INTERNATIONAL TELECOMMUNICATION UNION

CCITT

THE INTERNATIONAL
TELEGRAPH AND TELEPHONE
CONSULTATIVE COMMITTEE

RED BOOK

VOLUME V

TELEPHONE TRANSMISSION QUALITY

RECOMMENDATIONS OF THE P SERIES

VIIITH PLENARY ASSEMBLY
MALAGA-TORREMOLINOS, 8-19 OCTOBER 1984

Geneva 1985

INTERNATIONAL TELECOMMUNICATION UNION

CCITT

THE INTERNATIONAL
TELEGRAPH AND TELEPHONE
CONSULTATIVE COMMITTEE

RED BOOK

VOLUME V

TELEPHONE TRANSMISSION QUALITY

RECOMMENDATIONS OF THE P SERIES

VIII[TH] PLENARY ASSEMBLY

MALAGA-TORREMOLINOS, 8-19 OCTOBER 1984

Geneva 1985

ISBN 92-61-02131-X

CONTENTS OF THE CCITT BOOK
APPLICABLE AFTER THE EIGHTH PLENARY ASSEMBLY (1984)

RED BOOK

Volume IV – *(4 fascicles, sold separately)*

FASCICLE IV.1 – Maintenance; general principles, international transmission systems, international telephone circuits. Recommendations M.10-M.762 (Study Group IV).

FASCICLE IV.2 – Maintenance; international voice frequency telegraphy and fascimile, international leased circuits. Recommendations M.800-M.1375 (Study Group IV).

FASCICLE IV.3 – Maintenance; international sound programme and television transmission circuits. Series N Recommendations (Study Group IV).

FASCICLE IV.4 – Specifications of measuring equipment. Series 0 Recommendations (Study Group IV).

Volume V – Telephone transmission quality. Series P Recommendations (Study Group XII).

Volume VI – *(13 fascicles, sold separately)*

FASCICLE VI.1 – General Recommendations on telephone switching and signalling. Interface with the maritime mobile service and the land mobile services. Recommendations Q.1-Q.118 *bis* (Study Group XI).

FASCICLE VI.2 – Specifications of Signalling Systems Nos. 4 and 5. Recommendations Q.120-Q.180 (Study Group XI).

FASCICLE VI.3 – Specifications of Signalling System No. 6. Recommendations Q.251-Q.300 (Study Group XI).

FASCICLE VI.4 – Specifications of Signalling Systems R1 and R2. Recommendations Q.310-Q.490 (Study Group XI).

FASCICLE VI.5 – Digital transit exchanges in integrated digital networks and mixed analogue-digital networks. Digital local and combined exchanges. Recommendations Q.501-Q.517 (Study Group XI).

FASCICLE VI.6 – Interworking of signalling systems. Recommendations Q.601-Q.685 (Study Group XI).

FASCICLE VI.7 – Specifications of Signalling System No. 7. Recommendations Q.701-Q.714 (Study Group XI).

FASCICLE VI.8 – Specifications of Signalling System No. 7. Recommendations Q.721-Q.795 (Study Group XI).

FASCICLE VI.9 – Digital access signalling system. Recommendations Q.920-Q.931 (Study Group XI).

FASCICLE VI.10 – Functional Specification and Description Language (SDL). Recommendations Z.101-Z.104 (Study Group XI).

FASCICLE VI.11 – Functional Specification and Description Language (SDL), annexes to Recommendations Z.101-Z.104 (Study Group XI).

FASCICLE VI.12 – CCITT High Level Language (CHILL). Recommendation Z.200 (Study Group XI).

FASCICLE VI.13 – Man-Machine Language (MML). Recommendations Z.301-Z.341 (Study Group XI).

Volume VII – *(3 fascicles, sold separately)*

FASCICLE VII.1 – Telegraph transmission. Series R Recommendations (Study Group IX). Telegraph services terminal equipment. Series S Recommendations (Study Group IX).

FASCICLE VII.2 – Telegraph switching. Series U Recommendations (Study Group IX).

FASCICLE VII.3 – Terminal equipment and protocols for telematic services. Series T Recommendations (Study Group VIII).

Volume VIII – *(7 fascicles, sold separately)*

FASCICLE VIII.1 – Data communication over the telephone network. Series V Recommendations (Study Group XVII).

FASCICLE VIII.2 – Data communication networks: services and facilities. Recommendations X.1-X.15 (Study Group VII).

FASCICLE VIII.3 – Data communication networks: interfaces. Recommendations X.20-X.32 (Study Group VII).

FASCICLE VIII.4 – Data communication networks: transmission, signalling and switching, network aspects, maintenance and administrative arrangements. Recommendations X.40-X.181 (Study Group VII).

FASCICLE VIII.5 – Data communication networks: Open Systems Interconnection (OSI), system description techniques. Recommendations X.200-X.250 (Study Group VII).

FASCICLE VIII.6 – Data communication networks: interworking between networks, mobile data transmission systems. Recommendations X.300-X.353 (Study Group VII).

FASCICLE VIII.7 – Data communication networks: message handling systems. Recommendations X.400-X.430 (Study Group VII).

Volume IX – Protection against interference. Series K Recommendations (Study Group V). Construction, installation and protection of cable, and other elements of outside plant. Series L Recommendations (Study Group VI).

Volume X – *(2 fascicles, sold separately)*

FASCICLE X.1 – Terms and definitions.

FASCICLE X.2 – Index of the Red Book.

PAGE INTENTIONALLY LEFT BLANK

PAGE LAISSEE EN BLANC INTENTIONNELLEMENT

CONTENTS OF VOLUME V OF THE RED BOOK

Part I – Series P Recommendations

Telephone transmission quality

Part II – Supplements to Series P Recommendations

PRELIMINARY NOTES

1 This volume fully supersedes Volume V of the CCITT *Yellow Book* (Geneva, 1981).

It has been indicated (immediately after the titles of Recommendations or Supplements) whether the texts are new ones approved by the Plenary Assembly of Malaga-Torremolinos, 1984 or are texts amended at the same period. Texts without any such indication date from at least as far back as the Plenary Assembly of New Delhi, 1960, when Volume V was divided into numbered Recommendations; certain of these texts may be even older.

2 The units used in this Volume are in conformity with CCITT Recommendations B.3 and B.4 (Volume I of the *Red Book*).

The indication "amended Malaga-Torremolinos, 1984" has not been affixed to those Recommendations in which the only amendment has been an editorial change concerning units.

The following abbreviations are used, particularly in diagrams and tables, and always have the following clearly defined meanings:

dBm the absolute power level in decibels;

dBm0 the absolute power level in decibels referred to a point of zero relative level;

dBr the relative power level in decibels;

dBm0p the absolute psophometric power level in decibels referred to a point of zero relative level.

