Final draft ETSI EG 201 377-1 V1.3.2 (2009-08)

ETSI Guide

Speech and multimedia Transmission Quality (STQ); Specification and measurement of speech transmission quality; Part 1: Introduction to objective comparison measurement methods for one-way speech quality across networks

Reference
REG/STQ-00105-1

Keywords
interworking, quality, speech, testing, transmission, voice

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 ETSI Guide (EG) has been produced by ETSI Technical Committee Speech and multimedia Transmission Quality (STQ), and is now submitted for the ETSI standards Membership Approval Procedure.

The present document is part 1 of a multi-part deliverable covering the specification and measurement of speech transmission quality, as identified below:

Part 1: **"Introduction to objective comparison measurement methods for one-way speech quality across networks";**

Part 2: "Mouth-to-ear speech transmission quality including terminals";

Part 3: "Non-intrusive objective measurement methods applicable to networks and links with classes of services".

1 Scope

The present document is part 1 of a series of documents on the specification and measurement of mouth-to-ear (also end-to-end) speech transmission quality. Its main objective is to describe objective comparison-based methods and systems for measuring mouth-to-ear speech quality in networks. Apart from this, it gives an overview on other important aspects of mouth-to-ear speech quality. As the need arises, these other aspects will be covered in more detail in subsequent parts of the present document. Although some of the models described in the present document are superseded the description of the models is kept for information.

The present document gives an overview of the methods available for measuring one-way speech transmission quality. Its purpose is to give information and guidance primarily for operators, users, consumer organizations and regulators who wish to measure or compare the speech transmission quality provided by different networks. The need for the present document has been increased by:

- the liberalization of voice services, which has introduced alternative competing providers of voice services;
- the introduction of new mobile and IP based technologies;

which has increased the range of services and cost/quality options for users.

The present document applies to both fixed and mobile networks with or without terminal equipment connected to the network. It applies only for narrowband (i.e. between 300 Hz and 3 400 Hz) communications. In principle, comparison methods can be used for IP-based (internet protocol-based) networks, but further work is needed on the calibration of the methods for such networks. The present document describes:

- methods for measurements of individual impairments or combinations of impairments to be made at acoustic or electrical interfaces;
- methods for combining measures of different impairments into a single objective measure;
- methods for predicting the subjective effect of impairments that would be perceived by users.

The methods in the present document assume that subjects with normal hearing have been involved in the test. Therefore, the instrumental methods estimate the perceived speech quality of persons with normal hearing. For each method, the guide contains a general description to highlight the main points, and provides references for more detailed information. The present document does not contain detailed specifications of the individual methods.

The present document concentrates on *one-way* speech quality in networks. It gives no guidance on how to evaluate systems that include equipment such as echo cancellers or in which interactive impairments such as talker echo are significant. The perceived quality in such cases depends not only on the one-way performance, but very much on the behaviour of the equipment under duplex conditions; specifically, the influence of double-talk and delay needs to be considered.

Although all assessments of overall speech quality are ultimately subjective because they depend on the user's opinion, a distinction is made between:

- subjective methods, which involve real time user assessment; and
- objective methods, which use stored information on the user's assessment and therefore involve some degree of calibration.

Objective methods for the evaluation of speech quality fall into three categories:

a) *Comparison Methods:* Methods based on the comparison of transmitted speech signal and a known reference.

b) *Absolute Estimation Methods:* Methods based on the absolute estimation of the speech quality (i.e. there is no known reference signal); e.g. INMD (ITU-T Recommendation P.561 [i.16]).

c) *Transmission Rating Models:* Methods that derive a value for the expected speech quality from knowledge about the network; e.g. ETSI Model (ETR 250 [i.1], ITU-T Recommendation G.107 [i.14]).

The classification of assessment methods is depicted in figure 1.

Practical implementations of test equipment may include combinations of these methods. The focus of the present document is on comparison methods (intrusive methods), which currently yield the most accurate results. The other categories are only covered in short overviews, although they may be preferable for certain applications.

