

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

P.800.1

(07/2006)

SERIES P: TELEPHONE TRANSMISSION QUALITY,
TELEPHONE INSTALLATIONS, LOCAL LINE
NETWORKS

Methods for objective and subjective assessment of
quality

Mean Opinion Score (MOS) terminology

ITU-T Recommendation P.800.1

ITU-T P-SERIES RECOMMENDATIONS

TELEPHONE TRANSMISSION QUALITY, TELEPHONE INSTALLATIONS, LOCAL LINE NETWORKS

Vocabulary and effects of transmission parameters on customer opinion of transmission quality	Series	P.10
Subscribers' lines and sets	Series	P.30
		P.300
Transmission standards	Series	P.40
Objective measuring apparatus	Series	P.50
		P.500
Objective electro-acoustical measurements	Series	P.60
Measurements related to speech loudness	Series	P.70
Methods for objective and subjective assessment of quality	Series	P.80
		P.800
Audiovisual quality in multimedia services	Series	P.900
Transmission performance and QoS aspects of IP end-points	Series	P.1000

For further details, please refer to the list of ITU-T Recommendations.

ITU-T Recommendation P.800.1

Mean Opinion Score (MOS) terminology

Summary

This Recommendation provides a terminology which shall be used in conjunction with speech quality expressions in terms of Mean Opinion Score (MOS). The new terminology is motivated by the intention to avoid misunderstanding as to whether specific values of MOS are related to listening quality, talking quality or conversational quality, and whether they originate from subjective tests, from objective models or from network planning models. Furthermore, this revision provides identifiers regarding the bandwidth and the type of interface (electrical or acoustical).

Source

ITU-T Recommendation P.800.1 was approved on 14 July 2006 by ITU-T Study Group 12 (2005-2008) under the ITU-T Recommendation A.8 procedure.

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications. The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of ITU. ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Assembly (WTSA), which meets every four years, establishes the topics for study by the ITU-T study groups which, in turn, produce Recommendations on these topics.

The approval of ITU-T Recommendations is covered by the procedure laid down in WTSA Resolution 1.

In some areas of information technology which fall within ITU-T's purview, the necessary standards are prepared on a collaborative basis with ISO and IEC.

NOTE

In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

Compliance with this Recommendation is voluntary. However, the Recommendation may contain certain mandatory provisions (to ensure e.g. interoperability or applicability) and compliance with the Recommendation is achieved when all of these mandatory provisions are met. The words "shall" or some other obligatory language such as "must" and the negative equivalents are used to express requirements. The use of such words does not suggest that compliance with the Recommendation is required of any party.

INTELLECTUAL PROPERTY RIGHTS

ITU draws attention to the possibility that the practice or implementation of this Recommendation may involve the use of a claimed Intellectual Property Right. ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU members or others outside of the Recommendation development process.

As of the date of approval of this Recommendation, ITU had not received notice of intellectual property, protected by patents, which may be required to implement this Recommendation. However, implementors are cautioned that this may not represent the latest information and are therefore strongly urged to consult the TSB patent database.

© ITU 2006

All rights reserved. No part of this publication may be reproduced, by any means whatsoever, without the prior written permission of ITU.

CONTENTS

	Page
1 References.....	1
2 Recommended MOS terminology	1
2.1 MOS related to listening-only situations.....	2
2.2 MOS related to conversational situations.....	3
2.3 MOS related to talking situations	3
3 Relationship between some MOS qualifiers.....	4

ITU-T Recommendation P.800.1

Mean Opinion Score (MOS) terminology

1 References

The following ITU-T Recommendations and other references contain provisions which, through reference in this text, constitute provisions of this Recommendation. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; users of this Recommendation are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations and other references listed below. A list of the currently valid ITU-T Recommendations is regularly published. The reference to a document within this Recommendation does not give it, as a stand-alone document, the status of a Recommendation.

- [1] ITU-T Recommendation P.10/G.100 (2006), *Vocabulary for performance and quality of service*.
- [2] ITU-T Recommendation G.107 (2005), *The E-model, a computational model for use in transmission planning*.
- [3] ITU-T Recommendation G.113 Appendix I (2002), *Provisional planning values for the equipment impairment factor I_e and packet-loss robustness factor B_{pl}* .
- [4] ITU-T Recommendation P.562 (2004), *Analysis and interpretation of INMD voice-service measurements*.
- [5] ITU-T Recommendation P.800 (1996), *Methods for subjective determination of transmission quality*.
- [6] ITU-T Recommendation P.830 (1996), *Subjective performance assessment of telephone-band and wideband digital codecs*.
- [7] ITU-T Recommendation P.833 (2001), *Methodology for the derivation of equipment impairment factors from subjective listening-only tests*.
- [8] ITU-T Recommendation P.834 (2002), *Methodology for the derivation of equipment impairment factors from instrumental models*.
- [9] ITU-T Recommendation P.862 (2001), *Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs*.