The units for air pressure are related as follows:

1 Pa (Pascal) = 1 N (Newton)/m^2 = 10 dyne/cm^2 = 10 barye = 10 μbar

3 The Questions entrusted to each Study Group for the Study Period 1985-1988 can be found in Contribution No. 1 to that Study Group.

4 In this Volume, the expression "Administration" is used for shortness to indicate both a telecommunication Administration and a recognized private operating agency.

PART I

Series P Recommendations

TELEPHONE TRANSMISSION QUALITY

PAGE INTENTIONALLY LEFT BLANK

PAGE LAISSEE EN BLANC INTENTIONNELLEMENT

SECTION 1

VOCABULARY
AND
EFFECTS OF TRANSMISSION PARAMETERS ON CUSTOMER OPINION OF TRANSMISSION QUALITY AND THEIR ASSESSMENT

Recommendation P.10

VOCABULARY OF TERMS ON TELEPHONE TRANSMISSION QUALITY AND TELEPHONE SETS

(Geneva, 1980; amended at Malaga-Torremolinos)

1 Introduction

This Recommendation contains terms and definitions appropriate to the work of Study Group XII which were discussed within the Group of Experts N of the Joint Coordinating Group for the CCIs and the IEC.

Terms which appear in the International Electrotechnical Vocabulary (IEV) (Chapter 722) have their IEV number reproduced here for reference purposes. Terms of the CCITT have been classified in a manner similar to that used in the IEV.

2 Terms and definitions

02. *Telephone set components*

02.01 **Y-ratio**

F: rapport Y

S: relación Y

The ratio between the sending and receiving efficiencies of a passive *telephone set* circuit.

04. *Telephone set types*

04.01 **telephone set; telephone instrument**

F: poste téléphonique: appareil téléphonique; téléphone

S: aparato telefónico; teléfono

An assembly of apparatus for *telephony* including at least a *telephone transmitter*, a *telephone receiver* and the wiring and components immediately asociated with these transducers.

Note – A telephone set usually includes other components such as a *switchhook*, a built-in *telephone bell*, and a *dial*.

722.04.01

04.02 **telephone station**

F: poste téléphonique (installé)

S: estación telefónica

A *telephone set* with associated wiring and auxiliary equipment connected to a *telephone network* for the purpose of *telephony.*

Note – The auxiliary equipment may include, for example, an external *call indicating device*, a protector, a *local battery.*

722.04.02

04.03 **loudspeaking (telephone) set**

F: poste (téléphonique) à haut-parleur; téléphone à haut-parleur

S: aparato telefónico con altavoz; teléfono de altavoz

A *telephone set* using a *loudspeaker* associated with an amplifier as a *telephone receiver.*

722.04.10

04.04 **hands free (telephone) set**

F: poste (téléphonique) mains libre

S: aparato telefónico de manos libres; teléfono de manos libres

A *loudspeaking telephone set* which may be used without a *handset.*

722.04.11

05. *Telephone set accessories*

05.01 **acoustic shock suppressor (in telephony)**

F: anti-choc (en téléphonie)

S: supresor de choques acústicos; antichoque (en telefonía)

A device associated with a *telephone station* and intended to prevent *acoustic shocks*, by setting an upper limit to the absolute values of the instantaneous electrical voltage that can be applied to the *telephone earphone.*

722.05.07

13. *Private telephone systems*

13.01 **private (telephone) installation**

F: installation (téléphonique) intérieure

S: instalación telefónica privada

A *telephone network* installed on the premises of a single individual or organization.

Note – By convention, private telephone installations include sets of *telephone stations* which are connected to one *subscriber's line*

722.13.01

21. *Telephone calls description*

21.01 **call attempt (by a user)**

F: (tentative d')appel (par un usager)

S: (tentativa de) llamada (por un usario)

A sequence of operations made by a user of a telecommunication network trying to obtain the desired user or service.

Associated term: to *call*

722.21.01; identical to 701.03.04

21.02 **connection**

F: chaîne de connexion

S: cadena de conexión; conexión

A temporary association of transmission channels or telecommunication circuits, switching and other functional units set up to provide the means of a transfer of information between two or more points in a telecommunication network.

722.21.02; identical to 701.03.01

21.03 **(complete) connection**

F: chaîne de connexion complète; (chemin de) communication

S: cadena de conexión completa; conexión completa

A *connection* between users' terminals.

722.21.03; identical to 701.03.02

21.04 **call**

F: communication

S: comunicación

The establishment and use of a *complete connection* following a *call attempt*

722.21.04; identical to 701.03.05

31. *Local line networks*

31.01 **local line network**

F: réseau local de lignes (téléphoniques)

S: red local de líneas (telefónicas)

All the *subscribers' telephone lines* and ancillary equipment provided to connect *subscribers* to their *local switching entity.*

722.31.01

31.02 **subscriber's (telephone) line; subscriber loop (in telephony)**

F: ligne (téléphonique) d'abonné; ligne (de) réseau

S: línea (telefónica) de abonado; bucle de abonado (en telefonía)

A link between a public *switching entity* and a *telephone station* or a *private telephone installation* or another terminal using signals compatible with the *telephone network.*

Note – In French, the term "ligne de réseau" is used only when the private telephone installation is a *private branch exchange* or an *internal telephone system.*

722.31.02

32. *Telephone station usage*

32.01 **acoustic hood**

F: abri téléphonique; abriphone

S: cabina acústica; burbuja acústica

A hood lined with sound-absorbing material to faciliate the use of a *telephone station* by reducing the *ambient noise* level.

722.32.03

32.02 **telephone booth**

F: cabine téléphonique

S: cabina telefónica cerrada

A small cabin containing a *telephone station* and providing a certain measure of acoustic insulation and privacy for the user.

722.32.04

32.03 **telephone stall**

F: cabine téléphonique ouverte

S: cabina telefónica abierta

A *telephone booth* without a door.

722.32.05

41. *Transmission performance*

41.01 **acoustic shock (in telephony)**

F: choc acoustique (en téléphonie)

S: choque acústico (en telefonía)

Any temporary or permanent disturbance of the functioning of the ear, or of the nervous system, which may be caused to the user of a *telephone earphone* by a sudden sharp rise in the acoustic pressure produced by it.

Note – An acoustic shock usually results from the occurrence, in abnormal circumstances, of short-lived high voltages at the terminals of a *telephone set.*

722.41.20

41.02 **opinion score (in telephony)**

F: note d'opinion (en téléphonie)

S: nota de opinión (en telefonía)

The value on a predefined scale that a subject assigns to his opinion of the performance of the telephone transmission system used either for conversation or only for listening to spoken material.

