

International Telecommunication Union

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

P.807

(02/2016)

SERIES P: TERMINALS AND SUBJECTIVE AND
OBJECTIVE ASSESSMENT METHODS

Methods for objective and subjective assessment of
speech quality

Subjective test methodology for assessing speech intelligibility

Recommendation ITU-T P.807



International
Telecommunication
Union

Documents
Document Sharing Hub

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.807

Subjective test methodology for assessing speech intelligibility

Summary

Recommendation ITU-T P.807 describes a subjective testing methodology for assessing speech intelligibility in communications settings, systems and devices. The method provides a percent correct intelligibility score based on a two-alternative, forced-choice task where the stimulus is one of the two words from a pair of words, i.e., a test item. Half of the test items are rhyming word-pairs (i.e., they differ only in the initial consonant) and half are alliterative word-pairs (i.e., they differ only in the final consonant). The two critical consonants in each test item differ only in a single distinctive feature (see Annex A for a description of distinctive features). In addition to a score for overall intelligibility, the method provides scores for each of six distinctive features: voicing, nasality, sustention, sibilation, graveness and compactness. These scores may be used to diagnose the specific cause of impairments leading to degradation of speech intelligibility.

History

Edition	Recommendation	Approval	Study Group	Unique ID*
1.0	ITU-T P.807	2016-02-29	12	11.1002/1000/12750

Keywords

Diagnostic assessment of intelligibility, distinctive features, speech intelligibility testing.

* 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 2016

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.....	1
3 Definitions	1
3.1 Terms defined elsewhere	1
3.2 Terms defined in this Recommendation.....	1
4 Abbreviations and acronyms	2
5 Conventions	2
6 Description of the ITU-T P.807 testing methodology	2
6.1 Need for an ITU-T intelligibility testing method	2
6.2 Rhyme tests	3
6.3 Intelligibility test method – ITU-T P.807	4
6.4 ITU-T P.807 test results	6
Annex A – ITU-T P.807 and distinctive features	12
Appendix I – Example instructions for the ITU-T P.807 test.....	14
Bibliography.....	17

.....

Recommendation ITU-T P.807

Subjective test methodology for assessing speech intelligibility

1 Scope

This Recommendation describes a subjective testing methodology for assessing speech intelligibility in communications settings, systems and devices. The ITU-T P.807 test methodology has been tested and was found to be appropriate for assessing speech intelligibility of telecommunications systems (e.g., speech codecs), including channel impairments and background noise conditions and for terminals and devices (e.g., handsets, intercom systems). It is designed for use with naive subjects and requires little in the way of specialized equipment or software. The method applies to word intelligibility and does not address intelligibility of phrases or sentences. The method described here applies to North American English. However, the method could be adapted to other languages or dialects taking into account the appropriate set of distinctive features for the language.

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.800] Recommendation ITU-T P.800 (1996), *Methods for subjective determination of transmission quality*.

3 Definitions

3.1 Terms defined elsewhere

None.

3.2 Terms defined in this Recommendation

This Recommendation defines the following terms:

3.2.1 compactness: Distinctive feature that distinguishes compact phonemes from diffuse phonemes. Compact phonemes are produced by constriction toward the rear of the vocal tract; diffuse phonemes by constriction near the middle. Compact phonemes are characterized by the concentration of spectral energy in the mid-frequency range; diffuse phonemes by the distribution of energy over more-widely separated spectral peaks.

3.2.2 graveness: Distinctive feature that distinguishes grave phonemes from acute phonemes. Grave phonemes are produced by constriction toward the anterior of the vocal tract; acute by constriction in the middle of the tract. Grave phonemes are distinguished among other things by the origin and direction of second-formant transitions. Grave consonants always involve relatively steep upward transitions of the second formant. Acute consonants usually involve downward second-formant transitions, depending on vowel environment and the phoneme involved. In general, grave phonemes are characterized by greater concentration of low-frequency spectral energy than are acute phonemes.

3.2.3 nasality: Distinctive feature that distinguishes nasal phonemes from non-nasal phonemes. Nasals are produced by lowering of the velum, allowing air to escape through the nasal passages; non-nasals by closing the nasal passages. Nasal phonemes are distinguished by relatively pronounced resonances at *circa* 200, 800, and 2200 Hz and by the presence of nulls throughout the frequency range.

3.2.4 sibilant: Distinctive feature that distinguishes sibilant phonemes from non-sibilant phonemes. Sibilants involve extreme constriction of the vocal tract that produces turbulence and high-frequency noise. Sibilant consonants are characterized by higher-frequency noise and greater duration than their non-sibilant counterparts.

3.2.5 sustention: Distinctive feature that distinguishes sustained phonemes from interrupted phonemes. Sustained phonemes are produced by incomplete constriction of the vocal tract; interrupted phonemes by complete constriction of the tract at some point. Sustained phonemes are distinguished by their gradual onset and by the presence of mid-frequency noise, interrupted by their abrupt onset. Sustained phonemes have characteristic durational and high-frequency cues that distinguish them from their interrupted counterparts.

3.2.6 voicing: Distinctive feature that distinguishes voiced phonemes from unvoiced phonemes. Voiced phonemes involve free vibration of the vocal cords; unvoiced phonemes do not. Voiced phonemes are distinguished from their unvoiced counterparts, or cognates, by the presence of periodicity, and, in particular, by the time of onset in periodicity. In voiced consonants, preceding vowels tend to be of greater duration than in the case of unvoiced consonants.

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

2AFC	2-Alternative, Forced-Choice
ACR	Absolute Category Rating
AMR-WB	Adaptive Multirate Codec – Wideband
ANSI	American National Standards Institute
DALT	Diagnostic Alliteration Test
DRT	Diagnostic Rhyme Test
HATS	Head And Torso Simulator
IPA	International Phonetic Alphabet
MOS	Mean Opinion Score
MRT	Modified Rhyme Test
RF	Radio Frequency
SNR	Signal to Noise Ratio

5 Conventions

None.

6 Description of the ITU-T P.807 testing methodology

6.1 Need for an ITU-T intelligibility testing method

In recent years, there has been increased interest in testing systems or devices for intelligibility. This is especially relevant for "speech enhancement" techniques and algorithms, e.g., noise