INTERNATIONAL TELECOMMUNICATION UNION

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

P.831

(12/98)

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

Methods for objective and subjective assessment of quality

Subjective performance evaluation of network echo cancellers

ITU-T Recommendation P.831

(Previously CCITT Recommendation)

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

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

ITU-T RECOMMENDATION P.831

SUBJECTIVE PERFORMANCE EVALUATION OF NETWORK ECHO CANCELLERS

Summary

This Recommendation describes methods and procedures for conducting subjective performance evaluations of network echo cancellers.

The deployment of digital technology in the Public Switched Telephone Network (PSTN) has had numerous advantages for users of the network as well as for network operators. These new technologies come at the price of increased transmission time, which increases the likelihood that any echo impairment will be annoying to voice users of the network. Hence, the deployment of echo cancellers in the network is widespread. Recommendations G.165 and G.168 define certain instrumental tests that must be met to ensure minimum performance of an echo canceller. However, there has been some concern that those tests do not address fully the echo cancellation needs of voice users of the network.

Subjective testing is a commonly used method of assessing the performance of digital devices, including digital speech codecs and Digital Circuit Multiplication Equipment (DCME). This Recommendation defines natural extensions of those techniques to the subjective evaluation of echo cancellers.

Source

ITU-T Recommendation P.831 was prepared by ITU-T Study Group 12 (1997-2000) and was approved under the WTSC Resolution No. 1 procedure on the 3rd of December 1998.

Keywords

Echo cancellation, subjective performance, speech transmission quality.

FOREWORD

ITU (International Telecommunication Union) is the United Nations Specialized Agency in the field of telecommunications. The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of the ITU. The 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 Conference (WTSC), which meets every four years, establishes the topics for study by the ITU-T Study Groups which, in their turn, produce Recommendations on these topics.

The approval of Recommendations by the Members of the ITU-T is covered by the procedure laid down in WTSC Resolution No. 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 term *recognized operating agency (ROA)* includes any individual, company, corporation or governmental organization that operates a public correspondence service. The terms *Administration, ROA* and *public correspondence* are defined in the *Constitution of the ITU (Geneva, 1992)*.

INTELLECTUAL PROPERTY RIGHTS

The ITU draws attention to the possibility that the practice or implementation of this Recommendation may involve the use of a claimed Intellectual Property Right. The 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, the 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.

CONTENTS

Recommendation P.831

SUBJECTIVE PERFORMANCE EVALUATION OF NETWORK ECHO CANCELLERS

(Geneva, 1998)

1 General

1.1 Scope

This Recommendation describes procedures to be used to assess the subjective performance of echo cancellers. The methods defined here may be used to assess the extent to which an echo canceller operates effectively for voice users of the PSTN. In particular, the intent is not to define methods that may be used to assess the effects of delay, nor is it the intent to define rules for echo canceller application. These issues are addressed in Recommendations G.114 and G.131, respectively. Further, this Recommendation does not define specific values for echo canceller parameters (e.g. convergence time) to yield satisfactory subjective performance.

The procedures defined here may also be appropriate for evaluating the subjective performance of other signal processing devices that may be deployed in the PSTN (e.g. Automatic Level Control devices). These issues are under study in ITU–T Study Group 12.

1.2 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; all users of this Recommendation are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations listed below. A list of the currently valid ITU-T Recommendations is regularly published.

– CCITT Recommendation G.164 (1988), *Echo suppressors*.

– ITU–T Recommendation G.165 (1993), *Echo cancellers.*

– ITU–T Recommendation G.168 (1997), *Digital network echo cancellers.*

– ITU–T Recommendation P.51 (1996), *Artificial mouth*.

– ITU–T Recommendation P.56 (1993), *Objective measurement of active speech level.*

– ITU–T Recommendation P.57 (1996), *Artificial ears*.

– ITU–T Recommendation P.58 (1996), *Head and torso simulator for telephonometry*.

– ITU–T Recommendation P.501 (1996), *Test signals for use in telephonometry*.

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

– ITU–T Recommendation P.810 (1996), *Modulated Noise Reference Unit (MNRU).*

– ITU–T Recommendation P.830 (1996), *Subjective performance assessment of telephone-band and wideband digital codecs*.

– ITU *Handbook on Telephonometry*, 2nd edition, Geneva 1992.

1.3 Terms and definitions

This Recommendation defines the following terms:

1.3.1 double-talk: When near-end and far-end speech occur simultaneously.

1.3.2 far-end: The side of an echo canceller which does not contain the echo path on which the echo canceller is intended to operate.

1.3.3 live network: A telephony network in commercial use.

1.3.4 near-end: The side of an echo canceller which contains the echo. This includes all transmission facilities and equipment (including the hybrid and terminating telephone) which is included in the echo path.

1.3.5 syllable clipping or temporal clipping: Loss of speech energy caused by voice/speech activated devices. For echo cancellers, the primary source of temporal clipping is the NLP. In this instance, clipping does not refer to amplitude limiting.

1.3.6 talker echo loudness rating: The loudness loss in the talker echo path. (see Recommendation G.100)

1.3.7 third-party listening test: A listening-only subjective test (see Recommendation P.800) in which the listener hears recordings from the "center" of the connection under evaluation. In conventional listening-only tests, the listener is positioned at one end of the connection under study.

1.3.8 conversation test: A subjective test in which two participants have a conversation, as described in Annex A/P.800 and in the *Handbook on Telephonometry*.

1.3.9 talking-and-listening test: A subjective test in which a participant talks while simultaneously listening for impairments (e.g. echo).

1.3.10 untrained subject: See 3.3.1.

1.3.11 experienced subject: See 3.3.2.

1.4 Abbreviations

This Recommendation uses the following abbreviations:

ACR	Absolute Category Rating
ATM	Asynchronous Transfer Mode
DCME	Digital Circuit Multiplication Equipment
DCR	Degradation Category Rating
DMOS	Degradation Mean Opinion Score
EC	Echo Canceller
ERL	Echo Return Loss
ERLE	Echo Return Loss Enhancement
HATS	Head and Torso Simulator
LRGP	Lip Ring Guard Position
MNRU	Modulated Noise Reference Unit
MOS	Mean Opinion Score
MRP	Mouth Reference Point
NLP	Non-Linear Processor
OLR	Overall Loudness Rating

PCME Packetized Circuit Multiplication Equipment

PLMN Public Land Mobile Network

PSTN Public Switched Telephone Network

RLR Receiving Loudness Rating

SLR Sending Loudness Rating

TELR Talker Echo Loudness Rating

1.5 Conventions

As described in this Recommendation, subjective evaluation of echo cancellers may be conducted using listening-only, talking-and-listening, or conversational methods. The test procedures defined in this Recommendation may also be useful for evaluation of other active speech signal processing devices.

2 Overview of subjective testing procedures for Echo Cancellers

Table 1 lists four subjective testing methods that have been found suitable for evaluating the subjective performance of Echo Cancellers (ECs). Each testing method is described briefly in this clause. Detailed descriptions are found in other clauses.

Each listing in Table 1 includes the possible applications of that type of test. Four classes for application of the methods are identified:

- evaluation of overall opinion and/or quality of connections with ECs from the viewpoint of a typical voice user of the PSTN;
- identification of EC parameters that are important for voice users of the PSTN (e.g. tandem operation of ECs, echo return loss enhancement);
- selection of values for those parameters (e.g. speed of convergence, minimum amount of echo return loss enhancement);
- diagnostic evaluation of specific problems involving ECs.

Each of the testing methods may be used with untrained or experienced participants. However, each procedure may have different application depending on the amount of technical experience a particular participant has with ECs. For example, conversational tests with untrained participants would be especially useful for "global" evaluation of EC performance. On the other hand, a conversational test between experts might serve as a diagnostic tool.

In general, it is not recommended that Talking-and-Listening Tests, and Listening-Only Tests be performed in isolation. A complete evaluation of echo canceller performance must take into account conversational interactions between subjects.

