ETSI TR 102 526 V1.1.1 (2006-06)

Technical Report

Speech Processing, Transmission and Quality Aspects (STQ); Wideband telephony considerations

Reference

DTR/STQ-00057

Keywords

speech, telephony

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - NAF 742 C
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° 7803/88

Important notice

Individual copies of the present document can be downloaded from:
http://www.etsi.org

The present document may be made available in more than one electronic version or in print. In any case of existing or perceived difference in contents between such versions, the reference version is the Portable Document Format (PDF). In case of dispute, the reference shall be the printing on ETSI printers of the PDF version kept on a specific network drive within ETSI Secretariat.

Users of the present document should be aware that the document may be subject to revision or change of status. Information on the current status of this and other ETSI documents is available at
http://portal.etsi.org/tb/status/status.asp

If you find errors in the present document, please send your comment to one of the following services:
http://portal.etsi.org/chaircor/ETSI_support.asp

Contents

Intellectual Property Rights

IPRs essential or potentially essential to the present document may have been declared to ETSI. The information pertaining to these essential IPRs, if any, is publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (http://webapp.etsi.org/IPR/home.asp).

Pursuant to the ETSI IPR Policy, no investigation, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Foreword

This Technical Report (TR) has been produced by ETSI Technical Committee Speech Processing, Transmission and Quality Aspects (STQ).

1 Scope

The present document describes the state of the art of the research, tools and standards which are relevant for specifying, assessing and predicting wideband speech quality. The present document gives a summary of:

- Existing methods and specifications applicable for wideband telephony.
- The state of the art subjective testing procedures for wideband applications.
- The ongoing work relevant to define and assess the wideband terminal (and network) characteristics.
- The ongoing work on objective models for wideband speech quality assessment and prediction.

The present document furthermore gives an overview about the work needed to create a wideband transmission rating model.

Independent of speech coder used.

The present document focuses on wideband telephony (100 Hz to 8 kHz) but is not limited to this frequency range.

2 References

For the purposes of this Technical Report (TR), the following references apply:

[1] Barriac V.; Le Saout, J.-Y.; Lockwood, C.: "Discussion on unified methodologies for the comparison of voice quality of narrowband and wideband scenarios"; ETSI-Workshop on Wideband Speech Quality in Terminals and Networks: Assessment and Prediction; Mainz, June 8-9 2004.

[2] Drascher, T.: "A Subjective Sound Quality Assessment of Mobile Phones for Production Support"; ETSI-Workshop on Wideband Speech Quality in Terminals and Networks: Assessment and Prediction; Mainz, June 8-9 2004.

[3] Gierlich, H.W.; Völl, S.; Jax, P.; Kettler, F.: "Speech quality assessment for wideband communication scenarios"; ETSI-Workshop on Wideband Speech Quality in Terminals and Networks: Assessment and Prediction; Mainz, June 8-9 2004.

[4] Gros, L.; Monfort, J.-Y.; Quinquis C.: "Comparison of Narrow band and Wideband Speech Codecs in noisy environment"; ETSI-Workshop on Wideband Speech Quality in Terminals and Networks: Assessment and Prediction; Mainz, June 8-9 2004.

[5] Halkosaari, T.; Vaalgamaa, M.: "Radiation Directivity of Human and Artificial Speech"; ETSI-Workshop on Wideband Speech Quality in Terminals and Networks: Assessment and Prediction; Mainz, June 8-9 2004.

[6] Kitawaki, N.: "Perspectives on Multimedia Quality Prediction Methodologies for Advanced Mobile and IP-based Telephony"; ETSI-Workshop on Wideband Speech Quality in Terminals and Networks: Assessment and Prediction; Mainz, June 8.-9. 2004.

[7] Mahieux, Y.; Derval, G.; Deléam.D.: "Wide Band Speech introduction into VOIP solutions"; ETSI-Workshop on Wideband Speech Quality in Terminals and Networks: Assessment and Prediction; Mainz, June 8-9 2004.

[8] Möller, S.; Raake, A.; Barriac, V.; Quinquis, C.: "Deriving Equipment Impairment Factors for Wideband Speech Codecs"; ETSI-Workshop on Wideband Speech Quality in Terminals and Networks: Assessment and Prediction; Mainz, June 8-9 2004.

[9] Raake, A.: "How much better can wideband telephony be? - Estimating the necessary R-scale extension"; ETSI-Workshop on Wideband Speech Quality in Terminals and Networks: Assessment and Prediction; Mainz, June 8-9 2004.

[10] Rix, A.: "Subjective wideband speech quality and modelling issues"; ETSI-Workshop on Wideband Speech Quality in Terminals and Networks: Assessment and Prediction; Mainz, June 8-9 2004.

[11] Rix, A.: "Perceptual wideband speech and audio quality measurement"; ETSI-Workshop on Wideband Speech Quality in Terminals and Networks: Assessment and Prediction; Mainz, June 8-9 2004.

[12] Sydow, C.: "Practical Limitations of Wideband Terminals"; ETSI-Workshop on Wideband Speech Quality in Terminals and Networks: Assessment and Prediction; Mainz, June 8-9 2004.

[13] Ulseth, T: "A path towards common quality assessment of narrowband and wideband voice"; ETSI-Workshop on Wideband Speech Quality in Terminals and Networks: Assessment and Prediction; Mainz, June 8-9 2004.

[14] ITU-T Recommendation H.245 : "Control protocol for multimedia communication".

[15] Möller, S., Raake, A., Kitawaki, N., Takahashi, A., Wältermann, M. (2006): "Impairment Factor Framework for Wideband Speech Codecs", submitted to IEEE Trans. Audio, Speech and Language Processing.

[16] Völl, S.; Gierlich, H.W.: "High Quality Background Noise Simulation: the ETSI STF 273 Project"; 2nd ETSI-Workshop on Wideband Speech Quality in Terminals and Networks: Assessment and Prediction; Mainz, June 22-23 2005.

[17] Monfort, J.-Y.: "New STF 294: Improving the quality of eEurope wideband speech applications by developing a standardised performance testing and evaluation methodology for background noise transmission"; 2nd ETSI-Workshop on Wideband Speech Quality in Terminals and Networks: Assessment and Prediction; Mainz, June 22-23 2005.

[18] Beaugeant, C.; Varga, I.; Jax, P.: "Noise Reduction Preprocessing for the Adaptive Multi-Rate Wideband (AMR-WB) Speech Codec"; 2nd ETSI-Workshop on Wideband Speech Quality in Terminals and Networks: Assessment and Prediction; Mainz, June 22-23 2005.

