ETSI ES 202 738 V1.5.1 (2017-01)

**Speech and multimedia Transmission Quality (STQ);
Transmission requirements for narrowband
VoIP loudspeaking and handsfree terminals
from a QoS perspective as perceived by the user**

Reference
RES/STQ-243

Keywords
handsfree, loudspeaking, narrowband, quality, speech, terminal, VoIP

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

The present document can be downloaded from:
http://www.etsi.org/standards-search

The present document may be made available in electronic versions and/or in print. The content of any electronic and/or print versions of the present document shall not be modified without the prior written authorization of ETSI. In case of any existing or perceived difference in contents between such versions and/or in print, the only prevailing document is the print of the Portable Document Format (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
https://portal.etsi.org/TB/ETSIDeliverableStatus.aspx

If you find errors in the present document, please send your comment to one of the following services:
https://portal.etsi.org/People/CommiteeSupportStaff.aspx

Copyright Notification

No part may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm except as authorized by written permission of ETSI.
The content of the PDF version shall not be modified without the written authorization of ETSI.
The copyright and the foregoing restriction extend to reproduction in all media.

© European Telecommunications Standards Institute 2017.
All rights reserved.

DECT™, **PLUGTESTS**™, **UMTS**™ and the ETSI logo are Trade Marks of ETSI registered for the benefit of its Members.
3GPP™ and **LTE**™ are Trade Marks of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners.
GSM® and the GSM logo are Trade Marks registered and owned by the GSM Association.

Contents

Intellectual Property Rights 5
Foreword 5
Modal verbs terminology 5
Introduction 5
1 Scope 6
2 References 6
2.1 Normative references 6
2.2 Informative references 7
3 Definitions and abbreviations 8
3.1 Definitions 8
3.2 Abbreviations 9
4 General considerations 10
4.1 Default Coding Algorithm 10
4.2 End-to-end considerations 10
5 Test equipment 10
5.1 IP half channel measurement adaptor 10
5.2 Environmental conditions for tests 10
5.3 Accuracy of measurements and test signal generation 11
5.4 Network impairment simulation 11
5.5 Acoustic environment 12
5.6 Influence of terminal delay on measurements 12
6 Requirements and associated measurement methodologies 13
6.1 Notes 13
6.2 Test setup 13
6.2.1 General 13
6.2.2 Setup for terminal 14
6.2.2.1 Hands-free measurements 14
6.2.2.2 Measurements in loudspeaking mode 18
6.2.3 Test signal levels 18
6.2.3.1 Send 18
6.2.3.2 Receive 19
6.2.4 Setup of background noise simulation 19
6.3 Coding independent parameters 20
6.3.1 Send frequency response 20
6.3.2 Send Loudness Rating (SLR) 21
6.3.3 Mic mute 21
6.3.4 Send distortion 21
6.3.5 Out-of-band signals in send direction 22
6.3.6 Send noise 23
6.3.7 Terminal Coupling Loss weighted (TCLw) 23
6.3.8 Stability loss 23
6.3.9 Receive frequency response 24
6.3.10 Receive Loudness Rating (RLR) 26
6.3.11 Receive Distortion 27
6.3.12 Out-of-band signals in receive direction 28
6.3.13 Receive noise 28
6.3.14 Double talk performance 29
6.3.14.1 General 29
6.3.14.2 Attenuation range in send direction during double talk $A_{H,S,dt}$ 29
6.3.14.3 Attenuation range in receive direction during double talk $A_{H,R,dt}$ 30
6.3.14.4 Detection of echo components during double talk 31
6.3.14.5 Minimum activation level and sensitivity of double talk detection 32

6.3.15 Switching characteristics 32
6.3.15.1 Note 32
6.3.15.2 Activation in send direction 32
6.3.15.3 Silence suppression and comfort noise generation 33
6.3.16 Background noise performance 33
6.3.16.1 Performance in send direction in the presence of background noise 33
6.3.16.2 Speech quality in the presence of background noise 34
6.3.16.3 Quality of background noise transmission (with far end speech) 35
6.3.17 Quality of echo cancellation 35
6.3.17.1 Temporal echo effects 35
6.3.17.2 Spectral echo attenuation 36
6.3.17.3 Occurrence of artefacts 37
6.3.17.4 Variable echo path 37
6.3.18 Variant impairments; network dependant 38
6.3.18.1 Clock accuracy send 38
6.3.18.2 Clock accuracy receive 38
6.3.18.3 Send delay variation 38
6.3.19 Send and receive delay - round trip delay 39
6.4 Codec specific requirements 41
6.4.1 Objective listening speech quality MOS-LQO in send direction 41
6.4.2 Objective listening quality MOS-LQO in receive direction 42
6.4.3 Quality of jitter buffer adjustment 43

Annex A (informative): Processing delays in VoIP terminals 46

Annex B (informative): Bibliography 49

History 50

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 (https://ipr.etsi.org/).