Table 1/P.831 – Applications of subjective testing methods

Testing method	Untrained subjects	Experienced subjects
Conversational Test	Overall Opinion and/or Quality Parameter Value Selection Parameter Identification	Diagnostic Parameter Value Selection Parameter Identification
Talking-and-Listening Test	Parameter Value Selection Diagnostic	Diagnostic Parameter Value Selection Parameter Identification
Third-Party Listening Type A	Diagnostic Overall Opinion and/or Quality Parameter Value Selection Parameter Identification	Diagnostic Parameter Value Selection Parameter Identification
Third-Party Listening Type B	Diagnostic Overall Opinion and/or Quality Parameter Value Selection Parameter Identification	Diagnostic Parameter Value Selection Parameter Identification
NOTE – • Third-party listening Type A – see 2.3, below. • Third-party listening Type B – see 2.3, below. • In order to use the suite of test methods effectively, it will be helpful to understand the relationship among the various types of tests (e.g. MOS on conversational vs. MOS on Listening Type A). This topic is currently under study.		

2.1 Conversational test

Conversational test procedures for ECs are described in clause 4. Conversational tests have the advantage that they allow ECs to be evaluated under somewhat realistic conditions. On the other hand, conversational tests are time consuming, hence expensive, to run. Additionally, while such tests can be arranged to elicit episodes of double-talk, the number and duration of these episodes are hard to control.

2.2 Talking-and-listening test

Talking-and-listening tests are described in clause 5. Talking-and-listening tests were designed to focus on the initial part of a telephone call. They are relatively easy to run (as compared to conversational tests) and have the advantage that they focus on such important parameters as initial convergence of the EC.

2.3 Third-Party listening tests

In a conventional listening test (as described in Recommendation P.800), the listening point is at one end of the connection under study (the receiving end). Furthermore, the listener hears the signals as if he or she were actually participating in a conversation. Third-party listening tests differ from conventional listening tests in that the listener may hear signals from both endpoints of the connection. In particular, the third-party listening tests described here put the listener in the logical position of the talker. Since the subject (i.e. the listener in the listening test) will not, in general, be the actual person talking, the subject actually has the role of a third party who is "listening in" on a conversation. Two types of third-party listening tests are described in this Recommendation:

- Third-party listening Test A – uses recordings made with HATS (according to Recommendation P.58), one at each end of the connection.
- Third-party listening test B – similar to third-party listening Test A, but no HATS are used.

3 General considerations for subjective evaluation of echo cancellers

Unless otherwise noted, the general considerations described in this clause apply to each of the test methods described in 4.2.

3.1 Echo canceller parameters to evaluate

An example set of test conditions is provided in Appendix I.

Table 2 lists EC parameters that should be considered when evaluating echo canceller performance. For each parameter, the test conditions that should be evaluated are also shown. The column labelled "Test Conditions" contains a number of simple entries that are abbreviations for more complicated issues. A brief elaboration of these entries is as follows:

- *Background audio*
  - level;
  - type (car, babble, highly dynamic noise, etc.);
  - circuit noise;
  - injected noise.
- *Echo path circuit*
  - delay (possibly exceeding the tail capacity of the EC);
  - multiple echo paths;
  - frequency response;
  - unequal send and receive levels;
  - amplitude variation (due to level control in the network);
  - DCME/PCME (e.g. Comfort Noise Generation, Speech Coding);
  - Echo Return Loss (at hybrid);
  - conference bridge;
  - residual acoustic echo;
  - tandem Ecs;
  - mobile systems;
  - multiple hybrids;
  - non-linear tail circuit:
    - time variation:
      - continuous ("phase roll");
      - instantaneous large change;
    - low bit-rate coding;
    - PCM offset;
    - ATM;
    - delay variation as found in Internet telephony.

- *Tandem ECs*
  - low bit-rate codec between tandem Ecs.
- *Other*
  - "mixed" voice/DTMF call (voice mail, etc.);
  - call waiting/on-hold ("leakage").

It is neither necessary nor desirable to evaluate all of these parameters in a single subjective test. It is suggested that preliminary evaluation of an EC, using experts and/or laboratory personnel, be used to identify the kinds of problems that should be evaluated in a given subjective test. If the number of conditions remains large, multiple evaluations should be conducted.

The last column in Table 2 ("Type of test") is intended to show the test procedure(s) that would be appropriate for evaluating a given parameter (under the appropriate test conditions).

Table 2/P.831 – EC parameter and test conditions

Parameter to examine	Test conditions	Type of test
Impairments during Double-Talk	Level Differences Background Noise Talker Sequence Near-end Start Tandem Ecs Echo path circuit	Conversational Third-Party Listening Test A Third-Party Listening Test B
Impairments during Single Talk	Background Noise Echo path circuit Tandem Ecs	Conversational Third-Party Listening Test A Third-Party Listening Test B
Initial Convergence	Changing Echo path circuit Echo path change Near-end Start	Talking-and-Listening Conversational Third-Party Listening Test A Third-Party Listening Test B
Divergence	Level Differences Background Noise Talker Sequence Near-end Start Tandem Ecs Echo path circuit	Conversational Third-Party Listening Test A Third-Party Listening Test B
Background Noise and Comfort Noise Generation (noise pumping, etc.)	Level Differences Background Noise Talker Sequence Near-end Start Tandem Ecs Echo path circuit	Conversational Third-Party Listening Test A Third-Party Listening Test B

3.2 Test equipment and calibration

Selection of test equipment, and calibration of the equipment, will depend on the objectives of the test. It is, therefore, difficult to provide comprehensive guidance on these issues. However, those conducting subjective evaluations of ECs should pay particular attention to the following:

- TELR in the absence of a functioning EC;
- delay in the "network" segment and in the tail of the EC;

- hybrid characteristics such as return loss and impulse response (dispersion and frequency response);
- loudness ratings and frequency responses of telephone sets;
- speech levels;
- speech material for third-party listening tests;
- clearing the H-register:
  - before each trial in a conversational test;
  - before each test recording in third-party listening tests.

3.3 Selection of subjects

Some care should be taken when selecting subjects for evaluation of ECs. As with other speech signal processing equipment (e.g. speech codecs, DCME, etc.), some potential subjects will be more experienced than others. It is recognized that experience with ECs is a continuum ranging from those who are completely unfamiliar with EC operation ("non-experts") to those who are thoroughly conversant in the operation and maintenance of ECs ("experts"), such as EC designers. However, it is convenient to refer to two parts of this continuum: untrained subjects and experienced subjects.

3.3.1 untrained subjects: Untrained subjects are accustomed to daily use of a telephone. However, they are neither experienced in subjective testing nor are they experts in technical implementations of ECs. Ideally, they have no specific knowledge about the device that they will be evaluating.

3.3.2 experienced subjects: Experienced subjects (for the purpose of EC evaluation) are experienced in subjective testing, but do not include individuals who routinely conduct subjective evaluations. Experienced subjects are able to describe an auditory event in detail and are able to separate different events based on specific impairments. They are able to describe their subjective impressions in detail. However, experienced subjects neither have a background in technical implementations of ECs nor do they have detailed knowledge of the influence of particular EC implementations on subjective quality.

3.4 Analysis of results

Test results should be evaluated using standard statistical procedures, as noted in Recommendation P.800 and the *Handbook on Telephonometry*.

4 Conversation tests

Conversation-opinion test procedures are described in Annex A/P.800 and in the *Handbook on Telephonometry*. Details are not repeated here. However, some considerations for use of conversation tests to evaluate the subjective performance of ECs is provided.