[19] Varga, I.: "Standardization of the Adaptive Multi-Rate Wideband (AMR-WB) Speech Codec and its Applications"; 2nd ETSI-Workshop on Wideband Speech Quality in Terminals and Networks: Assessment and Prediction; Mainz, June 22-23 2005.

[20] Quinquis, C.: "A new codec within ITU-T: new bandwidth"; 2nd ETSI-Workshop on Wideband Speech Quality in Terminals and Networks: Assessment and Prediction; Mainz, June 22-23 2005.

[21] Jax, P.; Vary, P.: "On the use of artificial bandwith Extension Techniques in Wideband Speech Communications"; 2nd ETSI-Workshop on Wideband Speech Quality in Terminals and Networks: Assessment and Prediction; Mainz, June 22-23 2005.

[22] Pankaj K. R.; Ajit V. Rao: "A Scalable Wideband Speech Codec"; 2nd ETSI-Workshop on Wideband Speech Quality in Terminals and Networks: Assessment and Prediction; Mainz, June 22-23 2005.

[23] Iser, B.: "Bandwidth extension of telephone band- limited Speech Signals"; 2nd ETSI-Workshop on Wideband Speech Quality in Terminals and Networks: Assessment and Prediction; Mainz, June 22-23 2005.

[24] Meier, F.: "A wideband capable VoIP Chipset called INCA-IP"; 2nd ETSI-Workshop on Wideband Speech Quality in Terminals and Networks: Assessment and Prediction; Mainz, June 22-23 2005.

[25] Diethorn, E.: "Aspects of Wideband Speech in Enterprise Telephony"; 2nd ETSI-Workshop on Wideband Speech Quality in Terminals and Networks: Assessment and Prediction; Mainz, June 22-23 2005.

[26] Wältermann, M.; Möller, S.; Raake, A.: "Quality Dimensions of Narrow-Band and Wideband Telephone Connections"; 2nd ETSI-Workshop on Wideband Speech Quality in Terminals and Networks: Assessment and Prediction; Mainz, June 22-23 2005.

[27] Schmidtmer, C.: "Perceptual Wideband Audio Quality Assessments Using PEAQ"; 2nd ETSI-Workshop on Wideband Speech Quality in Terminals and Networks: Assessment and Prediction; Mainz, June 22-23 2005.

[28] Berger, J.: "Speech Quality in Cellular Networks - More or Less than GSM EFR"; 2nd ETSI-Workshop on Wideband Speech Quality in Terminals and Networks: Assessment and Prediction; Mainz, June 22-23 2005.

[29] Völl, S.; Gierlich, H.W.; Kettler, F.; Jax, P.: "Background Noise Transmission Quality for Wideband Systems"; 2nd ETSI-Workshop on Wideband Speech Quality in Terminals and Networks: Assessment and Prediction; Mainz, June 22-23 2005.

[30] ETSI EG 202 396-1: "Speech Processing, Transmission and Quality Aspects (STQ); Speech quality performance in the presence of background noise; Part 1: Background noise simulation technique and background noise database".

[31] ETSI EG 202 396-2: "Speech Processing, Transmission and Quality Aspects (STQ); Speech Quality performance in the presence of background noise Part 2: Background noise transmission - Network simulation".

[32] ITU-T Recommendation E.800: "Terms and definitions related to quality of service and network performance including dependability".

[33] ITU-T Recommendation P.581: "Use of head and torso simulator (HATS) for hands-free terminal testing".

[34] ITU-T Recommendation P.800: "Methods for subjective determination of transmission quality".

[35] ITU-T Recommendation P.831: "Subjective performance evaluation of network echo cancellers".

[36] ITU-T Recommendation P.832: "Subjective performance evaluation of hands-free terminals".

[37] ITU-T Recommendation P.833: "Methodology for derivation of equipment impairment factors from subjective listening-only tests".

[38] ITU-T Recommendation P.834: "Methodology for the derivation of equipment impairment factors from instrumental models".

[39] ITU-T Recommendation P.862: "Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs".

[40] ITU-T Recommendation P.862.2: "Wideband extension to Recommendation P.862 for the assessment of wideband telephone networks and speech codecs".

[41] ITU-T Recommendation P.50: "Artificial Voices".

[42] ISO/DIS 10845: "Acoustics - Frequency weighting "A" for noise measurement".

[43] ETSI ETR 250: "Transmission and Multiplexing (TM); Speech communication quality from mouth to ear for 3,1 kHz handset telephony across networks".

[44] ITU-R Recommendation BS.1116: "Methods for the subjective assessment of small impairments in audio systems including multichannel sound systems".

[45] ITU-R Recommendation BS.1534-1: "Method for the subjective assessment of intermediate quality levels of coding systems".

[46] ITU-T Recommendation G.722: "7 kHz audio-coding within 64 kbit/s".

[47] ITU-T Recommendation G.711: "Pulse code modulation (PCM) of voice frequencies".

[48] ITU-T Recommendation G.722.1: "Low-complexity coding at 24 and 32 kbit/s for hands-free operation in systems with low frame loss".

3 Abbreviations

For the purposes of the present document, the following abbreviations apply:

ACR	Absolute Category Rating
AMR	Adaptive Multi Rate
Ie	equipment Impairment
Ie,wb	wideband Ie
MOS	Mean Opinion Score
NB	Narrowband
NP	Network Performance
PCM	Pulse-Code Modulation
PESQ	Perceptual Evaluation of Speech Quality
QoS	Quality of Service
WB	Wideband

4 Overview about work in different areas

4.1 Subjective speech quality assessment

Generally the subjective testing procedures which are found in the relevant ITU-T Recommendations (P.800 [34], P.831 [35], P.832 [36], etc.) can be applied. The most important questions to be answered are:

- Is the overall speech quality improved when using wideband transmission systems?
- How is the perception of speech sound quality influenced by wideband systems?
- Can speech intelligibility and the interactivity be improved by wideband transmission systems?
- How is the wideband system performance in the presence of noise?
- What performance requirements have to be set for other the parameters known influencing the speech quality (delay, loudness, echo performance, etc.)?
- Are there new quality parameters influencing the speech quality for wideband systems?

Ongoing work for wideband concentrates on the relationship between the single talk speech quality for narrowband systems vs. wideband systems.

Figure 1 gives an overview about the different parameters which might be perceived different in wideband systems.

