

International Telecommunication Union

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

P.1311

(12/2014)

SERIES P: TERMINALS AND SUBJECTIVE AND
OBJECTIVE ASSESSMENT METHODS

Telemeeting assessment

Method for determining the intelligibility of multiple concurrent talkers

Recommendation ITU-T P.1311

ITU-T

ITU-T P-SERIES RECOMMENDATIONS

TERMINALS AND SUBJECTIVE AND OBJECTIVE ASSESSMENT METHODS

Vocabulary and effects of transmission parameters on customer opinion of transmission quality	Series	P.10
Voice terminal characteristics	Series	P.30
		P.300
Reference systems	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 speech 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
Communications involving vehicles	Series	P.1100
Models and tools for quality assessment of streamed media	Series	P.1200
Telemeeting assessment	Series	P.1300
Statistical analysis, evaluation and reporting guidelines of quality measurements	Series	P.1400
Methods for objective and subjective assessment of quality of services other than voice services	Series	P.1500

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

Recommendation ITU-T P.1311

Method for determining the intelligibility of multiple concurrent talkers

Summary

Recommendation ITU-T P.1311 describes a method for conducting a listening test that measures the intelligibility of multiple concurrent talkers in a teleconference. This Recommendation specifies how to conduct such a test, giving detail on stimulus design, creation of source material, selection of test conditions, calibration, selection and training of listeners, test administration, and reporting of results. This Recommendation includes an electronic attachment containing American English source material referenced in Appendix I.

History

Edition	Recommendation	Approval	Study Group	Unique ID*
1.0	ITU-T P.1311	2014-12-22	12	11.1002/1000/12326

Keywords

Concurrent speech, coordinate response measure, CRM, spatial release from masking, speech intelligibility.

* To access the Recommendation, type the URL <http://handle.itu.int/> in the address field of your web browser, followed by the Recommendation's unique ID. For example, <http://handle.itu.int/11.1002/1000/11830-en>.

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications, information and communication technologies (ICTs). 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, implementers are cautioned that this may not represent the latest information and are therefore strongly urged to consult the TSB patent database at <http://www.itu.int/ITU-T/ipr/>.

© ITU 2015

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

Table of Contents

		Page
1	Scope.....	1
2	References.....	2
3	Definitions	2
	3.1 Terms defined elsewhere	2
	3.2 Terms defined in this Recommendation.....	2
4	Abbreviations and acronyms	3
5	Conventions	3
6	Background.....	3
7	Test methodology	3
	7.1 Overview	3
	7.2 Test conditions.....	4
	7.3 Test stimuli	6
	7.4 Test environment	7
	7.5 Experiment design	7
	7.6 Response format and collection.....	8
	7.7 Listener selection.....	8
	7.8 Reporting and interpretation of results	9
	Appendix I – Examples of call signs, colours and numbers for English and standard Chinese	10
	Appendix II – Examples of test condition assignment	11
	Appendix III – Listener instructions and user interface (Example).....	13
	Bibliography.....	15

Electronic attachment: English speech samples

Copyright International Telecommunication Union/ITU Telecommunication Sector

Recommendation ITU-T P.1311

Method for determining the intelligibility of multiple concurrent talkers

1 Scope

Modern immersive telemeeting systems have enabled remote meeting experiences that begin to resemble in-person meetings, with rapid turn-taking and frequent occurrences of multiple people speaking at once. The shift to these natural behaviours is possible because telemeeting systems are increasingly capable of delivering contextual cues that allow interlocutors to stay oriented in a complex virtual meeting environment, much as they do in in-person meetings.

Telemeeting systems also are under pressure to allow access from a variety of endpoints and environments, including mobile endpoints, desks, home offices, conference rooms and dedicated telemeeting rooms, and to interconnect users serviced by different networks. All of these endpoints and networks potentially have different capabilities to capture, carry, and reproduce the contextual cues that make engaging and interactive conferences possible. The industry is in need of a standard method to quantify the intelligibility of multiple concurrent talkers that these systems provide and this Recommendation defines such a method.