2 Recommended MOS terminology

The abbreviation MOS (Mean Opinion Score) is defined in ITU-T Rec. P.10/G.100 in the following way:

- The mean of opinion scores, i.e., of the values on a predefined scale that subjects assign to their opinion of the performance of the telephone transmission system used either for conversation or for listening to spoken material.

Apart from subjective opinion, the abbreviation MOS is also used for scores that originate from objective models or network planning models. The following identifiers are recommended to be used together with the abbreviation MOS in order to distinguish the area of application, where N refers to narrow-band, W refers to wideband, LQ refers to Listening Quality, CQ refers to Conversational Quality, S refers to Subjective, O refers to Objective, and E refers to Estimated.

	Listening-only	Conversational	Talking
Subjective	MOS-LQSy	MOS-CQSy	MOS-TQSy
Objective	MOS-LQOy	MOS-CQOy	MOS-TQOy
Estimated	MOS-LQEy	MOS-CQEy	MOS-TQEy

NOTE – The letter "y" at the end of above acronyms is a placeholder for the descriptor of the respective audio bandwidth, see the following provisional instructions.

- N for MOS scores obtained for narrow-band (300-3400 Hz) speech relative to a narrow-band high quality reference. This is applicable for instance to narrow-band only subjective tests or to P.862.1 scores.
- W for MOS scores obtained for wideband (50-7000 Hz) speech relative to a wideband high quality reference. This is applicable for instance to wideband only subjective tests or to P.862.2 scores.
- M for MOS scores obtained for narrow-band or wideband speech relative to a wideband high quality reference in a mixed bandwidths context. This is applicable for instance to mixed bandwidths subjective tests.

The effects of audio bandwidth on MOS scores is currently under investigation in ITU-T Study Group 12. In cases where the bandwidth denominators N, W or M do not properly reflect the actual situation, it is suggested that provisionally the placeholder "y" be replaced by a proper notation.

2.1 MOS related to listening-only situations

These MOS scores are applicable to a listening-only situation. Three different cases have to be distinguished.

2.1.1 MOS-LQS

The score has been collected in a laboratory test by calculating the arithmetic mean value of subjective judgments on a 5-point ACR quality scale, as it is defined in ITU-T Rec. P.800. Subjective tests carried out according to ITU-T Recs P.830, P.835 and P.840 give results in terms of MOS-LQS.

2.1.2 MOS-LQO

The score is calculated by means of an objective model which aims at predicting the quality for a listening-only test situation. Objective measurements made using the model given in ITU-T Recs P.862.1 and P.862.2 give results in terms of MOS-LQO.

It should be noted that the method recommended by ITU-T in Recs P.862.1 and P.862.2 is validated between electrical interfaces only. Currently no ITU Recommendation exists which covers the measurement of listening quality including acoustical interfaces. Work on a new ITU-T Rec. P.OLQA which is intended to include acoustical interfaces is under progress.

In this context, it should be noted that there is a necessary distinction between intrusive end-to-end measurements of listening quality aiming at:

- **MOS-LQO (electrical)**

This kind of measurement is performed at electrical interfaces only. In order to predict the listening quality as perceived by the user, assumptions for the terminals are made in terms of IRS or corrected IRS frequency response; this implicitly includes the assumption of a sealed condition between the handset receiver and the user's ear. ITU-T Rec. P.862 falls into this category.

- **MOS-LQO (acoustical)**

This kind of measurement is performed at acoustical interfaces. In order to predict the listening quality as perceived by the user, this measurement includes the actual telephone set products provided by the manufacturer or vendor. In combination with the choice of the acoustical receiver in the lab test ("artificial ear"), there will be a more or less leaky condition between the handset's receiver and the artificial ear. Consequently, for more realistic test scenarios, there may be a degradation of the measured MOS value, while for more artificial test scenarios there may be a negligible difference.

2.1.3 MOS-LQE

The score is calculated by a network planning model which aims at predicting the quality in a listening-only application situation.

2.2 MOS related to conversational situations

These MOS scores are applicable to a conversational situation. Three different cases have to be distinguished.

2.2.1 MOS-CQS

The score has been collected in a laboratory test by calculating the arithmetic mean value of subjective judgments on a 5-point ACR quality scale, as it is defined in ITU-T Rec. P.800. Subjective conversation tests carried out according to ITU-T Recs P.800, P.831 and P.832 give results in terms of MOS-CQS.

2.2.2 MOS-CQO

The score is calculated by means of an objective model which aims at predicting the quality for a conversational test situation. Objective measurements made using the model given in ITU-T Recs P.562 and P.563 give results in terms of MOS-CQO.

2.2.3 MOS-CQE

The score is calculated by a network planning model which aims at predicting the quality in a conversational application situation. Estimates of conversational quality carried out according to ITU-T Rec. G.107, when transformed to mean opinion score, give results in terms of MOS-CQE.

2.3 MOS related to talking situations

Talking quality describes the quality of a telephone call as it is perceived by the talking party only. Talking quality will be mainly affected by the annoyance of the echo signal and effects like background noise switching and double talk.