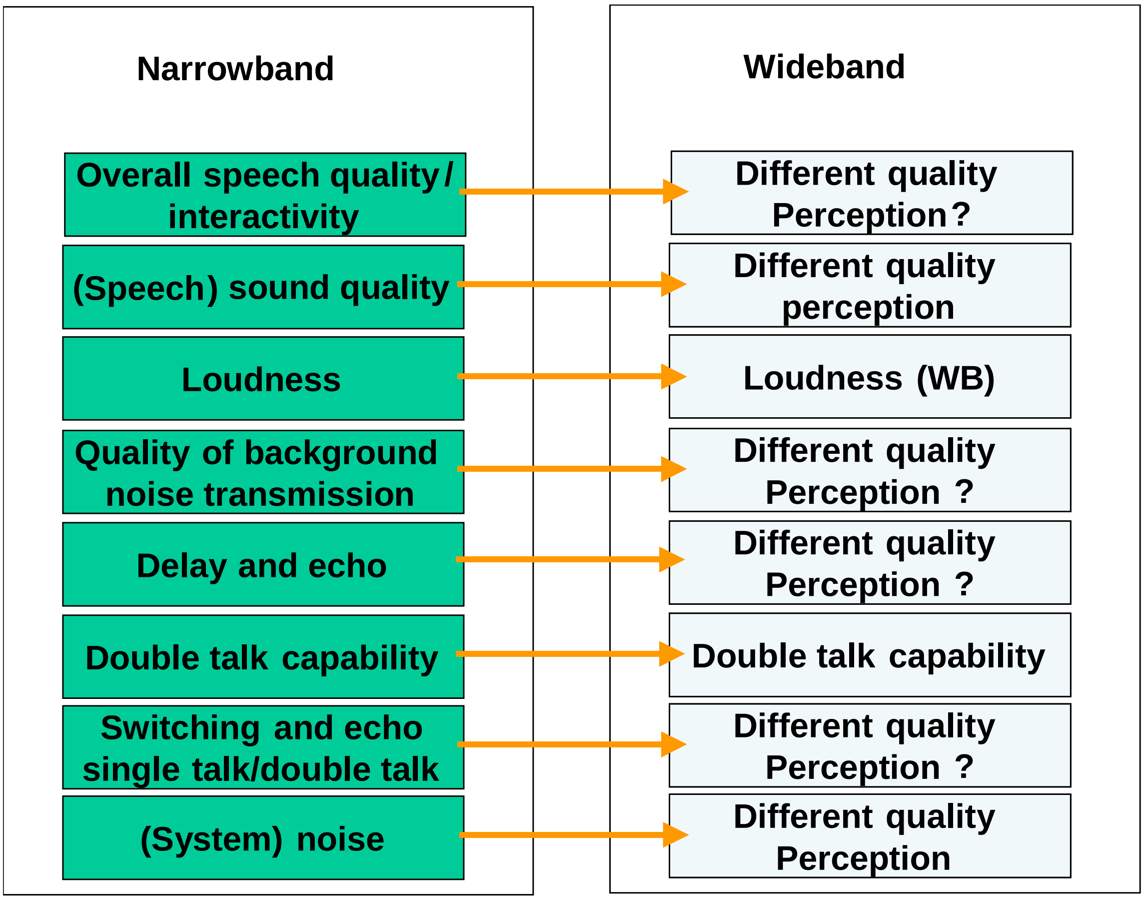


Figure 1: Wideband <-> narrowband: different performance parameters and their perception

In [10] a general discussion of subjective testing for wideband systems is found. P.800 [34] tests are compared to subjective tests described in ITU-R Recommendation BS.1116 [44] which originally was designed for high quality transmission system and ITU-R Recommendation BS.1534-1 [45] type of tests. ITU-R Recommendation BS.1534-1 [45] type of tests are designed for multimedia/mp3 or consumer headphones/loudspeaker tests, for systems providing acceptable quality with audible impairments. ITU-R Recommendation BS.1116 [44] tests are very sensitive and focussing on small impairments. [11] features 'P.800 wideband tests' for speech sound quality. In difference to the standard telephony tests high quality wideband equalized headphones are used (see also ITU-T Recommendation P.832 [36], [9] and [3]), the speech material is presented binaurally. The equalization of the headphones should be freefield or diffusefield (see [33]). This arrangement seems to be the most favourite for third party listening tests within wideband systems.

4.1.1 Conversational Tests: Comparison of narrowband and wideband speech codecs in noisy environment

Within the 3GPP characterization work for default codecs for packet switched conversational communications a variety of codecs were tested with (car noise, 60 dBPa) and without background noise [4]. Furthermore the experiments were conducted introducing a packet loss of 3 %. The following codecs were compared:

- AMR (6,7 and 12,2 modes)
- AMRWB (12,65 and 15,85 modes)
- G.723.1@ 6,3 kb/s
- G.729@ 8 kb/s
- G.711@ 64 kb/s
- G.722 @ 64 kb/s

Within theses experiments conversational tests were carried out, the following parameters were asked for:

- **Voice quality**

The experiments conclude that "the superiority of WB Codecs is higher for noisy and packet-loss conditions, but there is no systematic benefit of WB in noise" [4].

- **Intelligibility**

  In the experiments [4] intelligibility was measured by using an ACR type test asking for the "Quality of understanding the partner". The experiments conclude [4]: Tuckey tests show an almost systematic superiority of WB and G.711 [47] codecs on NB codecs (except AMR NB 12,2 kbs) in noisy and/or lossy conditions. Without noise and packet loss no significant difference can be observed.

- **Interactivity**

  Interactivity was measured using an ACR type test. Generally the interactivity decreases with car noise and packet loss, in packet loss conditions the interactivity is mostly determined by the packet loss the influence of the noise is low. Generally wideband codecs resist better to noise if no packet loss is present.

- **Default perception**

The experiments conclude [4]: 'Tuckey tests show no significant differences between codecs for any condition. The default perception criterion does not seem to be relevant to detect a possible advantage of WB codecs in noise.'

- **Global quality of the communication**

The experiments in [4] conclude for the global quality of the communication that "Tuckey tests show an almost systematic superiority of WB and G.711 [47] codecs on NB codecs (except AMR NB 12,2 kbs) for noisy and/or lossy conditions." Without noise and packet losses no significant differences can be observed. The superiority of WB Codecs is higher for noisy and packet-loss conditions, but there is no systematic benefit of WB in noise.