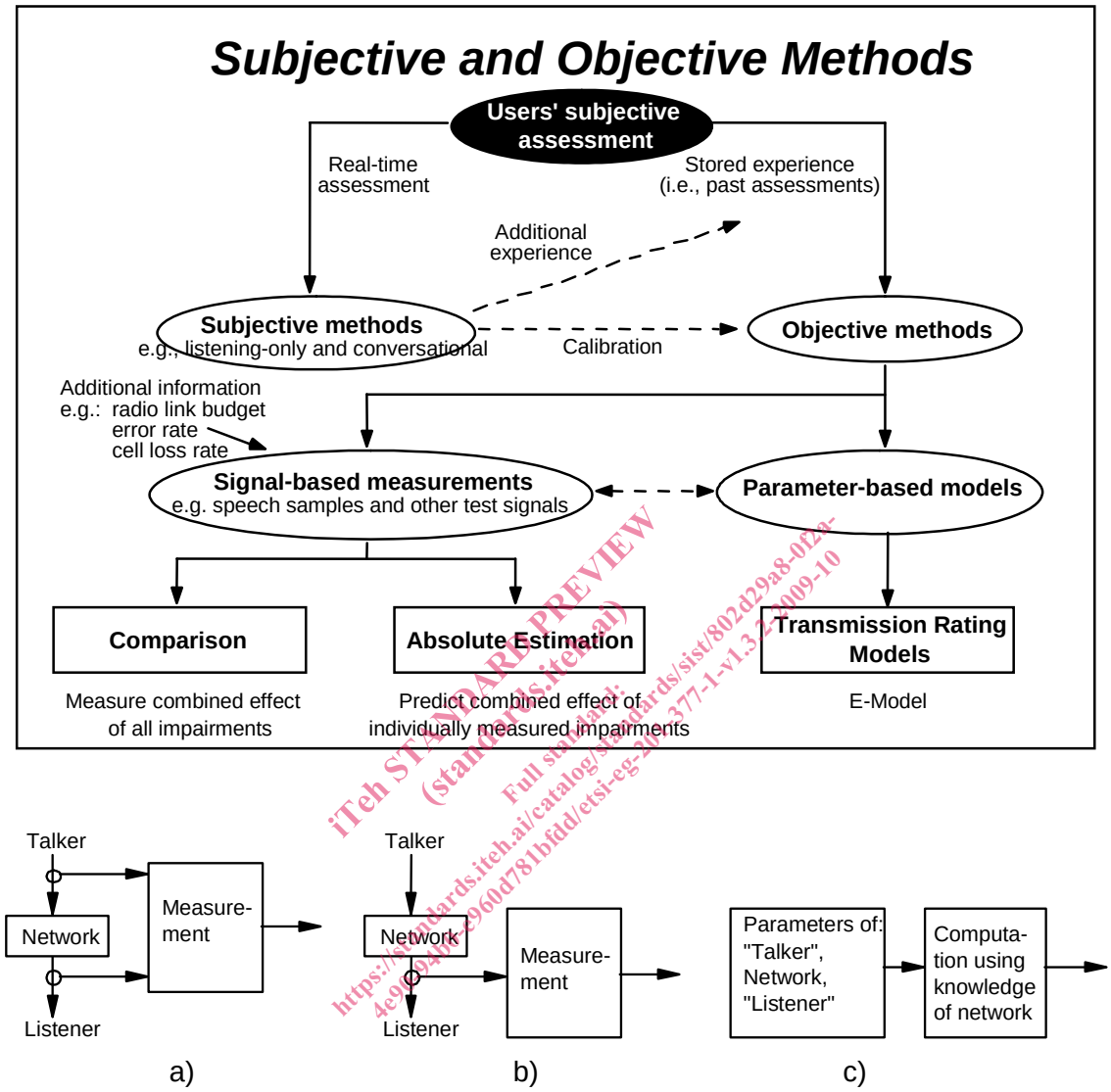

Figure 1: Classification of assessment methods showing:
a) Comparison methods,
b) Absolute estimation methods,
c) Transmission rating models

NOTE: As an ETSI Guide, the present document provides guidelines for test methods that may be implemented. However, a test method and especially quality models can only be applied in the way and within the scope defined in the reference standard. A "Warning" indicates when this applies.