4.1 Purpose

A conversation test involves two parties conversing over a connection. Depending on the purpose of the test, either experienced or untrained subjects can be used. Such tests can be useful to both manufacturers and operators, and are an important assessment tool because they provide the closest simulation of real telephone interactions between customers. The purpose of conversational testing will be different depending on whether experienced or untrained subjects are used. The differences are highlighted in the table below:

Untrained subjects	Experienced subjects
Overall opinion/quality, and difficulty	Diagnosis
Parameter Identification	Parameter Identification
Parameter Value Selection	Parameter Value Selection
	Choice of test conditions

Untrained subjects are used when it is important to obtain an indication of how the general population of telephone users would rate the overall quality and difficulty in using a connection. This can be used to give a "global" evaluation of echo canceller performance in a range of connections. However, untrained subjects are unable to describe and identify accurately the types of degradation associated with echo cancellers. Experienced subjects are therefore used in the following situations where it is necessary to obtain information about the subjective effects of individual degradations:

1) diagnosis of echo canceller problems;
2) identification of individual echo canceller parameters such as convergence time;
3) establishment of sensible echo canceller parameter values;
4) helping to choose suitable conditions for inclusion in a test to be performed by untrained subjects.

4.1.1 Advantages

The benefit of conversational testing is that it is the only way of realistically assessing the combined subjective effect of all the parameters affecting conversational quality. In particular, effects such as delay, echo and double-talk can have a marked effect on echo canceller performance.

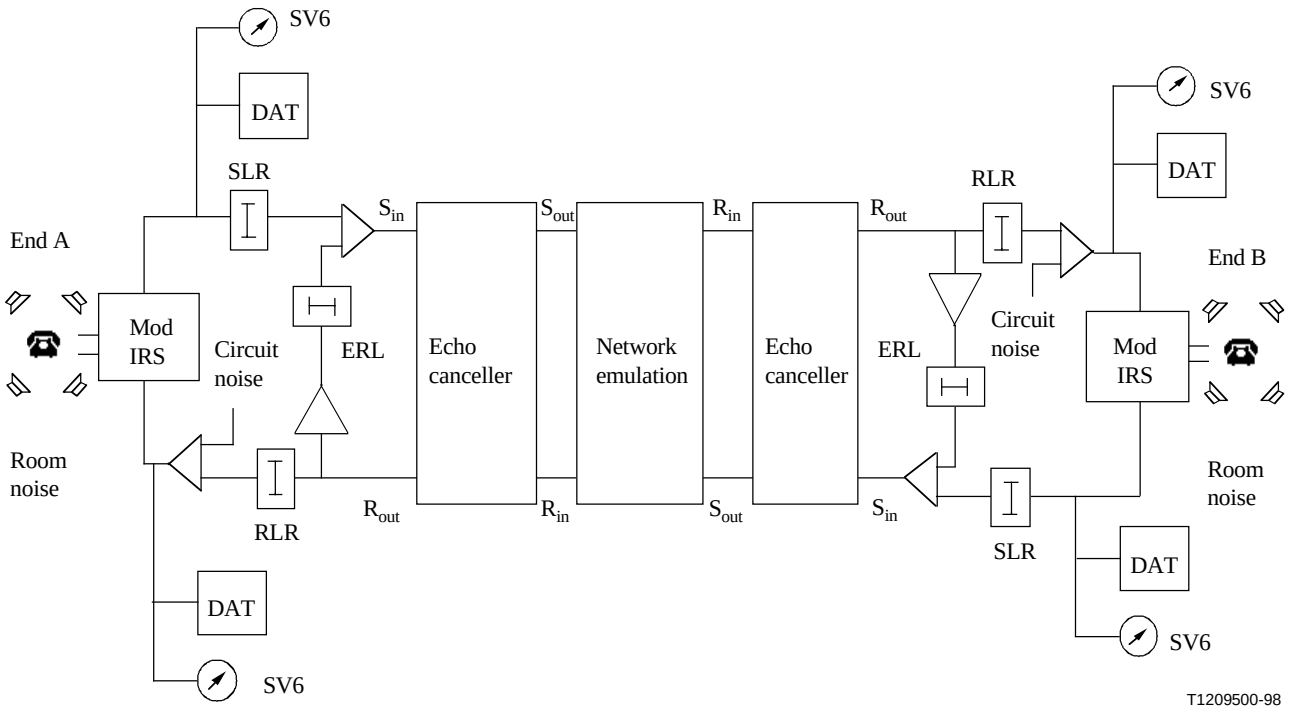
4.1.2 Disadvantages

The disadvantage of conversational testing is that it is time-consuming, and therefore expensive, compared with other forms of testing. The number of conditions that can be realistically tested in one experiment is limited because of the time required for typical conversations. It can also be quite complex to set up initially because of the need for full duplex operation. Studying echo canceller performance under double-talk conditions requires some strategy to force the two parties to speak simultaneously. While strategies for accomplishing this exist, controlling the number and duration of the double-talk episodes is difficult with the conversational test.

4.2 Test design, set-up and procedure

The general test design, set-up and procedure for full conversation tests are described in Recommendation P.800 and in the *Handbook on Telephonometry*, which should be consulted for further detail. The *Handbook on Telephonometry* also gives some guidance on "simplified conversation tests", where short cuts are suggested to reduce the time taken or to increase the number of treatments in one experiment. Some work has been done with a variation on simplified tests, where experts are asked to rate a number of individual degradations after they have given their opinions on quality and difficulty. These can be useful for diagnostic applications, but further work is necessary to provide correlation with subjective opinions of quality and difficulty gathered from full conversation tests and objective measurements. Some particular considerations for designing full echo canceller conversation tests are listed in subsequent clauses.

A circuit diagram of an example conversation test using ECs is shown in Figure 1.



SLR Sending Loudness Rating
RLR Receiving Loudness Rating
ERL Echo Return Loss
Mod IRS Modified Intermediate Reference System
DAT Digital Audio Tape machine
SV6 Speech Voltmeter (Rec. P.56)

Figure 1/P.831 – Example circuit for a conversation test

4.2.1 Test design and circuit conditions

The test should be designed with a range of good and bad conditions to ensure that the full opinion scale is used. Circuit conditions should be chosen to exercise the EC adequately and to cover the situations where it is likely to be deployed. See 3.1 for a list of possible parameters to investigate. Subclause 3.2 should be consulted for guidance on equipment calibration.

4.2.2 Reference conditions

Specific reference conditions for use in subjective evaluation of ECs is under study in ITU–T Study Group 12. General guidance on use of reference conditions is found in Recommendations P.800 and P.830. Reference conditions should be included so that tests performed on different echo cancellers at different times and by different test laboratories may be compared. Such reference conditions may include test set-ups without an echo canceller, but with well-defined residual echo levels (achieved by varying the echo path attenuation in steps) and other parameters.

4.2.3 Task

Different conversational tasks have been used by different Administrations, including one where subjects are asked to reach an agreement on an order of preference for a set of picture postcards (as described in the *Handbook on Telephonometry*). Another task has also been used where subjects are asked to describe to their partner the position of a set of numbers on a picture. Both subjects have similar pictures, but with some of the numbers in different positions. It is recommended that the picture should be designed for the task and that both the picture and the numbers are easy to describe. This can be achieved by using pictures consisting of coloured, geometrical figures (e.g. paintings by Kandinsky or others).

The primary considerations for choice of task are to ensure that it leads to a clear conclusion of the conversation, that the two participants are approximately equally active in the conversation (i.e., the conversation is not too one-sided), and that a reasonable range of vocabulary is used. In addition, it is important for echo canceller testing that the task leads to conversations where a realistic number of double-talk situations are generated. A specific value for the percentage of a conversation that is double-talk cannot be recommended at this time (since it almost certainly varies for different languages and cultures). This issue is under study in ITU–T Study Group 12.

4.2.4 Opinion scales and questionnaires

In a typical conversational test, the participants answer the following questions after each conversation:

What is your opinion of the connection you have just been using?

5 Excellent

4 Good

3 Fair

2 Poor

1 Bad

Did you or your partner have any difficulty in talking or hearing over the connection?

Yes

No

Further details on these scales are given in Recommendation P.800 and in the *Handbook on Telephonometry*. If either subject experienced difficulty, then they may be asked to describe the nature of the difficulty.

Some Administrations have found it useful to use a more detailed questionnaire when evaluating subjective performance of ECs. One such questionnaire is described in Annex A.