Note – According to the IEV, the scale generally consists of five values, for example: excellent, good, fair, bad, unfair. This example does not correspond to CCITT practice (see Notes 2 and 3 of Recommendation P.77).

722.41.24

42. *Measuring apparatus*:

42.01 **acoustic coupler (in telephonometry)**

F: coupleur acoustique (en téléphonométrie)

S: acoplador acústico (en telefonometría)

A cavity of defined shape and volume used for the testing of *telephone earphones* or *telephone transmitters* in conjunction with a calibrated microphone adapted to measure the pressure developed within the cavity.

722.42.12

42.02 **artificial ear**

F: oreille artificielle

S: oído artificial

A device for the calibration of earphones incorporating an *acoustic coupler* and a calibrated microphone for the measurement of sound pressure and having an overall acoustic impedance similar to that of the average human ear over a given frequency band.

722.42.13

42.03 **artificial mouth**

F: bouche artificielle

S: boca artificial

A device consisting of a *loudspeaker* mounted in an enclosure and having a directivity and radiation pattern similar to those of the average human mouth.

722.42.14

42.04 **artificial voice**

F: voix artificielle

S: voz artificial

A complex sound, usually emitted by an artificial mouth and having a power sound spectrum corresponding to that of the average human voice.

722.42.15

42.05 **local (telephone) system; local (telephone) circuit**

F: système (téléphonique) local; circuit (téléphonique) local

S: sistema (telefónico) local

The combination of *subscriber's station, subscriber's line* and *feeding bridge* if present.

Note 1 – This term is used in the context of *transmission* planning and *performance.*

Note 2 – In CCITT English texts, the term "local (telephone) system" is preferred.

722.42.15

42.06 **subscriber system (in transmission planning)**

F: système d'abonné

S: sistema de abonado

A *subscriber's line* associated with that part of the *private telephone installation* connected to this line during a telephone *call.*

Note – This term is used in the context of *transmission* planning and *performance.*

722.42.17

42.07 **SETED**

F: SETED

S: SETED

A working standard using high-quality moving coil microphones and receivers with an orthotelephonic response.

42.08 **balance return loss**

F: affaiblissement d'équilibrage

S: atenuación de equilibrado

Value of the return loss at a termination given by:

$$20 \log_{10} \left| \frac{Z_b + Z_a}{Z_b - Z_a} \right| \text{ dB}$$

where:

Z_a is the impedance of terminated line, or apparatus,

Z_b is the impedance of network.

42.09 **artificial voice**

F: voix artificielle

S: voz artificial

A mathematically defined signal which reproduces human speech characteristics, relevant to the characterisation of linear and nonlinear telecommunication systems. It is intended to give a satisfactory correlation between objective measurements and tests with real speech.

42.10 **electrical artificial voice**

F: voix artificielle électrique

S: voz eléctrica artificial

The artificial voice produced as an electric signal, for testing transmission channels or other electric devices.

42.11 **artificial mouth excitation signal**

F: signal d'excitation de la bouche artificielle

S: señal de excitación de la boca artificial

A signal applied to the artificial mouth in order to produce the acoustic artificial voice. It is obtained by equalizing the electrical artificial voice for compensating the sensitivity/frequency characteristic of the mouth.

42.12 **acoustic artificial voice**

F: voix artificielle acoustique

S: voz acústica artificial

Acoustic signal at the MRP (Mouth Reference Point) of the artificial mouth. It complies with the same time and spectral specifications as the electrical artificial voice.

43. *Telephonometry*

43.01 **reference equivalent**

F: équivalent de référence

S: equivalente de referencia

The loss, expressed in decibels, constant at all frequencies transmitted, which has to be introduced into the new *fundamental system for the determination of reference equivalents* or NOSFER in order to obtain in a given direction the same *loudness* as the *complete telephone connection* being considered, the *acoustical speech power* emitted by the talker being the same in both cases.

Note 1 – The reference equivalent is positive or negative according to whether it has been necessary for a loss to be added or removed from the NOSFER.

Note 2 – The reference equivalent is strictly defined by the measuring method described in Recommendation P.72.

722.43.14

43.02 **corrected reference equivalents**

F: équivalents de référence corrigés (ERC)

S: equivalentes de referencia corregidos (ERC)

Values of sending or receiving *reference equivalent* converted by a defined, nonlinear, transformation into corresponding values that obey the laws of algebraic addition.

Note – The conversion is performed to avoid some of the difficulties experienced in applying *reference equivalents*. It is defined in Annex A to Recommendation G.111.

722.43.17

43.03 **relative equivalent**

F: équivalent relatif

S: equivalente relativo

The loss, expressed in decibels, which has to be introduced into a *working standard* in order to obtain in a given direction the same *loudness* as the *telephone call* being considered, the *acoustic speech power* emitted by the talker being the same in both cases.

Note 1 – The relative equivalent is positive or negative according to whether it has been necessary for a loss to be added or removed from that standard.

Note 2 – The *reference equivalent* is approximately equal to the sum of the relative equivalent measured with respect to a working standard and of the reference equivalent of this working standard.

Note 3 – The relative equivalent is strictly defined by the measuring method described in Recommendation P.72.

722.43.22

43.04 **loudness rating**

F: équivalent pour la sonie

S: índice de sonoridad

A measure, expressed in decibels, for characterizing the *loudness* performance of *complete telephone connections* or of parts thereof such as *sending system, line, receiving system.*

Note – The CCITT is studying how to implement the principles described in Recommendation P.76 so as to overcome certain serious difficulties experienced in applying reference equivalents.

Note 2 – (added by the CCITT) – This definition is very general and corresponds to what is described as *loudness loss* in CCITT texts; in those texts, the term "loudness rating" should be confined to measurements in conformity with Recommendation P.76, and may be abbreviated as LR.

722.43.25

43.05 **R25 equivalent**

F: équivalent R25

S: equivalente R25

Loudness loss determined as a *reference equivalent* in accordance with Recommendation P.72, except that the listening level is constant, corresponding to 25 dB in NOSFER.

43.06 **planning equivalent**

F: équivalent de planification

S: equivalente de planificación

Result of a measurement with an objective meter which may be considered equal to an *R25 equivalent* or to a *corrected reference equivalent* with an accuracy which is sufficient for planning purposes.

43.07 **band sensation level**

F: niveau de sensation dans le bande

S: nivel de sensación en la banda

Difference, expressed in decibels, between the sound integrated over a frequency band and the sound pressure level in that band at the threshold of audibility, there being no other disturbing sound.

43.08 **earcap reference plane**

F: plan de référence écouteur

S: plano de referencia auricular

That plane formed by the contacting points of a flat surface against a telephone earcap.