2 References

References are either specific (identified by date of publication and/or edition number or version number) or non-specific.

- For a specific reference, subsequent revisions do not apply.
- Non-specific reference may be made only to a complete document or a part thereof and only in the following cases:
  - if it is accepted that it will be possible to use all future changes of the referenced document for the purposes of the referring document;
  - for informative references.

Referenced documents which are not found to be publicly available in the expected location might be found at http://docbox.etsi.org/Reference.

NOTE: While any hyperlinks included in this clause were valid at the time of publication ETSI cannot guarantee their long term validity.

2.1 Normative references

The following referenced documents are indispensable for the application of the present document. For dated references, only the edition cited applies. For non-specific references, the latest edition of the referenced document (including any amendments) applies.

Not applicable.

2.2 Informative references

The following referenced documents are not essential to the use of the present document but they assist the user with regard to a particular subject area. For non-specific references, the latest version of the referenced document (including any amendments) applies.

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

[i.2] ETSI EG 201 050: "Speech Processing, Transmission and Quality Aspects (STQ); Overall Transmission Plan Aspects for Telephony in a Private Network".

[i.3] ETSI TR 102 082: "Speech Processing, Transmission and Quality Aspects (STQ); Guidance on writing specifications and tests for non-linear and time variant telephony terminals".

[i.4] ETSI EG 202 396-1: "Speech and multimedia Transmission Quality (STQ); Speech quality performance in the presence of background noise; Part 1: Background noise simulation technique and background noise database".

[i.5] 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 - Subjective test database and results".

[i.6] ETSI EG 202 396-3: "Speech Processing, Transmission and Quality Aspects (STQ); Speech quality performance in the presence of background noise; Part 3: Background noise transmission - Objective test methods".

[i.7] ETSI ES 202 737: "Speech and multimedia Transmission Quality (STQ); Transmission requirements for narrowband VoIP terminals (handset and headset) from a QoS perspective as perceived by the user".

[i.8] ETSI ES 202 738: "Speech and multimedia Transmission Quality (STQ); Transmission requirements for narrowband VoIP loudspeaking and handsfree terminals from a QoS perspective as perceived by the user".

[i.9] ETSI ES 202 739: "Speech and multimediaTransmission Quality (STQ); Transmission requirements for wideband VoIP terminals (handset and headset) from a QoS perspective as perceived by the user".

[i.10] ETSI ES 202 740: "Speech and multimedia Transmission Quality (STQ); Transmission requirements for wideband VoIP loudspeaking and handsfree terminals from a QoS perspective as perceived by the user".

[i.11] EURESCOM Project P603 vol.1: "Quality of Service: Measurement Method Selection; Deliverable 2: Measurement Method; Volume 1 of 2: Main Report".

[i.12] EURESCOM Project P603 vol.2: "Quality of Service: Measurement Method Selection; Deliverable 2: Measurement Method; Volume 2 of 2: Annexes".

[i.13] ISO 532 (1975): "Acoustics - Method for calculating loudness level".

[i.14] ITU-T Recommendation G.107 (2002): "The E-model, a computational model for use in transmission planning".

[i.15] ITU-T Recommendation P.501: "Test signals for use in telephonometry".

[i.16] ITU-T Recommendation P.561 (2002): "In-service, non-intrusive measurement device - voice service measurements".

[i.17] ITU-T Recommendation P.562 (2004): "Analysis and interpretation of INMD voice-service measurements".

[i.18] ITU-T Recommendation P.563 (2004): "Single-ended method for objective speech quality assessment in narrow-band telephony applications".

[i.19] ITU-T Recommendation P.800 (1996): "Methods for subjective determination of transmission quality".

[i.20] ITU-T Recommendation P.830 (1996): "Subjective performance assessment of telephone-band and wideband digital codecs".

[i.21] ITU-T Recommendation P.835 (2003): " Subjective test methodology for evaluating speech communication systems that include noise suppression algorithm".

[i.22] ITU-T COM12-20: "Improvement of the P.861 perceptual speech quality measure".

[i.23] ITU-T COM12-24: "Proposed Annex A to Recommendation P.861".