The study concludes [4] that in general noise was the strongest experimental factor that affected the subjective data. Without packet loss, in noise, the advantage of WB compared to NB is that it enhances the comprehension and the interaction with the other partner, but not voice quality and global quality, since these criteria are highly affected by the presence of noise.

In [3] conversational tests were conducted for parameter identification. From the test subjects interviews it is concluded that the parameters influencing mostly the overall speech quality are:

- Sound of speech.
- Echo: level, masking, intelligibility.
- Quality of background noise transmission.
- Noise.
- Double talk.
- Switching/clipping.

4.1.2 Third party listening tests

- **Intelligibility**

In [3] the results of logatom tests indicate that speech intelligibility can be improved by wideband systems. In the test setup third party listening tests were used. Handsfree systems were used in the tests, the recordings were made by artificial heads (see ITU-T Recommendations P.831 [35], P.581 [33], etc.). For playback equalized high quality headphones were used. The result of the experiments is shown in figure 2.

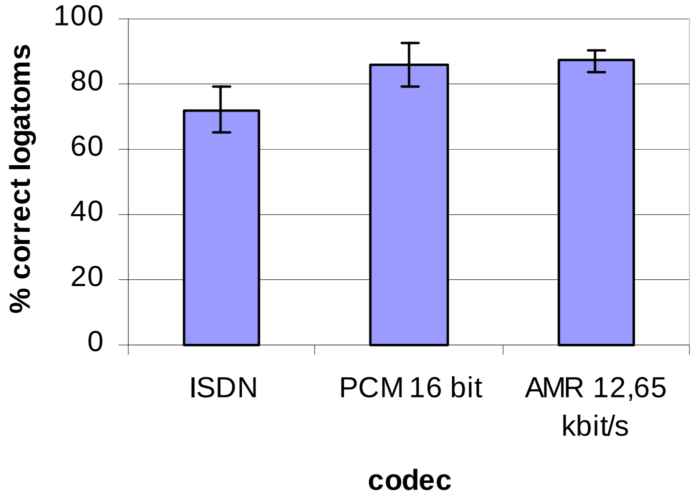


Figure 2: Wideband <-> narrowband: Speech intelligibility with logatom tests from [3]

- **Speech sound quality**

In the subjective experiments reported in [1] two types of P.800 based subjective tests were made, both using the ACR scale. One test includes only narrowband conditions, the other narrow-band and wideband conditions. In both tests headphones were used for playback. In the narrowband tests a clear channel narrowband reference was included, in the wideband tests a clear channel wideband reference was included in addition. The test results indicate that:

- the narrowband MOS values decrease in the mixed narrowband/wideband presentation;
- the relationship between the narrowband conditions used in the narrowband tests as compared to the mixed tests seems to be linear (see figure 3);
- no change could be observed for the wideband MOS scores achieved in a wideband only test as compared to a mixed narrowband/wideband test.

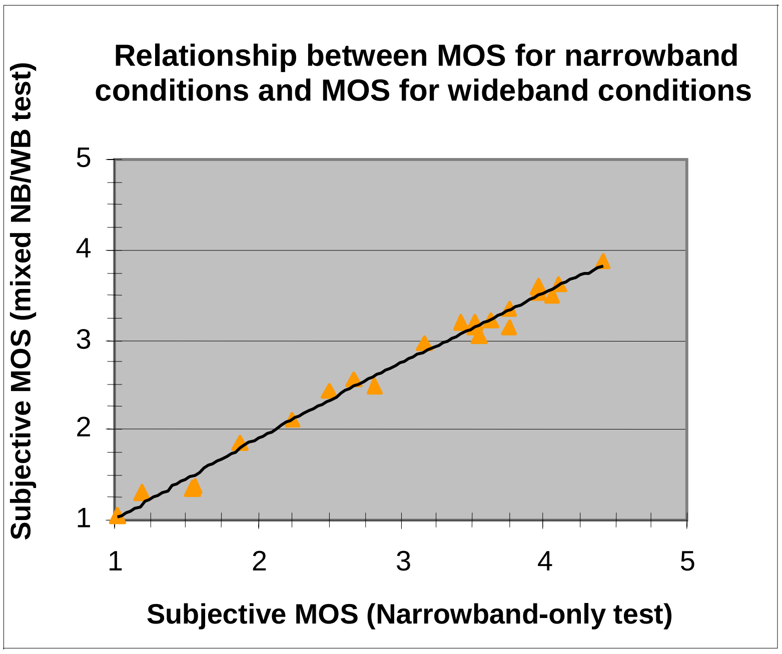


Figure 3: Relationship between narrowband MOS and wideband MOS from [1]

The subjective experiments conducted by [9] indicate that different result between wideband and narrowband tests can be expected depending on the type of acoustic interface used:

- Between a narrowband handset and monaural headset an MOS difference of 0.8 is reported, when a diotic headset is used the difference is about MOS 0,4.
- When increasing the bandwidth of the sound presentation the MOS difference between narrowband and wideband presentation is about 0,5 for undistorted signals which seems to be similar to the experimental results described above by [1].
- Other experiments of [9] indicate the "quality-difference" between narrowband and wideband may also be in the range of delta MOS 1,0 to 1,3, depending on the way of presenting the signals.

New experiments are reported in [26]. Instead of asking only for listening speech quality in the listening situation the authors describe a multidimensional approach. The experiments show that the following impairing factors in the listening situation can be found: interruptiveness, directness/clearness, frequency content and noisiness. The most important of the dimensions are frequency independent and it is pointed out that in most cases speech quality factors should be independent from narrowband or wideband transmission. Nevertheless, wideband speech could be different as in regard to the weighting of the factors.

- **Performance in the presence of background noise**

Background noise tests with music as the background noise signals are introduced in [3]. 16 different types of codecs were used in the experiments. The subjects were asked about their perception of the background noise transmission quality. The experiments conclude that generally 3 classes of performance seem to exist:

- Wideband codecs with high quality speech coding (PCM, MP3, G.722 [46]) are rated best (MOS 3,5 to 4).
- Narrowband codecs with low compression (G.711 [47], etc.) are rated reasonable (MOS 2 to 2,3).
- Codecs with high compression (wideband and narrowband) and bandwidth extension are rated unacceptable (MOS 1).

- **Echo annoyance tests**

Different experiments for echo annoyance are presented in [3]. The conclusions are:

- The AMR WB coder seems to be more sensitive to echo than others.
- High frequency echo in wideband systems contributes more to the echo performance than low frequency dominated echo.