43.09 **earcap reference point (ECRP)**

F: point de référence écouteur (PRE)

S: punto de referencia auricular (PRA)

Point in the *earcap reference plane*, used as a reference parameter.

43.10 **ear reference point (ERP)**

F: point de référence oreille (PRO)

S: punto de referencia oído (PRO)

A point located at the entrance of the ear canal of the listener's ear. (See figure A-1/P.64).

43.11 **earphone coupling loss (L_E)**

F: affaiblissement de couplage de l'écouteur (L_E)

S: pérdida de acoplamiento del auricular (L_E)

That quantity defined as the receiving sensitivity of a handset (usually as a function of frequency) when applied to an artificial ear minus the receiving sensitivity of the same handset on a human ear.

43.12 **equivalent lip position**

F: position équivalente des lèvres

S: posición equivalente de los labios

A plane perpendicular to, and on the axis of, an aritifical mouth 6 mm in front of the virtual source position.

43.13 **guard-ring**

F: anneau de garde

S: anillo de guarda

Annular ring fitted, during tests, onto the transmitter housing of a telephone handset, to localize the sound source in a prescribed position relative to the microphone.

43.14 **metre air path**

F: trajet d'un mètre à l'air libre

S: trayecto de un metro en el aire

Measured reference of sound pressure loss over a 1 metre air path. In an anechoic environment, the sound pressure attenuation of such a path is approximately 30 dB measured from the MRP.

43.15 **modal distance**

F: distance modale

S: distancia modal

Distance between the centre of the microphone protective grid or front sound opening on a handset, and the centre of the guard-ring.

43.16 **modal gauge**

F: jauge modale

S: calibre modal

Template used to check a guard-ring position on a handset relative to the receiver *earcap reference plane.*

43.17 **modal position**

F: position modale

S: posición modal

Prescribed position and inclination of a handset relative to a fixed sound source.

43.18 **mouth reference point (MRP)**

F: point de référence bouche (PRB)

S: punto de referencia boca (PRB)

Point 25 mm in front of and on the axis of the lip position of a typical human mouth (or artificial mouth) (see Figure A-1/P.64).

43.19 **no sidetone line impedance (Z_{S0})**

F: impédance de ligne à effet local nul

S: impedancia de línea de efecto local nulo (Z_{S0})

That circuit impedance which, when connected across the terminals of a telephone set, causes the sidetone to be reduced to zero.

43.20 **occlusion effect**

F: effet d'occlusion

S: efecto de oclusión

The change in human sidetone that occurs when the ear canal is occluded, e.g. by a telephone receiver.

43.21 **obstacle effect (obstruction effect)**

F: effet d'obstacle; effet d'obstruction

S: efecto de obstáculo; efecto de obstrucción

The change in the acoustic field close to a human or artificial mouth as obstacles (e.g. telephone transmitter) are brought into close proximity.

43.22 **sidetone path loss**

F: affaiblissement du trajet d'effet local

S: atenuación del trayecto de efecto local

Any path, acoustic, mechanical or electrical by which a telephone user's speech and/or room noise is heard in his own ear(s) (at ERP) and expressed as a loss compared with the speech at MRP. Symbols in common use are:

L_{MEHS} for sidetone paths within a human head,

L_{MEST} for electro-acoustic sidetone paths within the telephone set,

L_{MEMS} for mechanical sidetone paths within a telephone handset.

43.23 sidetone loudness rating (STLR)

F: affaiblissement pour la sonie de l'effet local (ASEL)

S: índice de sonoridad para el efecto local (ISEL)

The loudness of a telephone sidetone path compared with the loudness of the intermediate reference system (IRS) overall.

43.24 sidetone balance network

F: réseau d'équilibrage d'effet local

S: red equilibradora del efecto local

An electrical network as part of a 2- to 4-wire balance point within a telephone set circuit for the purpose of controlling the telephone sidetone path loss.

43.25 sidetone masking rating (STMR)

F: affaiblissement d'effet local par la méthode de masquage (AELM)

S: índice de enmascaramiento para el efecto local (IEEL)

The loudness of a telephone sidetone path compared with the loudness of the intermediate reference system (IRS) overall in which the comparison is made incorporating the speech signal heard via the human sidetone path L_{MEHS} as a masking threshold.

43.26 sidetone power rating (STPR)

F: affaiblissement d'effet local par addition de puissances (AELP)

S: atenuación del efecto local por suma de potencias (AELP)

The loudness of a telephone sidetone path compared with the loudness of the intermediate reference system (IRS) overall in which the comparison is made incorporating the speech signal heard via the human sidetone path L_{MEHS} by power addition.

43.27 speech volume penalty

F: pénalisation en volume sonore

S: penalización en volumen sonoro

The reduction in a subscriber's talking level [usually expressed as a function of a speech sidetone rating, e.g. STRE (reference equivalent of sidetone), STMR, STLR] due to the presence of sidetone.

43.28 talking resistance

F: résistance de conversation

S: resistencia de conversación

Fixed resistance used for test purposes, which has a resistance equal to that of a carbon microphone at a particular current.

43.29 virtual source position

F: position de la source virtuelle

S: posición de la fuente virtual

That position within a human or artificial mouth at which emitted sounds appear to have their source.

43.30 virtual source function

F: fonction de source virtuelle

S: función de la fuente virtual

The change in virtual sources position as a function of some other parameter, e.g. frequency, proximity of obstances.

EFFECT OF TRANSMISSION IMPAIRMENTS

(Geneva, 1980; amended at Malaga-Torremolinos, 1984)

1 Purpose

An essential purpose of the present transmission plan for international connections is to provide guidance on the control of transmission performance. Such guidance is contained in Recommendations related to complete connections and to the constituent parts of a connection. These Recommendations contain performance objectives, design objectives and maintenance objectives, as defined in Recommendation G.102 for various transmission impairments which affect the transmission quality and customer opinion of transmission quality[1)] Typical transmission impairments include transmission loss, circuit noise, talker echo, sidetone loss, attenuation distortion, group-delay distortion and quantizing distortion. Although not under the control of the transmission planner, room noise is another important factor which should be considered.

This Recommendation is concerned with the effect of transmission parameters, such as those listed above, on customer opinion of transmission quality. It is based on information contributed in response to specific questions which have been studied by the CCITT. Much of this information is based on the results of subjective tests in which participants have talked, listened or conversed over telephone connections with controlled or known levels of the impairments and rated the tansmission quality on an appropriate scale. General guidance for the conduct of such tests is provided in Recommendation P.74. In addition, Recommendation P.77 provides guidance on the use of telephone user surveys to assess speech quality on international calls.