[i.24] ITU-T COM12-34: "TOSQA - Telecommunication objective speech quality assessment".

[i.25] ITU-T COM12-62: "Results of Processing ITU speech database supplement 23 with the end-to-end quality assessment algorithm "PACE"".

[i.26] Journal of the Audio Engineering Society: "A Perceptual Audio Quality Measure Based on a Psychoacoustic Sound Representation", Beerends J.G., Stemerdink J.A. (1992), vol. 40, no. 12, pp. 963-978.

[i.27] BT Engineering Journal (1998): "Getting the message loud and clear: quantifying call clarity", Broom, S.; Coackley, P.; Sheppard, P., Vol. 17, p. 66-72.

[i.28] Speech Communication (1994): "Auditory Distortion Measure for Speech Coder Evaluation - Discrimination Information Approach", De A., Kabal P., 14(3):205-229.

[i.29] McMillan Publishing Company (1993): "Discrete Time Processing of Speech Signals", Deller J.R., Proakis J.G., Hansen J.H.L. Eaglewood Cliffs NJ.

[i.30] Report TA No. 92, KTH Karolinska Institutet(1979): "Statistical treatment of data from listening tests on sound-reproducing systems", Gabrielsson A, Department of Technical Audiology, S-10044 Stockholm, Sweden.

[i.31] Prentice Hall Press (1995): "Introduction to Mathematical Statistics", Hogg R.V., Craig A.T., Eaglewood Cliffs.

[i.32] IEEE Proceedings-Vision, Image and Signal Processing 141 (3) (1994): "Error activity and error entropy as a measure of psychoacoustic significance in the perceptual domain", Hollier, M.P.; Hawksford, M.O.; Guard, D.R. pp. 203-208.

[i.33] Proceedings of IEEE ICC '91: "Comparison of four objective speech quality assessment methods based on international subjective evaluations of universal codecs", Irii H., pp. 1726-1730.

[i.34] Proceedings of IEEE 5th International Workshop on Systems, Signals and Image Processing IWSSIP'98: "An Objective Speech Quality Measurement in the QVoice", Juriæ P., pp. 156-163.

[i.35] Springer-Verlag (1990): "Psychoacoustics, facts and models", Berlin, Heidelberg Zwicker E., Fastl H.

[i.36] ITU-T Recommendation P.862 (2001): "Perceptual evaluation of speech quality (PESQ), an objective method for end-to-end speech quality assessment of narrowband telephone networks and speech codecs".

[i.37] ITU-T Recommendation G.168: "Digital network echo cancellers".

[i.38] ITU-T Recommendation P.831: "Subjective performance evaluation of network echo cancellers".

[i.39] ITU-T COM12-6: "Subjective evaluation of hands-free telephones using conversational test, specific double talk test and listening only test".

[i.40] Speech Communication 20 (1996): "The Auditory Perceived Quality of Hands-Free Telephones: Auditory Judgements, Instrumental Measurements and Their Relationship", Gierlich, H.W. (1996) p 241-254.

[i.41] ITU-T Recommendation P.58 (1996): "Head and torso simulator for telephonometry".

[i.42] ITU-T Recommendation P.64 (1999): "Determination of sensitivity/frequency characteristics of local telephone systems".

[i.43] ITU-T Recommendation P.57 (2002): "Artificial ears".

[i.44] ITU-T Recommendation P.340 (2000): "Transmission characteristics of hands-free telephones".

[i.45] Gierlich, H.W.; Kettler, F., Diedrich, E.: "Speech Quality Evaluation of Hands-Free Telephones During Double talk: New Evaluation Methodologies"; EUSIPCO '98, Rhodos, Greece, Conference Proceedings, vol. 2, pp. 953 - 956, 1998.

[i.46] CCITT Supplement No. 5 to Recommendation P.74: "The SIBYL Method of Subjective Testing", Red Book, Volume V.

[i.47] ITU-T Recommendation P.82 (1988): "Method for evaluation of service from the standpoint of speech transmission quality".

[i.48] IEEE Signal Processing Magazine: "The bootstrap and its application in signal processing", pp. 56-76, Zoubir A.M., Boashash B., January 1998.