This Recommendation¹ describes a method for obtaining an objective measure of how well a telemeeting system allows users to follow a conversation when talk spurts of several talkers coincide. The method comprises a listening-only test that involves listeners observing several concurrent talkers, identifying one of them, and reporting what that talker said. The method is applicable to audio and audiovisual telemeeting systems with three or more users in two or more locations. The test outcome is determined by the degree to which the system under test (SuT) preserves the independence of voices and limits perceptual interference between voices. For example, in a system that spatially virtualizes different talkers, the method is sensitive to changes in the perceived angular separation of talkers and can be used to differentiate between alternative implementations of system components such as sound field capturing microphones or virtual spatial auditory displays. Implementing the method does not require access to internal components of the system under test, so that the method of this Recommendation can be used to test both system components and deployed, live systems. Practitioners should be aware that telemeeting systems can and should be evaluated on several metrics. Tests conducted using the method of this Recommendation provide an objective measure of how well users follow a conversation during times of concurrent talk. This is an important aspect of meetings, but it is not the only aspect that determines the acceptance of and user satisfaction with a telemeeting system. Practitioners must evaluate which other aspects, if any, are important for a complete evaluation of a system under test and are referred to [ITU-T P.1301] and other Recommendations currently in preparation for guidance on selecting measurement procedures. Practitioners should further be aware that the method specified in this Recommendation is applicable only to the evaluation of two or more concurrent talkers; it is not applicable to the assessment of intelligibility under conditions where only one talker is active at any one time. To evaluate the intelligibility of a single talker over a communication system practitioners are referred to [b-ANSI S3.2 (R2014)] In particular, the method does not aim at any representativeness of the phonetic representation of a chosen language.

¹ This Recommendation includes an electronic attachment containing speech samples.

2 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.

- [ITU-T P.58] Recommendation ITU-T P.58 (2013), *Head and torso simulator for telephonometry*.
- [ITU-T P.800] Recommendation ITU-T P.800 (1996), *Methods for subjective determination of transmission quality*.
- [ITU-T P.832] Recommendation ITU-T P.832 (2000), *Subjective performance evaluation of hands-free terminals*.
- [ITU-T P.1301] Recommendation ITU-T P.1301 (2012), *Subjective quality evaluation of audio and audiovisual multiparty telemeetings*.

3 Definitions

3.1 Terms defined elsewhere

This Recommendation uses the following terms defined elsewhere:

- 3.1.1 double talk** [ITU-T P.832]: When near-end and far-end speech occur simultaneously at a given point, typically the terminal under test.
- 3.1.2 diffuse field frequency response of HATS (sound pick-up)** [ITU-T P.58]: Difference, in dB, between the third-octave spectrum level of the acoustic pressure at the ear-drum reference point (DRP) and the third-octave spectrum level of the acoustic pressure at the HATS reference point (HRP) in a diffuse sound field with the HATS absent.

3.2 Terms defined in this Recommendation

This Recommendation defines the following terms:

- 3.2.1 concurrent talk:** When far-end speech of two or more talkers occurs simultaneously at a given point, typically the near end terminal.
- 3.2.2 seat:** An input to a conferencing system associated with a user. A seat is characterized by the signal path from the mouth reference point (MRP) of the talker in the seat to the listener.
- 3.2.3 talker:** A unique individual whose utterances are used in a listening test.
- 3.2.4 call sign:** A two-syllable word or phrase used to address a listener.
- 3.2.5 target:** The talker who utters the call sign the listener is listening for and the colour and number combination the listener is expected to report. May also refer to the seat the talker is in.
- 3.2.6 interferer:** Any talker that is not the target. May also refer to the seats these talkers are in.
- 3.2.7 condition:** A unique instantiation of a conferencing system, including the system under test, the seats, and environmental factors but excluding talkers and listeners.
- 3.2.8 trial:** The smallest unit of the test, including stimulus presentation, response collection, and feedback.
- 3.2.9 block:** A sequence of trials characterized by the constancy of condition and talkers.