5 Talking-and-listening test

As the name of this test suggests, a single subject must talk and listen simultaneously. This subject has the role of the far-end subscriber. There is no near-end subscriber during the test. The near-end can be simulated by different echo path realizations, by electrical injection of the background noise, or by different terminal equipment.

5.1 Purpose

This test procedure is specially designed to evaluate talking-related disturbances.

If the lack of a complete conversation can be tolerated, the subjective performance of echo cancellers can be investigated in an efficient way with talking-and-listening tests. All aspects of EC function that influence the transmission quality for subscribers while they are either talking-and-listening or only listening (without having a conversational partner on the other end of the connection) are covered by this procedure.

The test may be performed with either untrained or trained subjects, depending on the purpose of the test as summarized in the table below:

Untrained subjects	Experienced subjects
Diagnostic	Diagnostic
Parameter Value Selection	Parameter Identification
	Parameter Value Selection

5.2 Parameters

Typical parameters to be judged by subjects are:

- disturbances caused by echoes;
- disturbances caused by audible switching;
- quality of background noise transmission.

NOTE – These parameter examples are not necessarily independent of each other. For example, echoes and background noise may be simultaneously interrupted by the non-linear processor. Thus, the switching characteristic affects two completely different parameters, echo and background noise.

5.3 Set-up

A typical test set-up, which has been used extensively for subjective investigations of different echo cancellers, is shown in Figure 2. This measurement set-up shows an echo canceller on each side of the connection. The echo canceller under test is shown on the right in Figure 2, and the terms near-end and far-end are used with respect to this echo canceller.

The near-end subscriber is simulated by a HATS according to Recommendation P.58. If double-talk sequences are required, then the artificial mouth must be calibrated and equalized at the MRP in order to produce the correct sound pressure level at all frequencies. Use of the talking-and listening test for evaluating disturbances during double-talk is discussed further in the Note in 5.5

Environmental conditions will depend on the characteristics of the test room used, background noise and other factors.

NOTE 1 – The simulation of the near-end speaker may be made using an artificial mouth according to Recommendation P.51 and placing the handset in the LRGP position. However, where background noise is also simulated at the near-end, it is recommended that a HATS is used instead of the artificial mouth according to Recommendation P.51.

The test conditions (as indicated in Figure 2) may be changed by an operator or by automatic control between talking-and-listening sequences.

NOTE 2 – If a possible interaction between the near-end terminal and the echo cancellers can be excluded (e.g. no speech controlled devices) and if only the background noise influences the transmission quality, these conditions can be recorded before the test and stored on an appropriate medium. During the tests, these samples can be applied at the echo path.

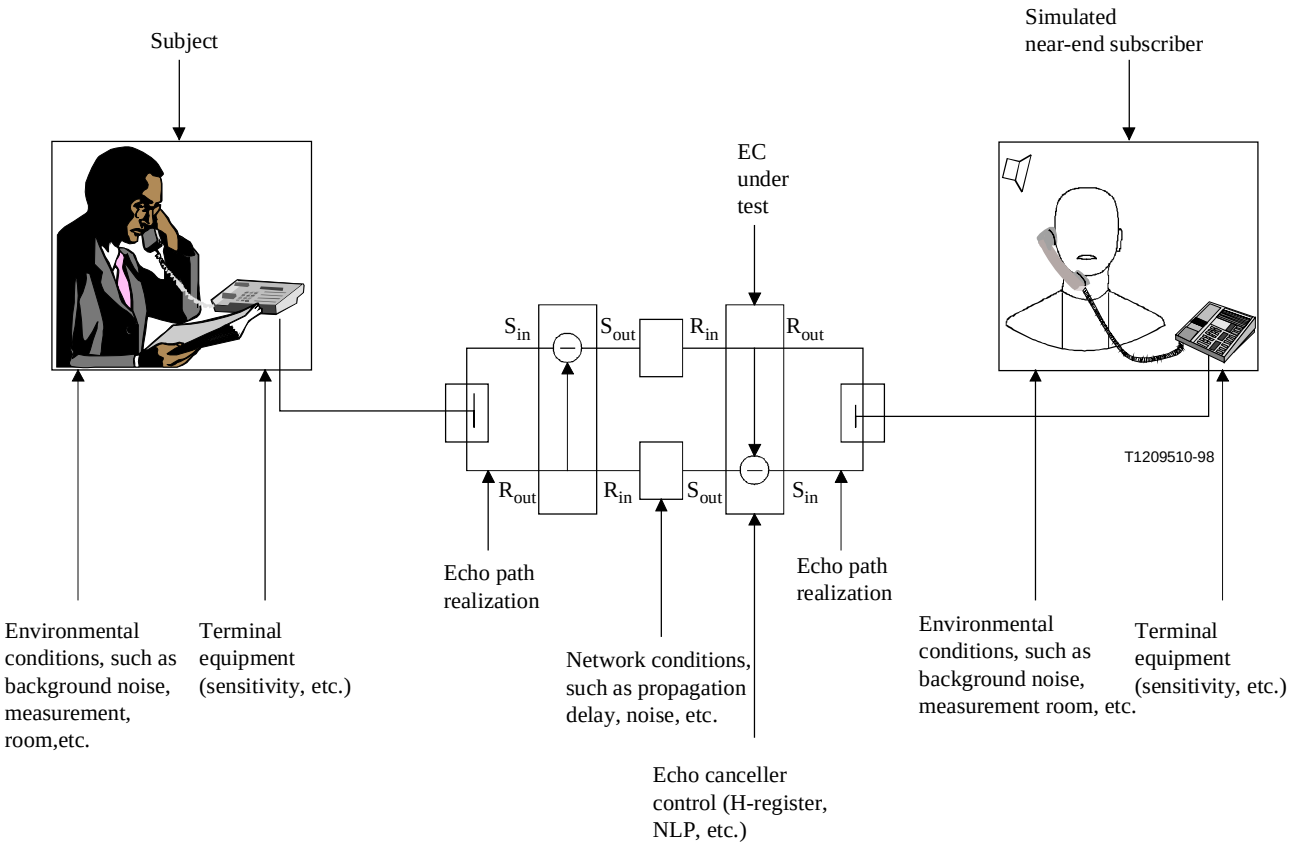


Figure 2/P.831 – Typical measurement setup for talking-and-listening tests

Only one subject takes part during the test, and the near-end subscriber is simulated by a HATS. The environmental conditions and terminal equipment on both sides can be chosen according to the objective of the test. The same applies for the echo path realizations and network conditions. The echo canceller itself can be controlled in order to reset the H-register, enable or disable the NLP or comfort noise features.

5.4 Description of test procedure

Description of the test procedure is separated into two parts: the first part is designed to examine initial convergence, and the second part to measure performance during steady state conditions. Note, however, that which of these procedures is followed will depend on the purpose of the test. Subjects are first instructed on how to conduct the test.

5.4.1 Initial convergence

Before subjects begin talking, the H-register is cleared and adaptation enabled. To reproduce realistic conditions, all subjects should answer an incoming telephone call with the same greeting:

’[company], [name], [greeting]’

After the greeting, the handset should be replaced, and subjects are asked to give their rating.

NOTE – For simplicity, the above description assumes that handsets are used. Other terminal equipment, such as hands-free terminals, may be used. In order to guarantee a natural simulated telephone call, the words in brackets represent typical expressions used in the laboratory carrying out the test.

5.4.2 Steady state conditions

The echo cancellers should first be fully converged. To avoid simple reading (which has some disadvantages), subjects are asked to perform a task, such as to describe the position of given numbers in a picture, which they have in front of them. Such a picture should be specially designed for the task, with numbers distributed in a well-defined way. It is recommended that the picture and the numbers are easy to describe. This can be done by using pictures that consist of coloured, geometrical figures (e.g. pictures from Kandinsky or others). After the handset has been replaced, the subjects are asked to give a rating.

NOTE 1 – If the handset is lifted at the beginning of this test, precautions must be taken to ensure that the echo canceller does not diverge or reset its H-register because of possible clicks or crackling sounds originating at the contact between handset and housing. To prevent such problems, it may be desirable to lift the handset from a smooth, soft pad.