[i.49] ETSI TBR 008: "Integrated Services Digital Network (ISDN); Telephony 3,1 kHz teleservice; Attachment requirements for handset terminals".

[i.50] ETSI TBR 009: "European digital cellular telecommunications system; Attachment requirements for Global System for Mobile communications (GSM) mobile stations; Telephony".

[i.51] ETSI TBR 010: "Digital Enhanced Cordless Telecommunications (DECT); General terminal attachment requirements: Telephony applications".

[i.52] ETSI ES 203 038: "Speech and multimedia Transmission Quality (STQ); Requirements and tests methods for terminal equipment incorporating a handset when connected to the analogue interface of the PSTN".

[i.53] InterNoise'96: "Objective Evaluation of Acoustic Quality Based on a Relative Approach", Genuit, K., Liverpool, UK.

[i.54] ITU-T Contribution COM 12-C178: "Towards a New E-Model Impairment Factor for Linear Distortion of Narrowband and Wideband Speech Transmission", Germany.

[i.55] ITU-T Worshop: "From Speech to Audio": "Echo perception in wideband telecommunication Scenarios - Comparison to E-Model's Narrowband Echo Findings", H.W. Gierlich, Silvia Poschen, Frank Kettler, Alexander Raake, Sascha Spors, Matthias Geier Sept.08, Lannion.

[i.56] ITU-T Recommendation P.805: "Subjective evaluation of conversational quality".

[i.57] ITU-T Recommendation P.861: "Objective quality measurement of telephone-band (300-3400 Hz) speech codecs".

[i.58] Klaus H., Berger J. (1997): "Die Bestimmung der Telefon-Sprachqualität für die Übertragungskette vom Mund zum Ohr - Herausforderungen und ausgewählte Verfahren. Deutsche Telekom".

[i.59] Berger, J. (1998): "Instrumentelle Verfahren zur Sprachqualitätsschätzung-Modelle auditiver Tests (Instrumental approaches for speech quality estimation-models of auditory tests)", Ph.D. thesis, Christian-Albrechts-University of Kiel, Shaker-Verlag. ISBN 3-8265-4092-3.

[i.60] ITU-T Recommendation P.10/G.100: "Vocabulary for performance and quality of service".

3 Definitions and abbreviations

3.1 Definitions

For the purposes of the present document, the following terms and definitions apply:

bark: frequency unit in the perceptual domain; e.g. frequencies at 3, 4, and 5 Bark are perceived as equally-spaced

cepstrum: cepstrum of a signal is defined as the inverse Fourier transform of the logarithm of the power spectrum of that signal

NOTE 1: See figure 5.

NOTE 2: Linear distortions of a signal (e.g. delay, echo) are additive in the cepstral domain.

cognitive: pertaining to higher layers of human reception; e.g. interpretation of speech

perceptual: pertaining to lower layers of human reception; e.g. processing of sound signals

psycho-acoustic: pertaining to acoustic processing particular to the human sound perception system; e.g. masking of adjacent frequency components

3.2 Abbreviations

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

ACR Absolute Category Rating
ANOVA ANalysis Of VAriances
ATM Asynchronous Transfer Mode
CCR Comparison Category Rating
CD Cepstral Distance
CDI Cochlear Discrimination Information

CMOS	Comparison Mean Opinion Scores
DC	Direct Current
DCME	Digital Circuit Multiplication Equipment
DCR	Degradation Category Rating
DFT	Discrete Fourier Transform
DMOS	Degradation Mean Opinion Scores
FFT	Fast Fourier Transform
FMNB	Frequency Measuring Normalizing Block
GSM	Global System for Mobile communication
HATS	Heads And Torso Simulator
INMD	In-service, Non-intrusive Measurement Device
IP	Internet Protocol
ISDN	Integrated Services Digital Network
LAR	Log-Area Ratios
LPC	Linear Prediction Coefficient
MNB	Measuring Normalizing Blocks
MOS	Mean Opinion Score
PAMS	Perceptual Analysis/Measurement System
PCM	Pulse Code Modulation
PESQ	Perceptual Evaluation of Speech Quality
POTS	Plain Old Telephony Service
PSQM	Perceptual Speech Quality Measure
PSTN	Public Switched Telephone Network
QoS	Quality of Service
QSDG	Quality of Service Development Group
SNR	Signal-to-Noise Ratio
TMNB	Time Measuring Normalizing Block
TOSQA	Telecommunication Objective Speech Quality Assessment

4 Overview

Today, telecommunication is strongly influenced by three major facts:

- the liberalization of telecommunication, i.e. the separation between regulatory bodies and operators;
- the splitting of operations into network providers and service providers; and
- the increase of international traffic due to the internationalization of trade and business.