4.2 Wideband codecs and mixed narrowband/wideband scenarios

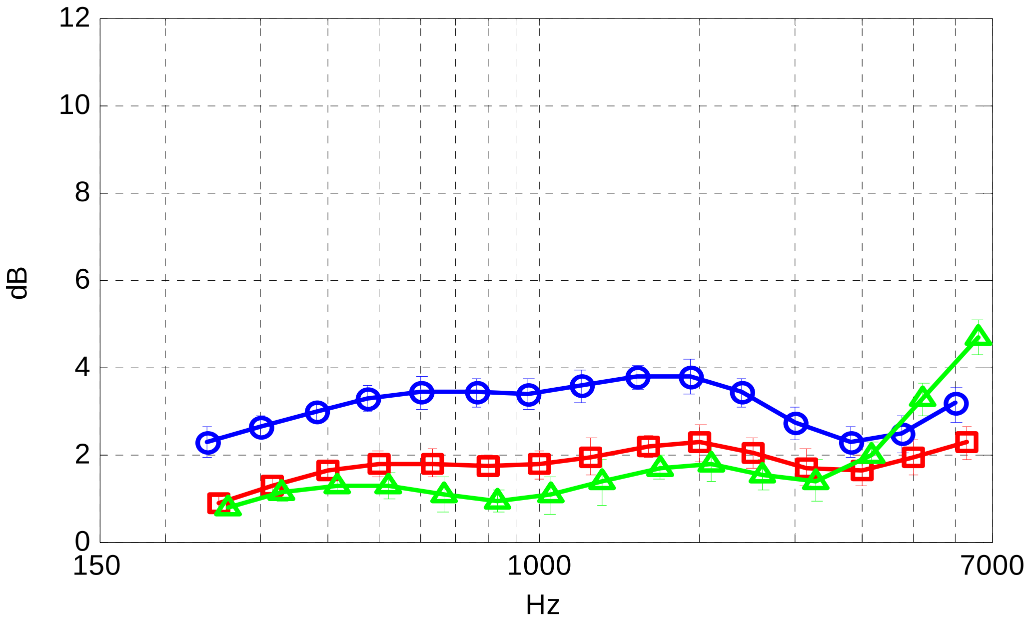
A specification for a linear 16-bit PCM codec for wideband implementations can be found in [14]. An overview about the standardization process of the AMR-WB is given in [19]. A review of wideband codecs is available in EG 202 396-2 [31]. Recent work on a new ITU-T extended wideband codec (higher bandwidth than 8 kHz, an extension of ITU-T Recommendation G.722.1 [48]) is described by [20].

In general the transition from wideband to narrowband needs to be considered in the codecs as well as in the terminal design. Methods for wideband extension of narrowband signals maybe very useful especially in mixed narrowband-wideband environments. Examples for possible solutions are described in [21] and [23]. The implementation of wideband codecs in combination with narrowband codecs is introduced in [24] and [22].

4.3 Objective speech quality assessment

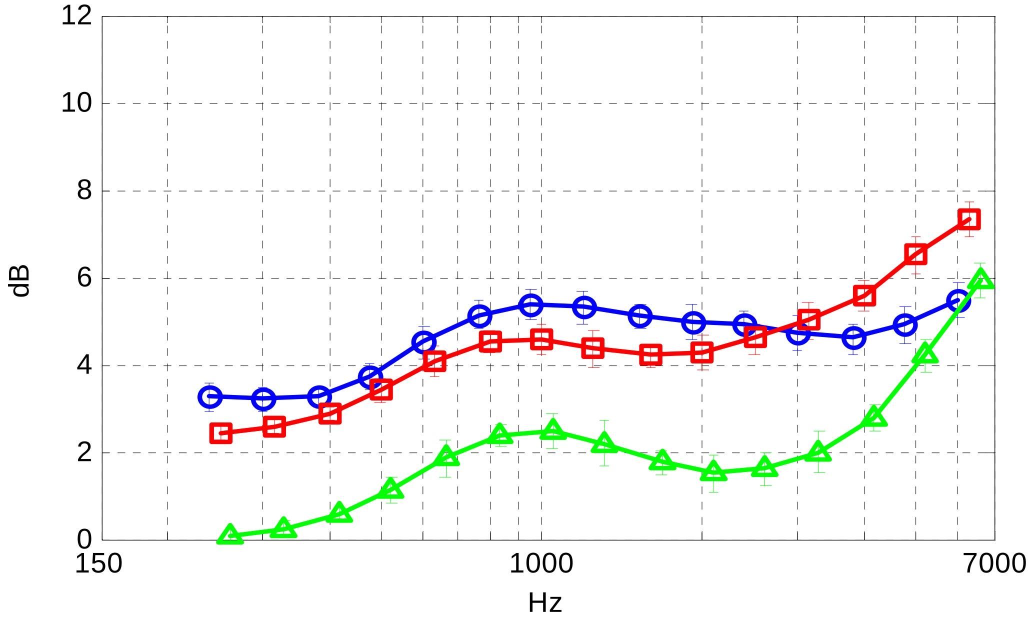
4.3.1 Radiation directivity of the artificial mouth

Intensive studies of the human mouth directivity especially for wideband were carried out in [5]. The study found differences in the frequency dependant directivity characteristics between humans and artificial heads, the measured differences are also dependant on the brand of the artificial head. 8 male and 5 female were used to access the human data, different test positions (microphone positions) were recorded simultaneously. The different speech samples, phonemes and active speech were segmented manually. For averaging and smoothing 1/3rd actave bands were used. Figures 4 and 5 (see [5]) show the difference as measured between two different artificial heads and the averaged results of humans.



- Large handset (blue circles).
- Small handset (red squares).
- Boom headset (green triangles).

Figure 4: Smoothed difference curves and 95% confidence limits between humans - Head Acoustics HMS II.3 (right) on near cheek positions



- Large handset (blue circles).
- Small handset (red squares).
- Boom headset (green triangles).

Figure 5: Smoothed difference curves and 95% confidence limits between human - B&K 4227 (left) on near cheek positions

The authors propose to compensate the differences at known positions by correcting the measured frequency response by a filter taking into account the average differences for those artificial heads where the observed difference is too high. It should be noted that former investigations done in ITU-T do not show the inconsistencies as indicated by [5].However it is not easy to compare these results against the ITU data as a different set of points have been tested. The only coincident points are the MRP, used as a reference, and point 3/P.58 (point 3.2 for [5]). However the results at point 3.2 have not been published. The work carried out could represent a basis for a possible extension of current ITU Recommendations and as such should be submitted to SG12. However, it shall be noted that both P.51 and P.58 do NOT specify the technical construction details of the devices themselves (e.g. the mouth opening section) but their acoustical performances (e.g. the sound generation patterns, distortion, linearity, etc.). The dimensional data provided in P.58 are only intended as general templates to be complied with by the overall dimensions of HATS, they are not provided as specific implementation guidances for the manufacturers.