NOTE 2 – The speaking duration can be influenced by controlling the quantity of numbers and the complexity of the picture. Care should be taken to ensure that speaking time is adequate for the needs of the experiment. To avoid speaking durations that are too short, it is recommended that the subjects are asked to describe the position of each number in such a way that a partner at the other end of the connection could identify the exact position in the same picture.

5.5 Advantages

The talking-and-listening test is designed to examine the performance of an echo canceller only during single-talk segments of a conversation. Thus, this test has some advantages compared to complete conversational tests. If parameters which are relevant only during single-talk need to be evaluated, then the procedure requires less time than conversational tests. The tests can be clearly separated into parts (e.g. the evaluation of initial convergence or steady state conditions). In addition, talking-and-listening tests are even better suited than conversational tests for evaluating specific parameters, because subjects can concentrate entirely on these parameters, without leading and following a discussion. The environmental conditions at the near-end can be changed easily during the test. Different ambient background noises can be recorded beforehand and applied at the near-end. In principle, even double-talk performance can be evaluated using these tests. In this case, the simulated speech of the near-end subscriber must be applied using an appropriate source, such as an artificial mouth or a HATS according to ITU-T specifications.

NOTE – The evaluation of double-talk performance with talking-and-listening tests has not been verified. Several difficulties, such as synchronization of the subject on one side and the artificial sources on the other side of the connection, and the missing interaction between the subject and artificial source, must be considered.

5.6 Disadvantages

The test procedure is more artificial than a real discussion between two subjects over a telephone connection. Talking-and-listening tests should not be used in isolation of conversation tests in order to evaluate echo cancellers.

5.7 Reference conditions

Reference conditions should be included so that tests performed on different echo cancellers at different times and by different test laboratories may be compared. Such reference conditions may include test set-ups without an echo canceller, but with well-defined residual echo levels (achieved by varying the echo path attenuation in steps) and other parameters.

NOTE – Reference conditions that include echo cancellers (comparable to MNRU conditions for speech codec tests) should be designed carefully to represent typical quality impairments introduced by echo

cancellers. Such conditions should include modulated background noise (typically caused by non-linear processors such as center clippers) and different echo disturbance simulations, such as switched, interrupted echoes, continuous echoes or time variant echoes (as they typically appear during initial convergence).

5.8 Precautions

Subjects must talk and listen simultaneously without having a conversational partner. The near-end is typically realized by simulated echo paths with different ambient background noises. Therefore, additional stimulation is necessary to encourage subjects to talk. Special precautions should be taken so that subjects are stimulated to talk in a natural way that is as close as possible to a real conversation. Therefore, reading of given sequences cannot be recommended. In addition, subjects should be carefully prepared for this test situation, so that on the one hand they do not expect any interaction from a conversational partner, but on the other hand, they behave in a similar way to a conversational situation.

If the talking-and-listening tests are to be applied to investigations during the initial convergence of echo cancellers, the talking duration must be limited. Otherwise, subjects give their ratings under the impression of steady state conditions, when the echo canceller is fully converged. In addition, this restriction should be the same for all subjects, to ensure reproducible results.

6 Third-party listening Test A

6.1 Purpose

This test procedure is designed to evaluate and compare the individual performance parameters of different echo cancellers, different algorithm implementations or different measurement conditions in one test.

Subjects judge the quality of conversational recordings made between a pair of correctly equalized HATS and reproduced by correctly equalized headphones, as third-party listeners. The test is applicable to situations where the recording procedure needs to reproduce the listening situation as realistically as possible.

The test may be performed with either untrained or trained subjects, depending on the purpose of the test. In either case (untrained or experienced subjects), this test procedure is appropriate for: diagnostic purposes, parameter identification and parameter value selection.

The test can also be used for the generation of a database of processed speech samples of different echo cancellers. Such a database may be used to perform comparisons against new implementations.

The test material simulates a complete or partial conversation using two HATS according to Recommendation P.58, equipped with P.58 artificial mouths and P.57 Type 3.4 pinnae with handset mounting devices (to reproduce the pressure force between handset and ear in accordance with normal use). All types of speech signal degradation can be investigated with this kind of third-party listening test. Specific conversational related parameters such as delay cannot be covered by this test, because this requires a complete conversational test with interaction between the two subjects.

6.2 Parameters

The test may be used to examine the following parameters:

- Under single-talk conditions:
  - disturbances caused by echoes;
  - quality of background noise transmission.

- Under double-talk conditions:
  - disturbances during double-talk;
  - disturbances caused by echoes;
  - disturbances caused by speech gaps (e.g. syllable clipping).

6.3 Set-up

A typical recording set-up, which was used during the development of the test procedure and which has subsequently been used for extensive subjective investigations of different echo cancellers is given in Figures 3 and 4. The recording set-up in Figure 3 shows an echo canceller on each side of the connection. Both subscribers are simulated by HATS according to Recommendation P.58. In addition to the P.58 description, the HATS must be equalized in order to produce the correct signals at the ear of the listener. Appropriate equalizations are: Free-Field equalization (FF), Independent of Direction equalization (ID) and Diffuse Field equalization (DF). Accordingly the headphones used for the reproduction need to be equalized in the same manner (FF, ID, DF). In order to produce the correct sound pressure level at all frequencies, the mouths need to be calibrated and equalized at the MRP.

All parameters (acoustic environment, speech levels, network parameters such as echo path loss and others) can be changed for different recording set-ups.