Specific purposes of this Recommendation are:

1) to provide a general, but concise, summary of the major transmission impairments and their effect on transmission quality which would serve as a central reference for transmission planners;

2) to provide for retention of basic information on transmission quality in support of relevant. Series P and Series G Recommendations with appropriate reference to these Recommendations and other sources of information such as Supplements and Questions under study;

3) to provide for the interim retention of basic information on transmission quality which is expected to be relevant in the formulation of future Recommendations.

§ 2 of this Recommendation provides a brief description of individual impairments which can occur in telephone connections, typical methods of characterization and general guidance on the acceptable levels of these impairments. More specific information is provided in Annexes to this Recommendation, in other Recommendations and in Supplements.

§ 3 of this Recommendation is concerned with the effect of combined impairments on transmission quality and the use of opinion models which permit estimates to be made of customer opinion as a function of combinations of transmission impairments in a telephone connection. Thus, they can be used to evaluate the transmission quality provided by the present transmission plan, the impact of possible changes in the transmission plan or the consequences of departures from the transmission plan. Such evaluations require certain assumptions concerning the constituent parts of a connection, and guidance is provided by the hypothetical reference connections which are the subject of Recommendations G.103 and G.104.

2 Effect of individual impairments

2.1 *General*

§ 2 describes invididually a number of the transmission impairments which can affect the quality of speech transmission in telephone connections. Information is provided on the general nature of each impairment, on methods which have been recommended to measure the impairment and on the acceptable ranges for the impairment. References are provided to Recommendations where more detailed information on measurement methods and recommended values can be found.

---

1) In this Recommendation, the term "impairment" is used in a general sense to refer to any characteristic or degradation in the transmission path which may reduce the performance or quality. It is not used to denote "equivalent loss" as was the case in some earlier CCITT texts.

2.2 *Loudness loss*

An essential purpose of a telephone connection is to provide a transmission path for speech between a talker's mouth and the ear of a listener. The loudness of the received speech signal depends on acoustic pressure provided by the talker and the loudness loss of the acoustic-to-acoustic path from the input to a telephone microphone at one end of the connection to the output of a telephone receiver at the other end of the connection. The effectiveness of speech communication over telephone connections and customer satisfaction depend, to a large extent, on the loudness loss which is provided. As the loudness loss is increased from a preferred range, the listening effort is increased and customer satisfaction decreases. At still higher value of loudness loss, the intelligibility decreases and it takes longer to convey a given quantity of information. On the other hand, if too little loudness loss is provided, customer satisfaction is decreased because the received speech is too loud.

Over the years, various methods have been used by transmission engineers to measure and express the loudness loss of telephone connections. The reference equivalent method is a subjective method which has been widely used in CCITT and is defined in Recommendations P.42 and P.72.

When Reference Equivalents (REs) have been used in Recommendations concerned with the loudness loss of connections, they have typically been stated in terms of the planning value of the overall reference equivalent of a complete connection which was defined for one direction of transmission as the sum of the following quantities:

- the nominal values of the reference equivalents of the sending and receiving local systems;
- the nominal value of the losses at 800 or 1000 Hz of the chain of lines and exchanges interconnecting the two local systems.

Because difficulties have been encountered in the use of reference equivalents, the planning value of the overall reference equivalent has been replaced by the Corrected Reference Equivalent (CRE) as defined in Recommendation G.111. This has required some adjustment in the recommended values of loudness loss for complete and partial connections.

Recommendations P.76, P.78 and P.79 provide information on subjective and objective methods for the determination of loudness ratings (LRs) which are currently under study. Eventually, these methods are expected to eliminate the need for the subjective determinations of loudness loss in terms of the corrected reference equivalent.

2.2.1 *Customer opinion*

Customer opinion, as a function of loudness loss, can vary with the test group and the particular test design. The opinion results presented in Table 1/P.11 are representative of laboratory conversation test results for telephone connections in which other characteristics such as circuit noise are contributing little impairment. These results indicate the importance of loudness loss control.

TABLE 1/P.11

Planning value of the overall corrected reference equivalent (dB)	Representative opinion results[a]	
	Percent "good plus excellent"	Percent "poor plus bad"
5 to 15	>90	<1
20	80	2
25	65	5
30	40	15

[a] Based on a composite opinion model (see Annex A).

2.2.2 *Recommended values of loudness loss*

Table 2/P.11 provides further information on selected values of loudness loss which have been recommended or are under study by the CCITT.

Note – Recommended values of loudness ratings are under study in Question 19/XII [1].

TABLE 2/P.11

Values (dB) of reference equivalent RE (q), corrected reference equivalent CRE (y) and loudness loss LR for various connections cited in Recommendations G.111 and G.121

		Previously recommended RE (q)	Presently recommended	
			CRE (y)	LR [b]
Optimum range for a connection (Rec. G.111 § 3.2)	min optimum max	6 9 18	5 [a] 7 [a] to 11 16	2 [a] about 5 [a] 11
traffic weighted mean values				
Long term objectives				
– connection (Rec. G.111 § 3.2)	min max	13 18	13 16	8 11
– national system send (Rec. G.121 § 1)	min max	10 13	11.5 13	6.5 8
– national system receive (Rec. G.121 § 1)	min max	2.5 4.5	2.5 4	−2.5 −1
Short term objectives				
– connection (Rec. G.111 § 3.2)	max	23	25.5	20.5
– national system send (Rec. G.121 § 1)	max	16	19	14
– national system receive (Rec. G.121 § 1)	max	6.5	7.5	2.5
Maximum values for national system (Rec. G.121 § 2.1) of an average-sized country	send receive	21 12	25 14	20 9
Minimum for the national sending system (Rec. G.121 § 3)		6	7	2

[a] These values apply for conditions free from echo; customers may prefer slightly larger values if some echo is present.

[b] See Note 1 of Recommendation G.121, § 2.1. If Administrations find difficulties in applying the values of LR in their network, they may report this in contributions to part B of Question 19/XII [1].

2.3 *Circuit noise*

The circuit noise in a telephone connection has a major effect on customer satisfaction and the effectiveness of speech communication. This noise may include white circuit noise and intermodulation noise from transmission systems as well as hum and other types of interference such as impulse noise and single frequency tones. Customer satisfaction depends on the power, the frequency distribution and the amplitude distribution of the noise. For a given type of noise, the satisfaction generally decreases monotonically with increasing noise power.

Circuit noise is generally expressed in terms of the indications given by a psophometer standardized by the CCITT in Recommendation O.41. With this apparatus, frequency-weighted measurements of noise power in dBmp can be made at various points in telephone connections.

2.3.1 *Opinion results*

Many tests have been conducted which demonstrate the effect of circuit noise on customer opinion. These tests have shown that opinion judgements of circuit noise are also highly dependent on the loudness loss of the connection and can be influenced by many other factors, particularly the room noise and sidetone loss.