It can be assumed that this topic needs to be investigated in more detail.

4.3.2 Limitations for wideband introduced by the terminal

The quality of wideband systems is mainly determined by the terminal and its performance in wideband operation. The acoustical components used in a terminal highly contribute to the quality perception of wideband systems. [12] highlights the constraints under which terminals and their components are to be constructed. Even for desktop hands-free the acoustic transducer performance is far away from the requirements one needs to put for wideband systems. As an example the frequency response of a 50 mm speaker is given (see figure 6).

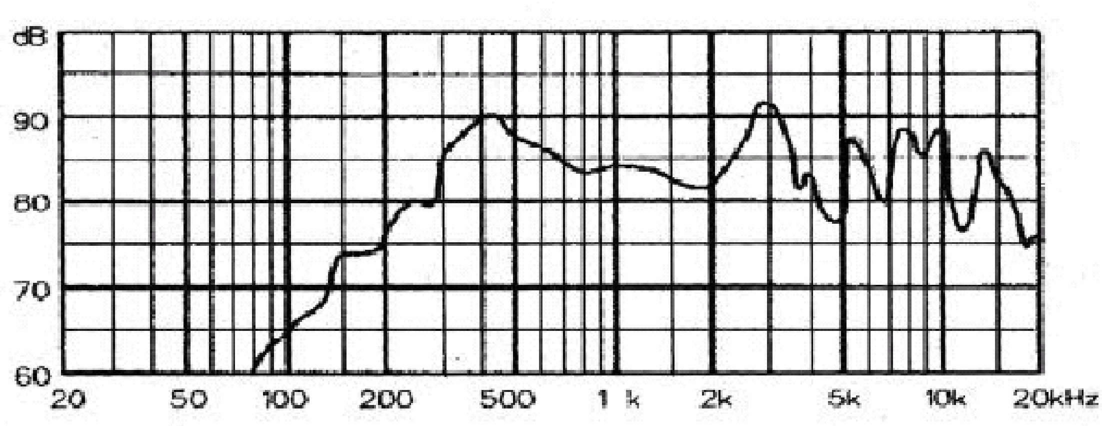


Figure 6: Frequency response of a typical 50mm speaker used for hands-free [12]

It is obvious that a lower frequency cutoff frequency of 200 Hz or even 100 Hz is mostly impossible to achieve with these types of speakers. For any type of mobile hands-free the situation is even worse.

[12] also highlights the different requirements on frequency response found in the different standards.

Figures 7 and 8 show the various requirements. Certainly a unification of the different requirements is needed.

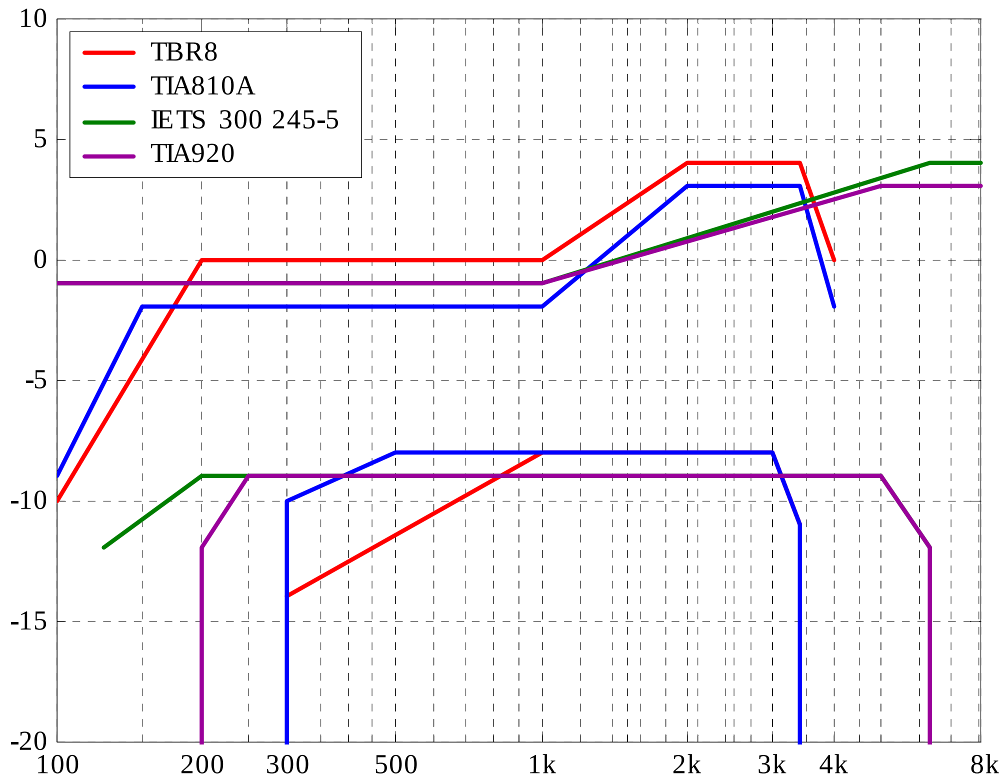


Figure 7: Different frequency response tolerance schemes in sending from the different standards [12], comparison narrowband <-> wideband