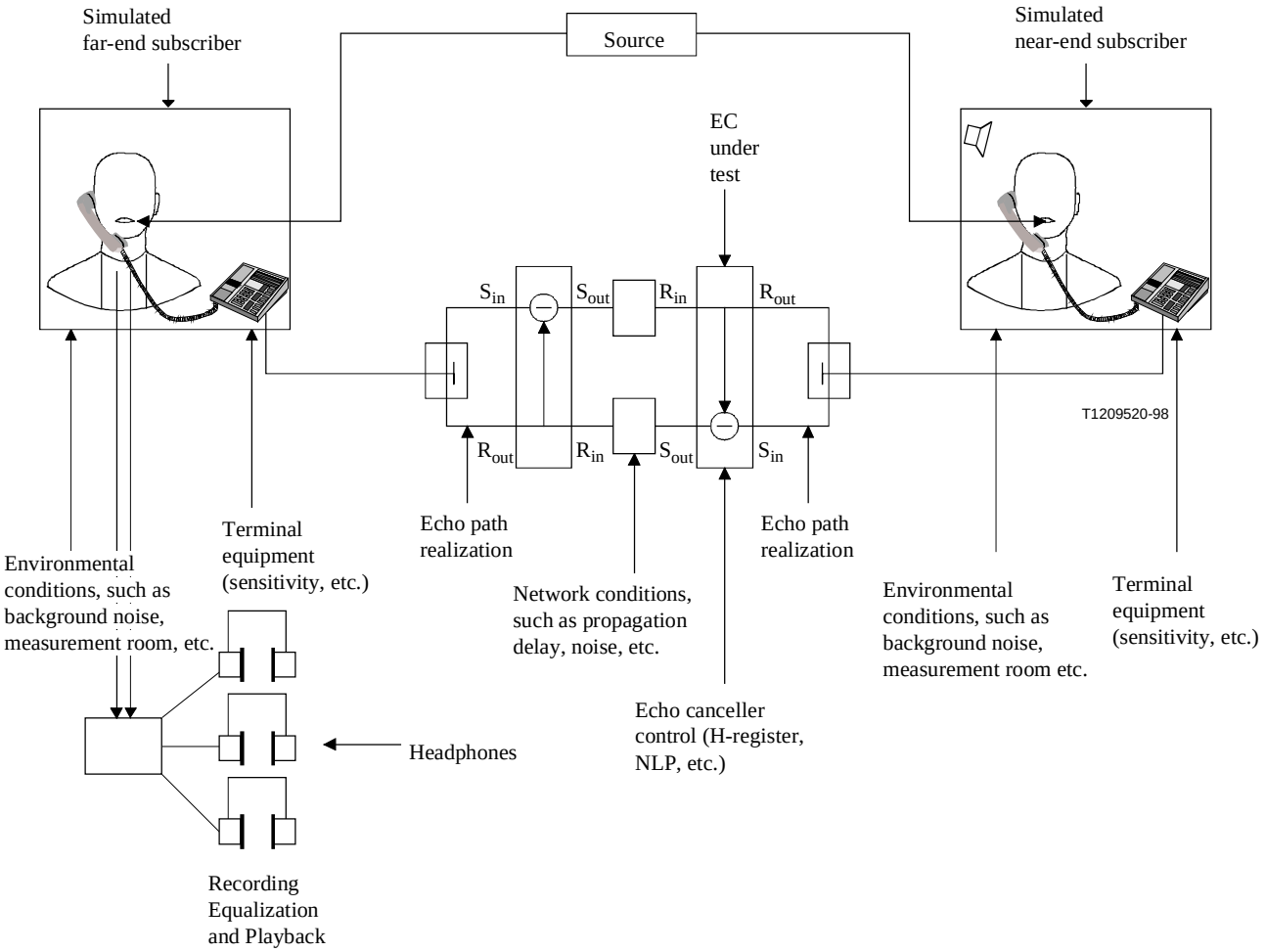


Figure 3/P.831 – Experimental set-up for recordings of speech material for listening test (Type A) using two P.58 HATS

6.4 Recording procedure for original source material

The artificial mouths are fed from appropriate pre-recorded source material which has typically been recorded and stored on a high quality digital recording medium, and is then played back. Recommendations P.800 and P.830 should be consulted for guidance on preparing recorded speech material. Such recordings allow the preparation and composition of various speech sequences including possible double-talk periods, if necessary. If double-talk sequences are used, it is recommended that a male and a female voice are used to distinguish between the talkers, as mentioned in 6.5. The subjects used for the original source material should be located in a quiet environment (e.g. quiet listening rooms), to avoid additional background noise (note that the complete acoustic environment is pre-recorded). Equalization before playback ensures that the listening situation is reproduced as closely as possible.

An alternative to using two speech signals (one male and one female talker) is to use one speech signal and the Composite Source Signal (CSS) defined in Recommendation P.501. Pairing a speech signal with an artificial speech signal such as the CSS allows the listener to gauge the amount of clipping and distortion suffered by the near-end party's speech signal during a double-talk episode.

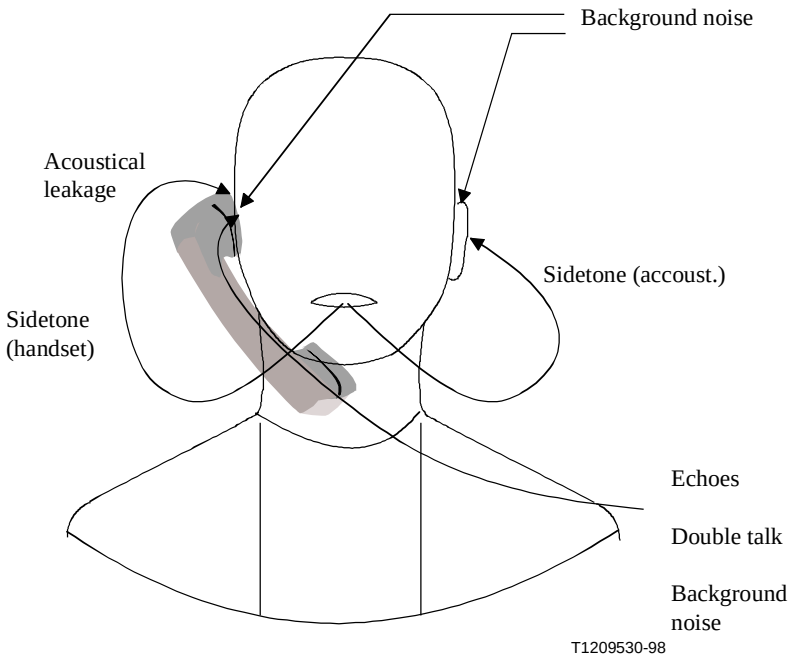


Figure 4/P.831 – Recordings of speech material for listening test Type A using a HATS with mounted handset – For clarity, the handset mounting device itself is not shown

6.5 Recording procedure for listening test material

Figure 4 demonstrates the procedure for making recordings with artificial heads. All signals and transmission paths which contribute to the ear signals are shown. The systems must be equipped with P.58 artificial mouths and P.57 Type 3.4 pinnae. The listening situation is as follows: the handset receives signals from the network (speech from the other subscriber, background noise and possible echoes) and transmits it to one ear. The listening examples to be evaluated are therefore presented monaurally. Speech from the local artificial mouth is transmitted to both ears, but in a different way for each ear. The ear covered by the handset receives its own voice through the leakage between handset and ear and additionally via the sidetone path of the handset. These transmission characteristics are pressure force dependent. The other ear, which is not covered by the handset, receives the voice directly from the mouth. The original signal is therefore presented binaurally, but with a significant difference for both ears. The signal to be evaluated is presented monaurally.

The speech samples used to feed the artificial mouths during the recordings for this listening test should be pre-recorded as described previously. Double-talk sequences should be composed in an appropriate way (note that the starting point of double-talk can strongly influence the echo canceller operation and hence the overall transmission quality). The start of the recordings should be synchronized with the control of other relevant parameters, e.g. control of the echo cancellers or the start of the background noise (note that in general, the long-term level and spectral characteristics of background noise may affect echo canceller operation). For recordings under double-talk conditions, care must be taken to distinguish between the near- and far-end speech in the listening test later. From the experience gained with the listening tests, it is recommended that a male and a female voice are used to distinguish between the talkers. One system plays back test sentences of a female voice and the double-talk signal is fed from the artificial mouth of the other measurement system, which simulates the near-end subscriber, using a male voice (or vice versa). If different speakers are used, it is easier for the subjects to concentrate only on the double-talk signal during the listening tests.

NOTE – At least one artificial head measurement system is necessary for the recordings. It should be placed in a suitable location, taking into account room characteristics and background noise conditions. The simulation of the near-end speaker (the location of which is not judged by test subjects) is not so critical, and can be made by using an artificial mouth according to Recommendation P.51 and placing the handset in the LRGP position. However, where background noise is also required at the near-end, it is recommended that a HATS is used instead of the artificial mouth according to Recommendation P.51.

6.6 Description of the listening test procedure

6.6.1 Playback

The playback procedure must ensure that the listening samples are exact reproductions of the ear signals. Equalized headphones are therefore necessary. The stimuli should be presented in a comparable way to that experienced by the subjects during normal telephone use: right-handed people normally use handsets with the left hand at the left ear (the right hand is often kept free). They are therefore used to listening with their left ear, and so the headphone channels should be chosen in the same way.

6.6.2 Subjects

Too much background information and explanations about the recording set-up should be avoided for untrained subjects. The recording procedure is very sophisticated. So far, no investigations have been made to verify the influence of different types of explanations for untrained subjects. From the experience gained with this test, it is therefore recommended that the only information given is that required to explain that the disturbances can only be heard in one ear. In addition, a sufficiently long period of training is recommended before the tests start, to familiarise the subjects with the listening situation. A minimum of 10 training sentences, representing the whole range of quality degradation (not including the extremes) is recommended.

6.7 Advantages

The measurement conditions in the set-up can be controlled accurately, and all echo cancellers may be tested under identical conditions. The number of test conditions or echo cancellers can be adjusted easily. If several echo cancellers, implementations or many environmental conditions are to be included, the procedure takes less time than other tests. The numbers of subjects can easily be increased, and only one set of recordings needs to be made. The tests may be separated into parts, for example the evaluation of initial convergence or steady state conditions. Also, the simulation of a whole conversation with two artificial head measurement systems allows recordings under single and double-talk conditions.