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 Standard (ES) has been produced by ETSI Technical Committee Speech and multimedia Transmission Quality (STQ).

Modal verbs terminology

In the present document "**shall**", "**shall not**", "**should**", "**should not**", "**may**", "**need not**", "**will**", "**will not**", "**can**" and "**cannot**" are to be interpreted as described in clause 3.2 of the ETSI Drafting Rules (Verbal forms for the expression of provisions).

"**must**" and "**must not**" are **NOT** allowed in ETSI deliverables except when used in direct citation.

Introduction

Traditionally, the analogue and digital telephones were interfacing switched-circuit 64 kbit/s PCM networks. With the fast growth of IP networks, terminals directly interfacing packet-switched networks (VoIP) are being rapidly introduced. Such IP network edge devices may include specifically designed IP phones, soft phones or other devices connected to the IP based networks and providing telephony service. Since the IP networks will be in many cases interworking with the traditional PSTN and private networks, many of the basic transmission requirements have to be harmonised with specifications for traditional digital terminals. However, due to the unique characteristics of the IP networks including packet loss, delay, etc. new performance specification, as well as appropriate measuring methods, will have to be developed. Terminals are getting increasingly complex. Advanced signal processing is used to address the IP specific issues. Also, the VoIP terminals may use other than 64 kbit/s PCM (Recommendation ITU-T G.711 [7]) speech algorithms.

The advanced signal processing of terminals is targeted to speech signals. Therefore, wherever possible speech signals are used for testing in order to achieve mostly realistic test conditions and meaningful results.

The present document provides speech transmission performance requirements for narrowband VoIP loudspeaking and hands-free terminals.

NOTE: Requirement limits are given in tables, the associated curve when provided is given for illustration.

1 Scope

The present document will provide speech transmission performance requirements for narrowband VoIP loudspeaking and hands-free terminals; it addresses all types of IP based terminals, including wireless, softphones and group audio terminals.

In contrast to other standards which define minimum performance requirements it is the intention of the present document to specify terminal equipment requirements which enable manufacturers and service providers to enable good quality end-to-end speech performance as perceived by the user.

In addition to basic testing procedures, the present document describes advanced testing procedures taking into account further quality parameters as perceived by the user.

NOTE: The present document does not concern headset terminals.

2 References

2.1 Normative references

References are either specific (identified by date of publication and/or edition number or version number) or non-specific. For specific references, only the cited version applies. For non-specific references, the latest version of the referenced document (including any amendments) applies.

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.

The following referenced documents are necessary for the application of the present document.

[1] ETSI I-ETS 300 245-3: "Integrated Services Digital Network (ISDN); Technical characteristics of telephony terminals; Part 3: Pulse Code Modulation (PCM) A-law, loudspeaking and handsfree telephony".

[2] ETSI EN 300 726: "Digital cellular telecommunications system (Phase 2+) (GSM); Enhanced Full Rate (EFR) speech transcoding (GSM 06.60)".

[3] ETSI TS 126 171: "Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); LTE; Speech codec speech processing functions; Adaptive Multi-Rate - Wideband (AMR-WB) speech codec; General description (3GPP TS 26.171)".

[4] Recommendation ITU-T G.108: "Application of the E-model: A planning guide".

[5] Recommendation ITU-T G.109: "Definition of categories of speech transmission quality".

[6] Recommendation ITU-T G.122: "Influence of national systems on stability and talker echo in international connections".

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

[8] Recommendation ITU-T G.723.1: "Dual rate speech coder for multimedia communications transmitting at 5.3 and 6.3 kbit/s".

[9] Recommendation ITU-T G.726: "40, 32, 24, 16 kbit/s Adaptive Differential Pulse Code Modulation (ADPCM)".

[10] Recommendation ITU-T G.729: "Coding of speech at 8 kbit/s using conjugate-structure algebraic-code-excited linear prediction (CS-ACELP)".