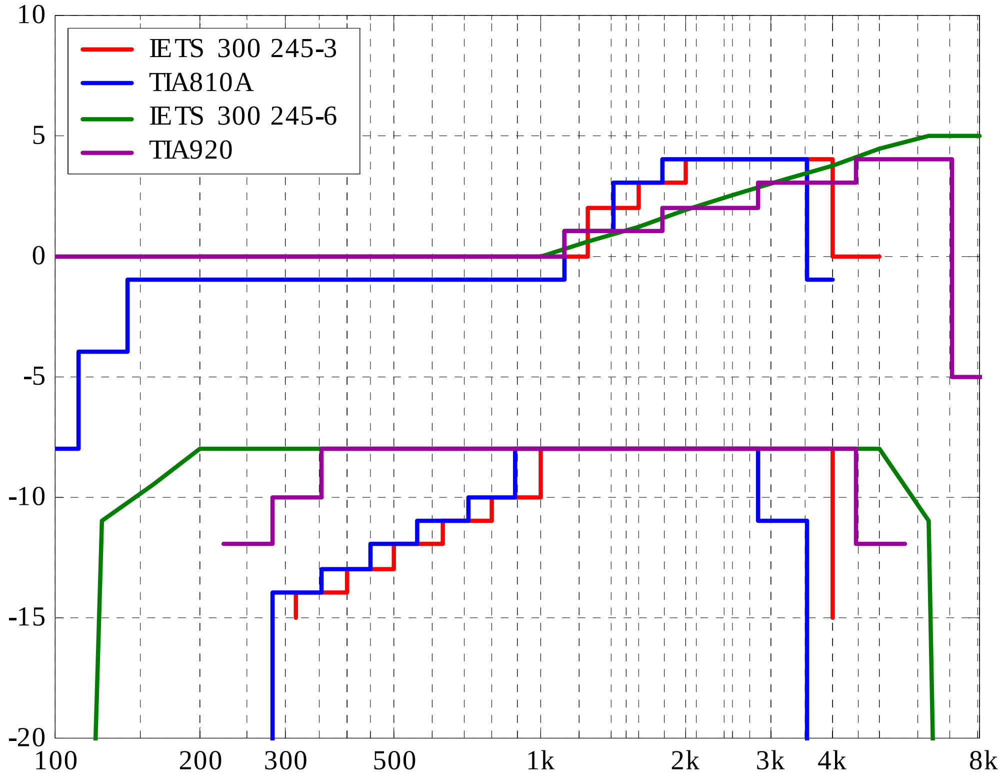


Figure 8: Different frequency response tolerance schemes in receiving from the different standards [12], comparison narrowband <-> wideband

Even more than in the narrowband case the terminal performance in the presence of background noise is of critical importance. An efficient method to simulate various background noise conditions is described in [16] and in EG 202 396-1 [30]. A database containing speech samples in various background noise situations processed with various noise cancellation techniques and transmitted over different network configurations is described in EG 202 396-2 [31]. Techniques for noise reduction preprocessing in wideband systems are described e.g. in [18].

4.4 Quality prediction and modelling

A general discussion about objective speech quality prediction in wideband systems and the available models is given in [11] and [27]. It is shown that BS. 1387 is not applicable for wideband speech, this is also true for the current ITU-T Recommendation P.862 [39].

4.4.1 Extension of objective speech quality measures (P.862) to wideband

[1] discusses possibilities for adaptation of the existing objective speech quality measure according to ITU-T Recommendation P.862 [39] for wideband. Basically two modifications were made: Adaptation of the input filter to wideband and use of a different mapping function. The same modification is described in [11].

These modifications led to a good match between the MOS scores and the P.862 [39] prediction for the speech material used in the subjective tests described in [1]. However many new experiments with new (and unknown) conditions are needed to validate the modification. [6] shows the correlation of Wideband P.862 [39] to subjective test data however not much detail on the experiments is given. [11] shows a good correlation of the modified P.862 [39] to some subjective tests, however a new mapping function to subjective tests results for WB P.862 [39] as well as more extended subjective testing is required.

ITU-T adopted the modifications of P.862 for wideband applications. The method is described in ITU-T Recommendation P.862.2 [40].

[6] introduces furthermore the possibility of using artificial voice according to ITU-T Recommendation P.50 [41] as a test signal for P.862 [39] (PESQ). The experiment focuses on different S/N conditions. It can be seen, that for S/N conditions > 5 dB the P.862 [39] results between real speech and the artificial voice signal are similar, for lower S/N conditions however the results differ significantly.

4.4.2 Extension of the E-model

Different approaches can be made in order to extend the existing E-model, a parametric quality prediction model to be used for network planning, towards wideband applications.

[13] proposes an intermediate E-model where the advantage of the wideband system is modelled by a wideband advantage factor WBa:

$$R = R_o - I_s - I_d - I_e + A + WB_a$$

This wideband advantage factor should be based on subjective tests. Furthermore, the equipment factors for wideband should be introduced. This might be possible based on the procedures described in ITU-T Recommendations P.833 [37] and P.834 [38]. It is recognized that the effect of delay on speech interactivity is probably not different for wideband speech compared with narrowband speech. However, the echo effects might be a challenge, particularly because the echo characteristics probably are different for a wideband terminal.

Although this approach is rather pragmatic, subjective validation is needed to get the right range of WBa and validated new Ie values.

A more basic detailed analysis of the requirements for a wideband E-model is given in [9].

Table 1: Overview about E-model parameters and their applicability for wideband [9]

Parameter	WB coverage	Reference	Remarks
SLR / RLR / OLR	✓ / ✓ / ✓	ITU-T Recommendation P.79, Annex G	–
Nc / Nfor	(✓?) / (✓?)	ITU-T Recommendation P.83, Annex A, O.41	Problem: Psophometric weighting O.41 → NB
Ps / Pr	✓ / ✓	ISO/DIS 10845 [42]	–
STMR / LSTR	(✓?) / (✓?)	ITU-T Recommendations P.79 ; P.64	–
qdu / Q	– / ✓	– / ITU-T Recommendation P.810	–
TELR / WEPL	– / –	– / –	Problem: Recommendation like ITU-T Recommendation G.122 → NB
Ta / T / Tr	✓ / ✓ / ✓	ITU-T Recommendation G.114 / ETR 250 [43]	–
Ie	✓	Various approaches	–
Bpl	–	–	So far too few tests
Ppl	✓	–	Bandwidth-independent

Similar to [13], also [9] discusses the introduction of an additional a factor Cwb leading to an R-scale extension. [9] suggests to use the following procedure:

- Conduct reference experiments including wideband and narrowband codecs as well as different types of degradation (see also results from [1]), include different bandwidth limitations.
- Derive a new R(wideband) value form the wideband experiments.
- Define a new R0(wideband)-value by plotting R(narrowband) vs R(wideband) as shown in figure 9.