The test is suited to the evaluation of specific parameters because subjects can concentrate better on these parameters. The perception of subjectively relevant parameters is in general highly influenced by various parameters like sensitivity, linear and non-linear distortions of terminal equipment, coupling between handset and ear (leakage), handset sidetone, masking effects and others. The recordings ensure that a very high degree of realism is reproduced for third-party listening tests. Subjects judge listening examples, which are recorded at the acoustic interface. Thus all the parameters mentioned above (including masking by the original voice) are included. Echo cancellers may be directly judged by A/B comparisons. The test is a suitable method for evaluating even small differences between different implementations or different measurement conditions.

6.8 Disadvantages

The test procedure is artificial compared to other tests, where subjects are allowed to talk. Although masking effects, terminal equipment (including leakage) and other parameters are considered with

this procedure, subjects are asked to listen and judge recordings of unknown speakers. The naturalness of hearing their own voice speaking is therefore missing. These listening tests are intended to supplement overall quality evaluations. They allow only detailed parameter investigations, and require comprehensive preparation, but provide a very efficient test procedure to evaluate echo canceller differences.

6.9 Reference conditions

Reference conditions can be included. These listening samples can be presented with the real recordings during the test. Reference conditions allow results from different laboratories to be compared, and may include test set-ups without an echo canceller, but with well-defined residual echo levels (achieved by varying the echo path attenuation in steps) and other parameters.

NOTE – Reference conditions that include echo cancellers (comparable to MNRU conditions for codec tests) should be carefully designed to represent typical quality impairments introduced by echo cancellers. Such conditions should include modulated background noise (typically caused by non-linear processes like center clippers) and different echo disturbance simulations, such as switched, interrupted echoes, continuous or time variant echoes (as they typically appear during initial convergence). The same remarks are applicable to reference conditions under double-talk conditions.

6.10 Precautions

To guarantee an exact acoustic reproduction of the recordings, equalized headphones should be used. Subjects should be carefully briefed before the test because the listening situation is quite sophisticated. The recordings are binaural, therefore both ear signals are different. Subjects hear the original speech signal in both ears in a different way than if a handset were used. Typically the disturbances (like echoes, modulated background noise) can only be heard in one ear (the one normally covered by the handset). Background information and explanations of the recording set-up are complicated for untrained subjects. From experience gained with this test, it is therefore recommended that subjects are only told that the disturbances can be heard only in one ear.

7 Third-party listening Test B

7.1 Purpose

This test describes an easy procedure for comparison of different echo cancellers. Primarily the method is used to judge the relative difference between echo cancellers. It is also possible to use this method for individual evaluation of echo cancellers.

Echo cancellers must work under a wide range of conditions and they can produce many kinds of speech transmission impairments. Possible conditions are of course a wide range of live-networks, but also simulated environments where impairments are isolated to fully understand the performance of the echo canceller under test.

A perfect echo canceller would be perceived as fully transparent by the user. The intention is that the user shall not be able to perceive that echo is present and that an echo canceller device has been inserted to handle the echo. The near-end signal, including background noise, shall be transmitted and no other impairments should be noticeable.

Subjective evaluation can be done with this in mind. All types of distortion on the near-end signal are faults and can be evaluated as such.

An efficient, and a discriminating method for evaluation of echo canceller performance is to use subjective listening tests. Recorded examples of sequences of far-end and near-end speech, or near-end speech and far-end CSS, are fed to the echo canceller under test (to terminals R_{in} and S_{in}). The

test subjects listen to the recorded output of terminal S_{out} and make judgements of the speech quality. To obtain reactions similar to those of subscribers in live networks, the test teams should consist of untrained listeners. Trained listeners can be used for judgement of specific details of the echo canceller performance. With untrained listeners it is important that they are not aware that the test concerns echo cancellers.

7.2 Impairments to evaluate

This test procedure can be used for evaluation of the following impairments:

During far-end single talk:	Echo Lack of background noise transparency Tandem
During double-talk:	Echo Distortion of the near-end signal Clipping of the near-end signal
During near-end single talk:	Distortion of the near-end signal Clipping of the near-end signal

7.3 Set-up

The test set-up, shown in Figure 5, is intended to be simple. All speech samples should be processed by all of the echo cancellers that will be tested. The S_{out} signals should be recorded for evaluation. To simulate operational conditions, a live network or some kind of laboratory network with known performance can be used.

All the test material can be prepared in advance and inserted electrically in the set-up.

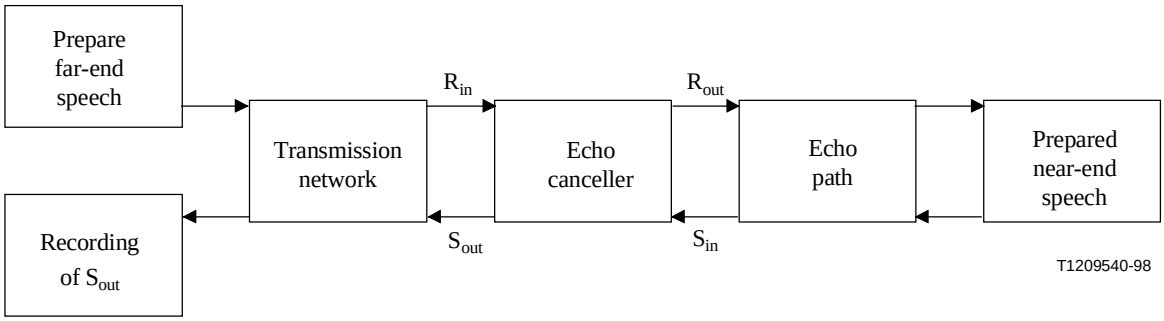


Figure 5/P.831 – Echo canceller test set-up for speech input and output

7.4 Advantages

The main advantage with this method is the simple recording procedure. The use of digital tape recorders is sufficient, if the material has been properly prepared. Since an acoustic environment is not needed during recording, the processing through the echo cancellers can be accomplished using electrical interfaces. This is very useful in a live network.

In a test environment, all network conditions, such as line levels, ERLs and delays, can easily be controlled. If a digital interface is used, bit exact repetition of recordings for the different echo cancellers can be obtained.

During the listening sessions, each listener will be exposed to all impairments present in the test. This means that the impairments will not be weighted, which is important since the relative importance of the different impairments is not known a priori.

7.5 Disadvantages

The main disadvantage is that the correlation between recorded distortions and their perception during a live conversation is not known. This correlation needs further study.

7.6 Reference conditions

It is easy to introduce reference conditions. The obvious one is a near-end signal without echo or echo cancellers. This will give a reference for the maximum quality possible with the network scenarios under consideration. A valid reference system for typical impairments in the echo path or impairments introduced by the echo canceller (e.g. NLP clipping) is not available. MNRU reference conditions have been used in the past, even though the modulated distortion produced by the MNRU is not representative of typical echo canceller impairments. A more appropriate reference impairment system is therefore required and needs further study.

7.7 Precautions

As with regular subjective tests, it is important not to prepare the untrained test subjects. They shall evaluate the distortion of the desired signal, the near-end signal. If the test subjects are told in advance that they are going to evaluate echo cancellers, this will automatically give echo impairments more focus when compared to other impairments. Preliminary instructions, appropriate for the evaluation method, can be used. For examples, see Recommendation P.800 and the *Handbook on Telephonometry*.

7.8 Description of test procedure

7.8.1 Recording

The source speech samples should be prepared in advance, which may include level equalization, spectral weighting of the speech signal, electrical summation of background noise to clean speech, etc. If different scenarios (different telephone handsets, background noise types, etc.) are to be tested, the speech material can be recorded under the environmental conditions of these scenarios. It is not necessary to use an acoustic interface to record signals processed through the echo canceller under test. If possible, it is preferable to have a digital interface to the different echo cancellers, which will assure that all echo cancellers will be processed through identical circuit conditions. If it is not possible to use a digital interface, as in the case of a live system, analogue electrical interfaces can be carefully used.