The subjective effect of circuit noise measured at a particular point in a telephone connection depends on the electrical-to-acoustical loss or gain from the point of measurement to the output of the telephone receiver. As a convenience in assessing the contributions from different sources, circuit noise is frequently referred to the input of a receiving system with a specified receiving CRE or loudness rating. A common reference point is the input of a receiving system having a Receiving CRE of 0 dB. When circuit noise is referred to this point, circuit noise values less than −65 dBmp have little effect on transmission quality in typical room noise environments. Transmission quality decreases with higher values of circuit noise.

The opinion results presented in Table 3/P.11 are representative of laboratory conversation tests and illustrate the effect of circuit noise when other connection characteristics such as loudness are introducing little additional impairment. When the loudness loss is greater than the preferred range, the effect of a given level of circuit noise becomes more severe.

Note – The effect of circuit noise on transmission quality is under study in Question 4/XII [2] (see Annex A of this Recommendation for further information on the effects of circuit noise).

TABLE 3/P.11

Circuit noise at point of 0 dB receiving CRE (dBmp)[b]	Representative opinion results[a]	
	Percent "good plus excellent"	Percent "poor plus bad"
-65	90	<1
-60	85	1
-55	70	3
-50	50	10
-45	30	20

[a] Based on a composite opinion model (see Annex A).

[b] The noise values apply for a receiving CRE = 0.46 dB which has been rounded off to 0 dB for simplicity and to take account of the precision of the calculations.

2.3.2 *Recommended values of circuit noise*

Contributions to circuit noise from the various parts of a connection should be kept as low as practical. The major source of circuit noise on medium or long connections is likely to occur in analogue transmission facilities where the noise power is typically proportional to the circuit length. In Recommendation G.222, a noise objective of 10 000 pW0p or −50 dBm0p is recommended for the design of carrier transmission systems of 2500 km. When referred to a point of 0 dB receiving reference equivalent (assuming a loss of 6 to 12 dB), this corresponds to a noise level in the range from −62 to −56 dBmp, which is sufficiently high to affect the transmission quality.

The decrease in quality is larger on longer circuits or in connections with several such circuits in tandem. The CCITT states in Recommendation G.143 that it is desirable that the total noise generated by a chain of six international circuits should not exceed −43 dBm0p when referred to the first circuit in the chain. This corresponds to approximately −46 dBm0p at the end of the chain or −58 to −52 dBmp at a point with a 0 dB receiving reference equivalent. Other sources of circuit noise in international connections should be controlled such that their contribution is small compared to that permitted on analogue transmission facilities. Specific guidance is provided in a number of Recommendations.

The limits for a single tone or narrow bands of noise should be more stringent than the limits for wideband noise in order to avoid customer annoyance. As a general rule to limit annoyance from single frequency tones, the power in any individual tone should be 10 dB less than the psophometric noise power in the circuit. To avoid audibility, an additional 5 dB of margin is recommended where practical [3].

2.4 *Sidetone loudness loss*

Sidetone loudness loss is the loudness loss of the acoustic-to-acoustic transmission path from the telephone microphone to the telephone receiver in the same telephone set. Thus, the sidetone loudness loss defines one of the paths through which the talker hears himself as he speaks. Other such paths are the head conduction path and the acoustic path from the mouth to the ear through ear cap leakage. The presence of these other paths affects the customer's perception of the sidetone loudness loss and consequently his reaction to it.

Sidetone loudness loss affects telephone transmission quality in several ways. Too little sidetone loudness loss causes the returned speech levels to be too loud and this reduces customer satisfaction. Another aspect of insufficient sidetone loudness loss is that talkers tend to reduce their speech levels and/or move the handset away from the mouth, thus reducing the received levels at the far end of the connection. Handset movement can also reduce the seal at the ear and so make it easier for room noise to reach the ear through the resulting leakage path, as well as reducing the level of the received signal from the far end of the connection. In addition, the sidetone path provides another route by which room noise can reach the ear. Very low levels of sidetone loudness loss can affect transmission quality adversely. As the sidetone loudness loss is increased there is a general region of preferred loss values. Excessive sidetone loss can make a telephone set sound dead as one is talking and, for many connections, the absence of sidetone would not be a preferred condition.

Sidetone loudness loss can be rated in much the same manner as connection loudness loss, and this is frequently done in terms of sidetone reference equivalent (Recommendation P.73), but it may also be done in terms of a sidetone rating method (Recommendation P.76) which takes into account the head conduction and direct acoustic paths in terms of the overall effect on the customer, and appears to yield ratings which correlate better with the subjective effects of talker sidetone than sidetone reference equivalent.

The sidetone loudness loss is influenced by the telephone set design and the impedance match between the telephone set and the subscriber line. Impedance variations at the far end of the subscriber line can also have significant mismatch effects on short subscriber lines with low loss. Impedance mismatches at other points in the connection will also affect the returned signal but, as the delay in the return path becomes significant, the effect is generally considered as talker echo. (See § 2.9.).

2.4.1 *Recommended values of sidetone loudness loss*

Recommendation G.121, § 5 provides guidance on the sidetone reference equivalent. Tests have shown that a value of at least 17 dB is desirable under some adverse conditions. This is not easily achieved and values between 7 and 10.5 dB are to be expected in most cases.

Subjective test results of customer opinion as a function of sidetone loss in terms of reference equivalents indicate a preferred range of 10 to 15 dB and for sidetone masking rating (STMR) of 7 to 10 dB (see Supplement No. 11). Lower values cause a substantial reduction in customer opinion and should be avoided. High values, up to 20 dB, are acceptable in the absence of talker echo. With moderate levels of talker echo, sidetone loss greater than 10 to 15 dB tends to increase the adverse opinion due to talker echo [3].

Note 1 – Suitable values for sidetone ratings in accordance with Recommendation P.76 (STMR) are under study in Question 9/XII [4].

Note 2 – Sidetone subjective test results are presented in COM XII-No. 158 (1981-1984 Study Period) and predictions for the effect of sidetone are included in the opinion model of Supplement No. 3 at the end of this volume (see also Supplement No. 11).

2.5 *Room noise*

Room noise is the term used to describe the background noise in the environment of the telephone set. In a residential location it may consist of household appliances, radio or phonograph noise, conversations or street noise. In an office location, business equipment, air conditioning equipment and conversations are likely to predominate. In many situations, the effect of room noise may be inconsequential compared to the effects of circuit noise. In noisy locations such as call offices in public places, however, the effects of room noise may have a substantial effect on the ease of carrying on a conversation or even in being able to hear and understand properly.