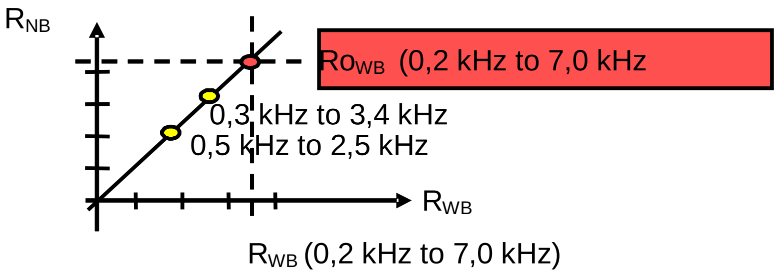


Figure 9: Proposal for derivation of a new wideband Ro(wideband) [9]

Depending on the underlying experiments, the Ro(WB) value was found to be in the range of 113 to 154. However there is still al lack of consistent experimental data, because most tests have not been carried out for the purpose opf an extension of the E-model. Furthermore the linear extrapolation as indicated in figure 9 may not be ideal; other data suggest a curvilinear relationship. As a conservative estimate however an Ro(WB) of 113 could be used as a starting point.

This proposal is further extended by [15], using the most conservative estimate, i.e. a 10 % extension of the R-scale. In [15], equipment impairment factors Ie,wb for wideband speech codecs are calculated, based on a procedure which is similar to the one used for deriving Ie values for narrowband codecs. The necessary steps are as follows:

- Transformation of MOS-values to the R-scale, using the standard E-model formula.
- Expansion of the R-scale range to 110.
- Calculation of the Ie-values as the difference to the Ie value for the G.711 [47] codec.
- Definition of a new Ie,wb as the difference to the direct (wideband) case (R := 110).

Results from France Telecom (tests for 3GPP) are used to derive first Ie,wb values according to this procedure, for the G.722 [46] and two AMNR-WB codecs. The values are further stabilized using test results from the AMR-WB characterization test phase. This leads to a provisional list of Ie,wb values, each of which is stabilized by the number of tests given in parentheses:

G.722 (64):	***Ie,wb*** := 14	(5 Tests)
G.722 (56):	***Ie,wb*** := 18	(2 Tests)
G.722 (48):	***Ie,wb*** := 24	(4 Tests)
G.722.1 (24):	***Ie,wb*** := 12	(3 Tests)
AMR-WB (6,6):	***Ie,wb*** := 29	(4 Tests)
AMR-WB (8,85):	***Ie,wb*** := 18	(4 Tests)
AMR-WB (12,85):	***Ie,wb*** := 10	(5 Tests)
AMR-WB (14,25):	***Ie,wb*** := 7	(4 Tests)
AMR-WB (15,85):	***Ie,wb*** := 6	(5 Tests)
AMR-WB (18,25):	***Ie,wb*** := 5	(4 Tests)
AMR-WB (19,85):	***Ie,wb*** := 2	(4 Tests)
AMR-WB (23,05):	***Ie,wb*** := 2	(3 Tests)
AMR-WB (23,85):	***Ie,wb*** := 5	(3 Tests)

Using this approach, a fixed relationship between Ie values for narrowband codecs and Ie,wb values for wideband codecs is implied:

$$Ie,wb = Ie + 15$$

Finally, an additivity check was made leading to some questions about the additivity property of the E-model: The check indicates that assuming a sum of Ie,wb impairment factors in case of codec tandems mostly leads to a too pessimistic estimation of the perceived quality, but not in all cases. In addition, the tandem order of the coders seems to be important. However, the results are provisionary only, and will have to be complemented by additional – independent test data. Such data is provided in a recent document summarizing results from different sources [15].

[8] discusses the question whether a proportional extension of the narrowband R-scale is at all allowed, and whether a direct transformation of the MOS values collected in a wideband context to R values on the extended R-scale would be possible. It is acknowledged that other scaling methods may be more appropriate for scaling the quality of wideband speech transmission. An answer to the raised question is still open. Furthermore, it should be noted that all proposals and discussions in [8] refer to impairments introduced by codecs alone; no information is yet available on the influence of other (wideband) impairments, and on the applicability of the additivity theorem fore these impairments in wideband systems.

4.4.3 The definition of QoS in conjunction with wideband

4.4.3.1 General

For a common understanding of QoS and Network Performance (NP) the main principles are reflected in the present document as follows:

- ITU-T Recommendation E.800 [32] definition of **Quality of Service**:

  "The collective effect of service performance which determine the degree of satisfaction of a user of the service".

- ITU-T Recommendation E.800 [32] definition of **Network Performance**:

  "The ability of a network portion to provide the functions related to communication between users".

The relationship between customer satisfaction, QoS and NP is shown in figure 10. The present document has its focus on the technical aspects related to customer satisfaction.

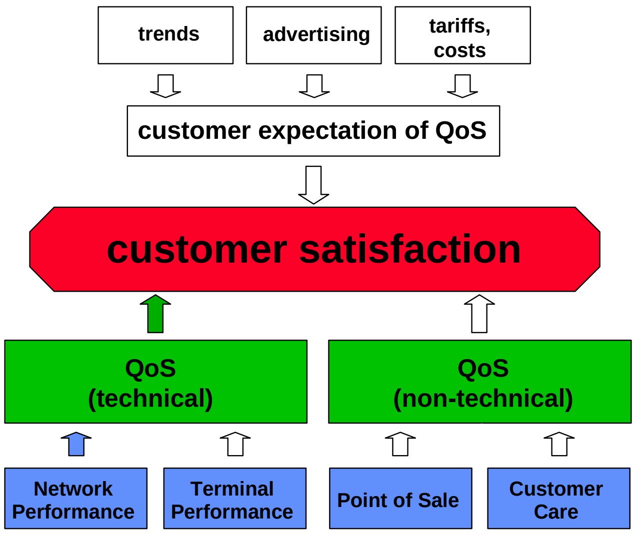


Figure 10: Relationship between customer satisfaction, QoS and NP

4.4.3.2 Parameters in relation to QoS

There is the need to specify independent QoS indicators for each service part. The selected indicators are considered:

- to have main influence on the customers satisfaction with regard to the service;
- to identify technical QoS aspects, which can be influenced by the performance of the network or the terminal;
- to be measurable by technical means;
- to be relevant for network operator's national and international benchmarking.

All parameters related to the same kind of service (traditional telephony, mobile telephony, etc.) should remain the same for wideband as compared to the narrowband case, except speech quality.

4.4.3.3 Speech Quality

In general the same parameters apply than in the narrowband case. However there is benefit of the higher bandwidth with respect to speech sound quality which will result in higher ratings with respect to sound quality for wideband systems. Other factors contributing to speech quality may be different as well (see clause 4.1) investigation on those is still ongoing.

It should be noticed that wideband systems when used e.g. in background noise situations may provide worse speech quality than narrowband systems if not properly implemented.

History

Document history		
V1.1.1	June 2006	Publication