level versus time analysis of the transmitted signal referred to the level versus time of the original signal applied at the near end. This "sensitivity vs. time" remains constant (straight line in figure 4) if the signal processing implemented in the phone is not affecting the speech transmission.

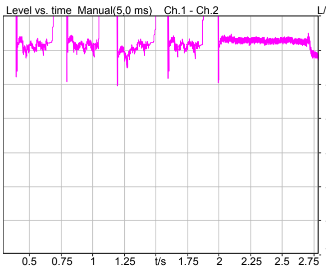

Fig. 4: Double talk in sending direction, mobile A

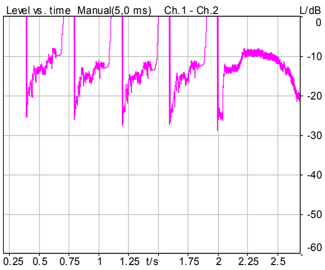

Fig. 5: Double talk in sending direction, mobile B

The 15 dB level variation introduced by mobile B (figure 5) is caused by an echo suppression inserting audible level variations in sending direction. This may lead to audible syllable clipping. Moreover these tests can be used in order to optimize performances (see figure 6 and 7).

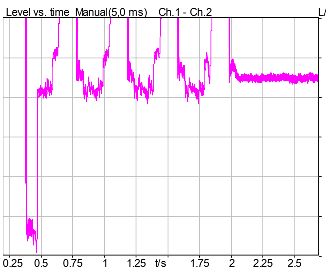

Fig. 6: Double talk in sending direction, mobile C (initial implementation)

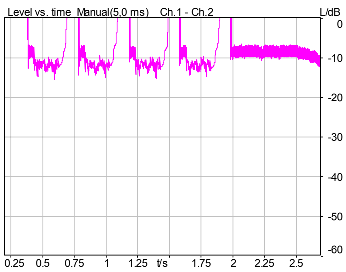

Fig. 7: Double talk in sending direction, mobile C (optimized version)

Background noise transmission

The coincident transmission of speech and noise in sending direction leads to speech quality degradations for the B subscriber. This problem is even more enhanced by the small size of the mobile phones: The microphone is positioned away from the mouth, more background noise than speech is captured by it. Therefore noise reduction algorithms have to be implemented in the phones in order to lower the high noise floor. These algorithms often deteriorate the quality of the transmitted speech.

Figure 8 shows an example of a background noise transmission test with simultaneous application of near end test signal (CSS according to [5]). The level vs. time analysis shows the modulation of the transmitted signal. Although the signal-to-noise ratio is high enough the resulting transmission quality is insufficient.

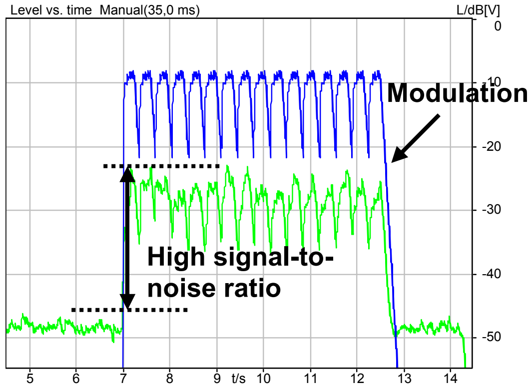

Fig. 8: Background noise transmission with near end speech, mobile D

The speech transmission measured for another mobile phone during presence of background noise is shown in figure 9. Although a noise reduction algorithm is also active in this example (see high signal-to-noise ratio) the CS signal bursts which are applied at the near end are this time transmitted without any degradation.

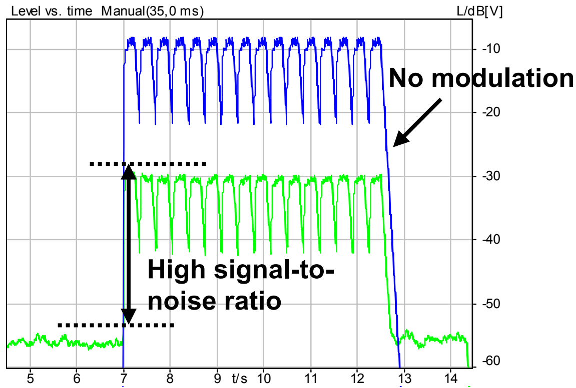

Fig. 9: Background noise transmission with near end speech, mobile E

Summary

In the past few years advanced signal processing algorithms were implemented in mobile phones. Advanced testing methods give new opportunities for evaluating and optimizing the resulting conversational quality.

References

[1] ETSI ETS 300 607-1 ed.12 (GSM 11.10-1), Digital cellular telecommunications system (Phase 2), 08/99
[2] ITU-T-Recommendation P.800.1, Mean Opinion Score (MOS) terminology
[3] ITU-T: "Results of Objective Speech Quality Assessment Including Receiving Terminals Using the Advanced TOSQA2001", 12/00
[4] ITU-T-Recommendation P.340, Transmission Characteristics of Hands-free Telephones
[5] ITU-T-Recommendation P.501, Test Signals for Use in Telephonometry