

I n t e r n a t i o n a l T e l e c o m m u n i c a t i o n U n i o n

ITU-T P.830 Implementer's Guide

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

(28 March 2013)

SERIES P: TERMINALS AND SUBJECTIVE AND
OBJECTIVE ASSESSMENT METHODS
Methods for objective and subjective assessment of
speech quality

**Implementer's Guide for P.830 (Subjective
performance assessment of telephone-band and
wideband digital codecs)**

ITU-T

Contact Information

ITU-T Study Group 12 /
Question 7/12 Rapporteurs

Alan Sharpley
Dynastat Inc.
USA

Paolo Usai
ETSI
France

Tel: + 1 512 554 2732
Fax: + 1 512 472 2883
Email: asharpley@dynastat.com

Tel: + 33 4 92 94 42 36
Fax: + 33 4 93 65 28 17
Email: paolo.usai@etsi.org

Revised Implementer's Guide for Recommendation ITU-T P.830: Subjective performance assessment of telephone-band and wideband digital codecs

1 Introduction

Nowadays the terms of reference often asks for evaluating the quality of codecs under development in presence of music even when the codec is only capable of narrowband or wideband. The foreseen application is mainly music on hold.

Some test plans were drafted and testing was performed. In all cases, Absolute Category Rating was chosen as methodology and to ensure better reliability it was decided to increase the number of votes.

One practical way to increase the number of votes was to run two different randomization orders per subject. Following this way, the test results presented acceptable standard deviation as expected.

In case two laboratories were involved in the testing exercise, possible discrepancies between different test labs and databases may still appear. In order to lower the differences coming from the use of different music items, another way to increase the number of votes was to run the test with both databases, a group of subject run the experiment with the database of the listening lab following the randomization order and then run the same with the database of the other listening lab.

2 Proposal

It is proposed to introduce at the end of clause 8.1.10 of Recommendation ITU-T P.830, the following section:

When performing a quality evaluation of speech codec with regards to music, the ACR method is recommended. When only one laboratory is involved in the testing exercise, each group of subjects runs two randomization orders. When two laboratories are involved, each group of subjects runs the test once with the database of their own laboratory and once with the database of the other laboratory. In case two laboratories are involved, care should be taken to avoid selecting the same music samples or samples that may not be appropriate for the other laboratory (e.g., due to differences in language or culture).

3 On reference Signals

The terms of reference often ask for evaluating the quality of codecs under development capable of narrow band, wideband, super-wideband and also fullband. When using super-wideband or fullband signals, P.810 MNRU is not appropriate because speech signals present a low level energy in the high frequency domain. To avoid artifacts, a modified MNRU using P.50 spectrally shaped noise is used.

It is proposed to introduce at the end of the clause 8.2.2 of Recommendation ITU-T P.830 the following sentences:

For super-wideband testing, when using DCR methodology, a modified MNRU using P.50 spectrally shaped noise is used. It is recommended to span the range of Q values from 5 to 45 dB at least with 5 to 7 different values of Q.

A software tool to construct a modified MNRU using P.50 spectrally shaped noise is to be made available in ITU-T Recommendation G.191.