In addition to these facts, there is also a strong influence due to technical evolution. The most important trends are the move from fixed networks to mobile networks, but also from conventional switched PSTN and ISDN networks to packet-based networks such as the Internet. These technical trends will make it necessary to extend the applicability of the methods described below in order to cover speech quality impairments from "new" types of degradations, such as packet losses and variable delay.

The liberalization as well as the splitting of operations lead to new legal/commercial/technical interfaces, which need a definition both in the contractual and technical sense:

- regulators need a measurement basis in order to specify the requirements which "their" network operators have to fulfil;
- operators of private networks (e.g. corporate networks, closed user groups) need a measurement basis as well for double-checking transmission planning issues for the interconnection of private networks with the public ISDN/PSTN; and
- service providers want to compare different network providers concerning their price/performance ratio.

In all cases the traditional methods for speech quality assessment based on subjective rating of speech samples are far too expensive, too slow and lack the precise repeatability.

The internationalization of traffic as well as the multitude of network providers lead to the fact that in many cases a phone call is routed through several networks, where these networks are based on different technologies (fixed analogue or digital, ATM, Internet, mobile networks, satellite links, etc.). The concatenation of multiple different networks is no longer restricted, and the resulting effects on speech quality are not well covered up to now.

4.1 Objective

The aim of the present document is to give:

- general information on mouth-to-ear speech quality, and the factors to be included in its evaluation (see clause 5);
- information on subjective reference assessment methods, which are essential to calibrate objective methods, showing what results can be obtained (see clause 6 and annex C);
- information on the objective comparison measurement methods available and how they work, especially the most recent methods (see clause 7);
- overview of other assessment methods (see clauses 8 and 9).

In a second part of the present document (to be developed later), the criteria for the evaluation of such objective measurement systems will be specified, namely:

- requirements concerning the technical characteristics of speech quality measurement;
- methods to test the conformity of these methods to the subjective reference assessments; and finally
- criteria to compare and evaluate the current methods.

4.2 Related work in standardization

On all of the above mentioned topics a lot of work has already been done in the past by a number of standards bodies:

- ETSI TC STQ http://portal.etsi.org/STQ
  This Technical Committee is responsible for the "co-ordination, production (where appropriate) and maintenance of end-to-end speech quality related deliverables" (TC/STQ Terms of Reference).
- 3GPP SA
  The work done in 3GPP SA 4 concentrates on codec quality in mobile networks (in particular for Half Rate, Enhanced Full Rate and Adaptive Multi-Rate codecs) and therefore is not primarily oriented towards mouth-to-ear speech quality aspects. However, it is a very important source of information especially for the subjective rating of speech samples and for the characteristics of speech samples to be used for assessment and measurement. Note that this work done in SA 4 was previously performed by ETSI SMG 11.
- ITU-T Study Group 12
  The work in ITU-T Study Group 12 is focused both on terminal and acoustic tests and on mouth-to-ear network aspects. Several questions are addressing mouth-to-ear speech quality issues. The details of the questions can be found:

  http://www.itu.int/ITU-T/studygroups/com12/index.asp
- ITU-T Study Group 2/QSDG
  The "Quality of Service Development Group" is a subgroup of ITU-T Study Group 2. Its members are network operators and manufacturers from all over the world.
  According to their Terms Of Reference, the tasks of QSDG are the following:
  - encourage participation in QoS activities;
  - identify and develop performance monitoring and evaluation;
  - improve QoS, include practices in TSS documentation;
  - disseminate information about QoS techniques and procedures;