The source speech material length has to be carefully selected. Too little near-end speech (which will be perceived as silence) will confuse the subjects, since they will be evaluating the distortion of the near-end speech. The speech samples should be as short as possible. However, to get a good representation of all impairments, samples that are too short should not be used for echo canceller evaluation.

Figure 6 shows an example of the speech signals at the far- and near-end speech inputs, respectively, where a 20-second duration has been used. The far-end starts with one sentence pair, while the near-end is silent. The first sentence-pair of far-end speech will allow the echo canceller to converge (Segment A indicated in Figure 6). After that, a double-talk period occurs with two sentence-pairs of far-end speech and two sentence-pairs of near-end speech (Segment B in Figure 6). Each sentence-pair is approximately 10 seconds long. The recording of the processed material begins after the initial

far-end period (Point C in Figure 6). The speech quality assessment period lasts for about 20 seconds. Figure 6 shows an example of the speech signals at the far- and near-end speech inputs respectively.

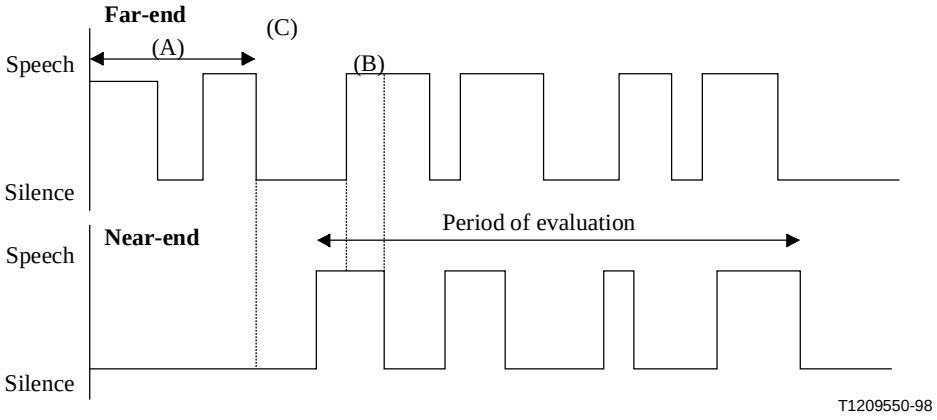


Figure 6/P.831 – Example of far-end and near-end speech input signal

7.9 Evaluation

Different types of known evaluation methods can be used. ACR and pair-comparison tests have been extensively used in the past. It should be noted, however, that if an ACR scale is used for echo canceller evaluation, the scores may be low. The example in Figure 6 shows a forced impairment test to simulate difficult periods of a conversation. Since it represents, for a long time, an impairment which would occur for much shorter periods of time in a real conversation, it is natural that the scores be low for the echo canceller under test.

Therefore, the scores obtained from listening tests of echo cancellers shall be used as relative values for comparison of different echo cancellers, implementations, conditions, etc. For maximum likelihood that such comparisons would be valid, however, it would be important that all cancellers, conditions, etc. to be compared be present in the same subjective test, or be present in different subjective tests with similar structure. It should be noted that, in general, direct comparison of scores obtained in different listening tests is not recommended.

7.9.1 Playback

The listening sessions should be performed in an environment conforming to the description in Recommendation P.800 for listening tests, and presentation of test material for a subject can be done according to standard procedures, see Recommendation P.800 and the *Handbook on Telephonometry*.

7.9.2 Subjects

Subjects shall be selected from the normal user population. No specific preparations shall be made. Preliminary instructions to be given to the subjects should follow the guidelines in Recommendations P.800 and P.830 and the *Handbook on Telephonometry*.

ANNEX A

Questionnaires for use in conversational tests

A.1 Introduction

Some Administrations have found that multiple questions following conversations can be used effectively in evaluating the subjective performance of ECs. This annex shows one such questionnaire using a quality scale and two impairment scales (an echo annoyance scale and a noise annoyance scale). The impairment scale leads to a Degradation Mean Opinion Score (DMOS) and is based on the DCR method (see Recommendation P.800). Application of the DCR to conversation tests differs from the procedure recommended for listening tests in that no explicit high-quality reference is introduced prior to each evaluation. In some languages, it may be appropriate to use the word ’impairment’ or ’disturbance’ in place of ’degradation’.

It may be desirable to ask other questions in addition to (or in place of) these example questions.

A.2 Sample questions

Subjects answer each of the following questions after each trial.

Q1: What is your opinion of the connection you have just been using?

Excellent

Good

Fair

Poor

Bad

Q2: Did you or your partner have any difficulty in talking or hearing over the connection?

Yes

No

Q3: How would you qualify the communication?

Unacceptable

Acceptable

Q4: How would you judge the degradation from echo of your own voice?

Imperceptible

Perceptible but not annoying

Slightly annoying

Annoying

Very annoying

Q5: How would you judge other degradations (clipping, various noises ...)?

Imperceptible

Perceptible but not annoying

Slightly annoying

Annoying

Very annoying

Q6: How did you find the voice of your partner ?

Unnatural

…

Natural

APPENDIX I

Example test conditions for echo canceller evaluations

Table I.1 shows parameters for five test conditions. These conditions are appropriate for use in conversational tests, in talking-and-listening tests, and in the third-party listening tests. These conditions are not intended to provide a comprehensive evaluation of an EC. However, they do illustrate the kinds of experimental manipulations that should be considered when evaluating ECs.

Table I.1/P.831 – Example test conditions for echo canceller evaluations

Parameter	Condition 1 ("Good")	Condition 2 ("Bad" #1)	Condition 3 ("Bad" #2)	Condition 4 ("Noisy" #1)	Condition 5 ("Noisy" #2)
Room Noise (West)	None	None	None	Babble at 50 dBA	None
Room Noise (East)	None	None	None	None	Babble at 50 dBA
ERL	14-17 dB, "Flat"	6-8 dB, "Flat"	6-8 dB, "Flat"	14-17 dB, "Flat"	14-17 dB, "Flat"
Tail Delay (Round Trip)	20-24 msec	20-24 msec	20-24 msec	20-24 msec	20-24 msec
Bulk Delay (Round Trip)	100-150 msec	100-150 msec	100-150 msec	100-150 msec	100-150 msec
SLR (East)	8 dB	0 dB	14 dB	8 dB	8 dB
RLR (East)	2 dB	5 dB	6 dB	2 dB	2 dB
SLR (West)	8 dB	14 dB	0 dB	8 dB	8 dB
RLR (West)	2 dB	6 dB	5 dB	2 dB	2 dB

NOTE 1 – The terms "East" and "West" are used to distinguish the ends of a connection.

NOTE 2 – ERL values will depend on the network in which the EC is to be deployed. Exact values will depend on the distribution of ERL for a specific network.

NOTE 3 – Tail delay is the round trip delay in the tail of the echo canceller, i.e. from the Receive Out port to the Send In port, through the hybrid.

NOTE 4 – Bulk delay is the delay inserted between two echo cancellers, i.e. in the "network"

When conducting listening tests, it is advisable to include MNRU conditions according to Recommendation P.810 (MNRU settings of 6, 12, 18, 24, 30, 36 dBQ are suggested). Inclusion of MNRU conditions will facilitate comparisons between tests that are conducted in different laboratories or in the same laboratory but at different times.

For third-party listening tests, the speech samples should be prepared using at least two male and two female talkers. The duration of the speech samples will depend on the objectives of the test. As a guideline, it is suggested that speech samples should be at least 12 s in duration. If the objective of the test is to study double-talk performance, then the speech samples should contain a minimum of 10-20% double-talk.

ITU-T RECOMMENDATIONS SERIES

Series A	Organization of the work of the ITU-T
Series B	Means of expression: definitions, symbols, classification
Series C	General telecommunication statistics
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	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	TMN and network maintenance: international transmission systems, telephone circuits, telegraphy, facsimile and leased circuits
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 and open system communications
Series Y	Global information infrastructure
Series Z	Programming languages