Room noise can manifest itself in several ways. One is through leakage around the earcap of the receiver. Another is through the sidetone path of the telephone set if the sidetone loudness loss is sufficiently low in comparison with leakage past the earcap. A third way is through the other ear, although the effect of this on telephone reception is usually less than that of noise entering the "telephone ear", unless the sound in the room causes distraction (a baby crying, for example). A fourth way is through the transmitter over the connection to the receiving telephone set.

The previous discussion applies primarily to conventional telephone sets. Loudspeaking telephone sets are more susceptible to room noise.

Noise present in stationary or moving vehicles (not commonly referred to as a room noise) may also have a substantial effect on the ease of carrying on a conversation or in being able to hear and understand properly over telephone connections involving mobile station.

Note 1 – The effect of room noise on customer opinion is under study in connection with circuit noise in Question 4/XII [2]; the effect of noise in vehicles is under study in connection with Question 24/XII [5].

Note 2 – Supplement No. 13 at the end of this volume describes noise spectra which can be used in evaluation of tranmission performance. Spectra for two different environments are described, one for room noise and two for internal vehicle noise, depending on whether the vehicle is stationary or moving.

2.6 *Attenuation distortion*

Attenuation distortion is characterized by transmission loss (or gain) at other frequencies relative to the transmission loss at 800 or 1000 Hz. Thus, attenuation distortion includes the low-frequency and high-frequency rolloffs which determine the effective bandwidth of a telephone connection, as well as in-band variations in loss as a function of frequency. The loudness loss and articulation of a telephone connection are respectively a function of the attenuation distortion. Even when the loudness loss is maintained at a constant value, opinions of the transmission quality as determined by subjective tests usually get worse as the amount of attenuation distortion increases.

The effect of attenuation distortion on loudness is greater at the lower end of the frequency band than at the higher end. The effect of attenuation distortion on sound articulation is, on the contrary, more marked at the higher frequencies. For both loudness and articulation impairments due to bandpass characteristics, it can be assumed that the impairment values due to highpass and lowpass characteristics add directly if each attenuation distortion slope is greater than 15 dB/octave.

The effect of attenuation distortion on listening and conversation opinion scores decreases noticeably as the overall loudness loss of a connection increases, particularly when circuit noise also exists. The effect of attenuation distortion on opinion scores is typically less than that of loudness loss, particularly at high values of loudness loss, but may be comparable to that of noise when the values of loudness loss and noise are both low.

The current network performance objectives for attenuation distortion in the electrical transmission elements of a worldwide 4-wire chain of 12 circuits are given in Recommendation G.132 but, of course, the frequency characteristics of the telephone sets themselves have some influence.

Note – Further information on the effects of attenuation distortion on transmission quality are provided in Annex B. Recommended objectives are under study in Question 14/XII [6].

2.7 *Group-delay distortion*

Group-delay distortion is characterized by the group delay at other frequencies relative to the group delay at the frequency where the group delay has its minimum value. Although the effect of group-delay distortion is usually a more significant impairment for data transmission than for speech transmission, large amounts of group-delay distortion can cause noticeable distortion for speech signals.

The effect of group-delay distortion at the upper and lower edges of the transmitted band can be described as "ringing" and "speech blurred", respectively. In the absence of noise or attenuation distortion, the effect is conspicuous throughout the entire range of typical loudness loss values. However, the effect in a typical 4-wire circuit chain is usually not serious since the group-delay distortion is normally accompanied by closely related attenuation distortion which tends to reduce the effect.

The current performance objectives for group-delay distortion for a worldwide chain of 12 circuits are given in Recommendation G.133.

Note – Further information on the effect of group-delay distortion is provided in Annex C.

2.8 *Absolute delay*

Values of absolute delay typical of those present in terrestrial transmission facilities have little effect on speech transmission quality if there is no talker or listener echo (4-wire connections, for example) or if the talker and listener echo are adequately controlled. Satellite facilities introduce larger amounts of delay (approximately 300 ms in each direction of transmission) and, again, the available opinion data indicates that there is little effect on the transmission quality of connections with a single satellite circuit, provided talker and listener echo are adequately controlled. Less data are available on the effects of one-way delays of approximately 600 ms (two satellite circuits in tandem) and the results are not entirely consistent. Therefore, caution is recommended with regard to the introduction of one-way absolute delay significantly greater than 300 ms.

Note – The effects of echo, echo control and propagation time are under study in Questions 6/XII and 11/XII [7].

2.9 *Talker echo*

Talker echo occurs when some portion of the talker's speech signal is returned with enough delay (typically more than about 30 ms) to make the signal distinguishable from normal sidetone. Talker echo may be caused by reflections at impedance mismatches or by other processes such as go-to-return crosstalk. The effect of talker echo is a function of the loss in the acoustic-to-acoustic echo path and the delay in the echo path. In general, customer satisfaction is decreased as the loss of the echo path is decreased or the delay of the echo path is increased.

Loss in the echo path is frequently expressed in terms of the corrected reference equivalent (CRE) of the talker echo path. This is defined in Recommendation G.131 as the sum of:

- the values of the transmission loss in the two directions of transmission between the 2-wire end of the talking subscriber's line in the terminal local exchange and the 2-wire terminals of the 4-wire – 2-wire terminating set at the listener's end;
- the value of the echo balance return loss at the listener's end; and
- the simultaneous-minimum sending and receiving CREs of subscriber's telephone sets and lines at the talker's local exchange.

Echo tolerance curves are provided in Figure 2/G.131 which indicate the recommended CRE of the echo path to control the probability of objectionable echo.

Note – The effect of echo and propagation time is under study in Question 6/XII [7].

2.10 *Listener echo*

Listener echo refers to a transmission condition in which the main speech signal arrives at the listener's end of the connection accompanied by one or more delayed versions (echoes) of the signal. Such a condition can occur as the result of multiple reflections in the transmission path. A simple, yet common, source of listener echo is a low loss 4-wire transmission path which interconnects two 2-wire subscriber lines. In such a connection, reflections can occur as the result of impedance mismatch at the hybrids at each end of the 4-wire section. A portion of the main speech signal can thus be reflected at the far end of the 4-wire path, return to the near end and be reflected again. The result is a listener echo, whose magnitude, relative to the main signal, depends on the two return losses and the two-way loss or gain of the 4-wire transmission path. The delay of the echo is determined primarily by the two-way delay of the 4-wire transmission path. For small delays, the listener echo results in a change in the spectral quality of the speech. For longer delays, the echo is more pronounced and is sometimes referred to as a "rain barrel" effect.

Listener echo may be characterized by the additional loss and additional delay in the listener echo path relative to that in the main signal path. The minimum value of the additional listener echo path loss over the frequency band of interest provides a margin against instability or oscillation. As a result, listener echo is frequently referred to as near singing distortion. Recommendation G.122 provides guidance on the influence of national networks on stability in international connections.