2.3.1 MOS-TQS

The score has been collected in a laboratory test by calculating the arithmetic mean value of subjective judgments on a 5-point ACR quality scale, as it is defined in ITU-T Rec. P.800.

2.3.2 MOS-TQO

The score is calculated by means of an objective model which aims at predicting the quality for a talking-only test situation. Methods generating a MOS-TQO are currently under development and not yet standardized.

2.3.3 MOS-TQE

The score is calculated by a network planning model which aims at predicting the quality in a talking-only application situation. No methods generating a MOS-TQE are currently standardized.

3 Relationship between some MOS qualifiers

Figure 1 gives an overview of the relation between MOS-LQS, MOS-LQO and MOS-LQE.

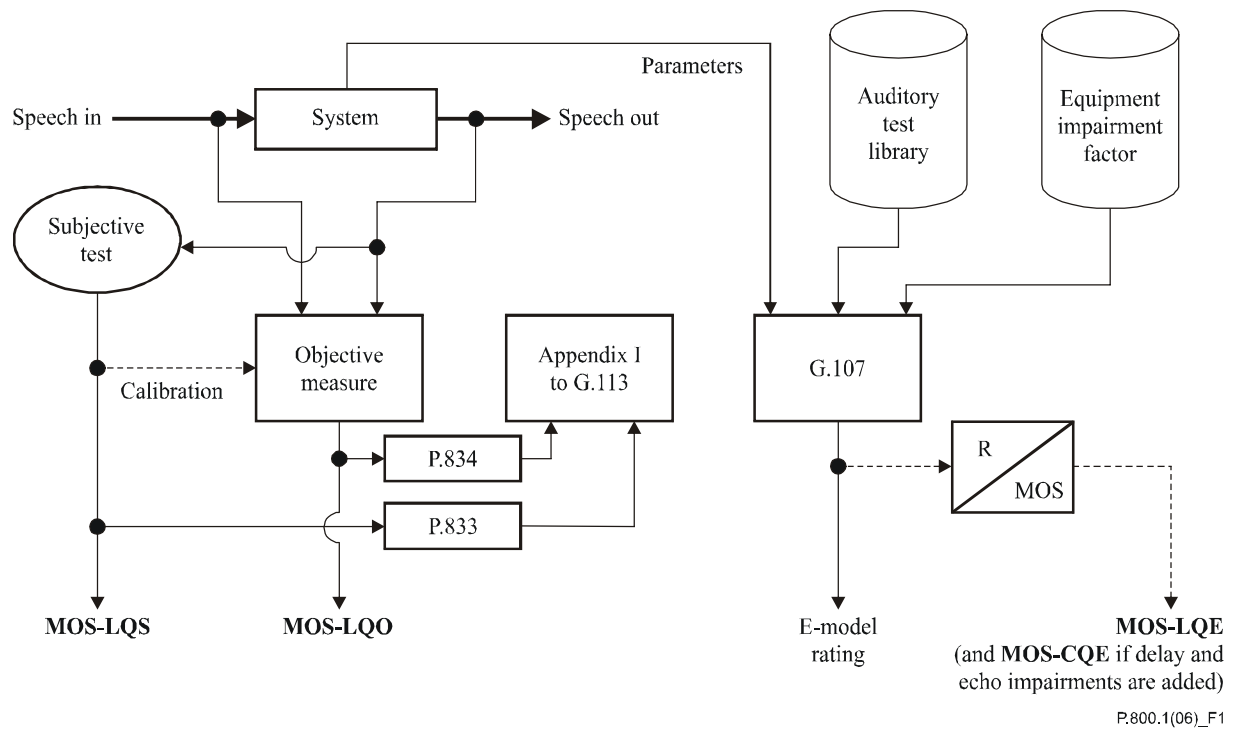


Figure 1/P.800.1 – Relationship between some MOS qualifiers

SERIES OF ITU-T RECOMMENDATIONS

Series A	Organization of the work of ITU-T
Series D	General tariff principles
Series E	Overall network operation, telephone service, service operation and human factors
Series F	Non-telephone telecommunication services
Series G	Transmission systems and media, digital systems and networks
Series H	Audiovisual and multimedia systems
Series I	Integrated services digital network
Series J	Cable networks and transmission of television, sound programme and other multimedia signals
Series K	Protection against interference
Series L	Construction, installation and protection of cables and other elements of outside plant
Series M	Telecommunication management, including TMN and network maintenance
Series N	Maintenance: international sound programme and television transmission circuits
Series O	Specifications of measuring equipment
Series P	Telephone transmission quality, telephone installations, local line networks
Series Q	Switching and signalling
Series R	Telegraph transmission
Series S	Telegraph services terminal equipment
Series T	Terminals for telematic services
Series U	Telegraph switching
Series V	Data communication over the telephone network
Series X	Data networks, open system communications and security
Series Y	Global information infrastructure, Internet protocol aspects and next-generation networks
Series Z	Languages and general software aspects for telecommunication systems