Note – The study of listener echo, Question 5/XII for the 1981-1984 Study Period, was essentially completed [8]. These results are taken into account in Recommendation G.122.

2.11 *Nonlinear distortion*

Nonlinear distortion occurs in systems in which the output is not linearly related to the input. A simple example is a system in which the output signal can be represented by, as a function of the input signal $e_i(t)$, a power series of the form,

$$e_o(t) = a_1 e_i(t) + a_2 e_i^2(t) + a_3 e_i^3(t) + \dots,$$

which, in the case of a sinusoidal input, creates second and third harmonics in the output. For more complex signals, the nonlinear terms are frequently referred to as intermodulation distortion. Nonlinear distortion is normally more significant for data transmission than it is for speech transmission. At present, one of the major sources of nonlinear distortion in telephone connections is from telephone sets with carbon microphones. However, other devices such as syllabic compandors and overloaded amplifiers may be significant contributors.

Note – Further information is provided in Annex D. Nonlinear distortion of telephone apparatus is being studied under Question 13/XII [9].

2.12 *Quantizing distortion*

Quantizing distortion occurs in digital systems when an analogue signal is sampled and each sample is encoded into one of a finite set of values. The difference between the original analogue signal and that which is recovered after quantizing is called quantizing distortion or quantizing noise. For many digital encoding algorithms, such as A-law or μ-law PCM, which have a nearly-logarithmic companding law, the subjective effect of quantizing distortion can be approximated by adding signal-correlated noise (white noise which has been modulated by the speech signal). Such a signal can be generated in a modulated-noise reference unit which can be adjusted to provide a reference signal with a selected and nearly constant signal to signal-correlated-noise ratio. Recommendation P.70 describes the modulated-noise reference unit recommended by CCITT for use in evaluating digital codecs for telephone speech applications.The signal to signal-correlated-noise ratio, when expressed in dB, is called Q. The effective Q of an unknown digital system can be determined by subjective comparison with the modulated-noise reference unit. (Supplement No. 14 provides guidelines on use of the modulated noise reference unit of Recommendation P.70.)

Subjective test results have been reported by some Administrations which have evaluated the effects of both circuit noise and Q on customer opinion. Results from tests of this type permit estimates to be made of the circuit noise level, which could provide approximately the same transmission quality ratings as a given level of quantizing distortion.

Note – Further information is provided in Annex E. The transmission performance of digital systems is under study in Question 18/XII [10].

2.13 *Phase jitter*

Phase jitter occurs when the desired signal, during transmission, is phase or frequency modulated at a low-frequency rate. If such distortion is present in sufficient quantity, the transmission quality is degraded. Table 4/P.11 summarizes the threshold data for single-frequency phase jitter which have been reported by one Administration. The results are in terms of the mean threshold expressed in terms of the signal-to-first order-sideband (C/SB) ratio in dB. The average standard deviation across subjects was about 4 dB.

TABLE 4/P.11

Phase jitter modulation rate (Hz)	Mean threshold C/SB ratio (dB)	
	Male talkers	Female talkers
25	10.9	13.8
80	14.4	16.3
115	12.3	18.3
140	13.8	20.0
200	17.0	18.0

2.14 *Intelligible crosstalk*

Intelligible crosstalk occurs when the speech signal from one telephone connection is coupled to another telephone connection such that the coupled signal is audible and intelligible to one or both of the participants on the second telephone connection. Although the level of the intelligible crosstalk may be high enough to degrade the transmission quality, the major concern is the loss of privacy.

A number of factors influence the intelligibility of a signal which is coupled from one telephone connection to another. They include the characteristics of the telephone apparatus (including sidetone), circuit noise, room noise, the coupling loss, the interfering talker's speech level and the hearing acuity of the listener.

Information is provided in Recommendation P.16 on the intelligibility threshold for crosstalk and on methods for calculating the probability of intelligible crosstalk. Design objectives for the various apparatus in telephone connections should be selected such that the probability of intelligible crosstalk is sufficiently low. Typically, objectives are intended to keep the probability below one percent in connections where the interfering and interfered-with parties are unlikely to know each other and unlikely to suffer the same coupling again. A more stringent objective of 0.1 percent is typical for use in local equipment such as subscriber lines where the two parties may be neighbours.

3 Effect of multiple impairments and the use of opinion models

Transmission performance of a practical connection can be affected by several transmission impairments which are likely to coexist. Although results for customer opinion in the form described in § 2 are useful in many studies involving one or two types of transmission impairments, they become increasingly cumbersome as the number of impairments under study increases. This has led to the study of more extensive analytical models of customer opinion which can be based on the composite results of a number of individual tests and studies. The formulation and use of these more comprehensive models are aided by the availability of modern digital computers. Ideally, such models might eventually include the effects of all or most of the significant types of transmission impairment mentioned in § 2 above.

Note – Although some Administrations have reported on efforts directed toward this goal, the subject of models for predicting transmission quality from objective measurements is still under study in Question 7/XII [11]. Examples of opinion models used by the Bell Communications Research, Inc. and by British Telecom are described in Supplements Nos. 3 and 4 at the end of this Volume.

ANNEX A

(to Recommendation P.11)

Composite opinion model

Laboratory test results showing the subjective effects of circuit noise and overall reference equivalent were received during the 1973-1976 Study Period from the United Kingdom Post Office (UK), Italy, the USSR Telecommunication Administration (USSR) and the American Telephone and Telegraph Company (AT&T) (USA) [12, 13, 14]. (This information was contributed to the study of Question 4/XII, effect of circuit noise on transmission performance.) These data were compared to each other to an AT&T opinion model for the effects of circuit noise and overall reference equivalent [15]. This comparison led to the formulation of a composite opinion model.

This Annex presents comparisons between performance estimates obtained using the composite opinion model and the laboratory test data [16]. The comparison shows reasonably good agreement. For this reason, the information is retained even though it is in terms of reference equivalents rather than corrected reference equivalents now recommended for planning purposes by the CCITT Recommendation G.111. It is expected that this annex will be revised when updated information in terms of corrected reference equivalents becomes available.

Test data from each of the four laboratory conversation tests are shown in Figures A-1/P.11 and A-2/P.11 in terms of percent of ratings which were "good plus excellent" and percent ratings which were "poor plus bad" as a function of circuit noise. The overall reference equivalent for each test condition is also identified. The solid curves show the corresponding results from the composite opinion model. These comparisons indicate that some of the data points fall above the composite model while other data points fall below it. In some cases, the differences are systematic for a complete test or for certain values of overall reference equivalent. As a result, a first conclusion is reached that no single set of opinion curves will provide exact prediction of test results from different tests and test groups in the various Administrations.

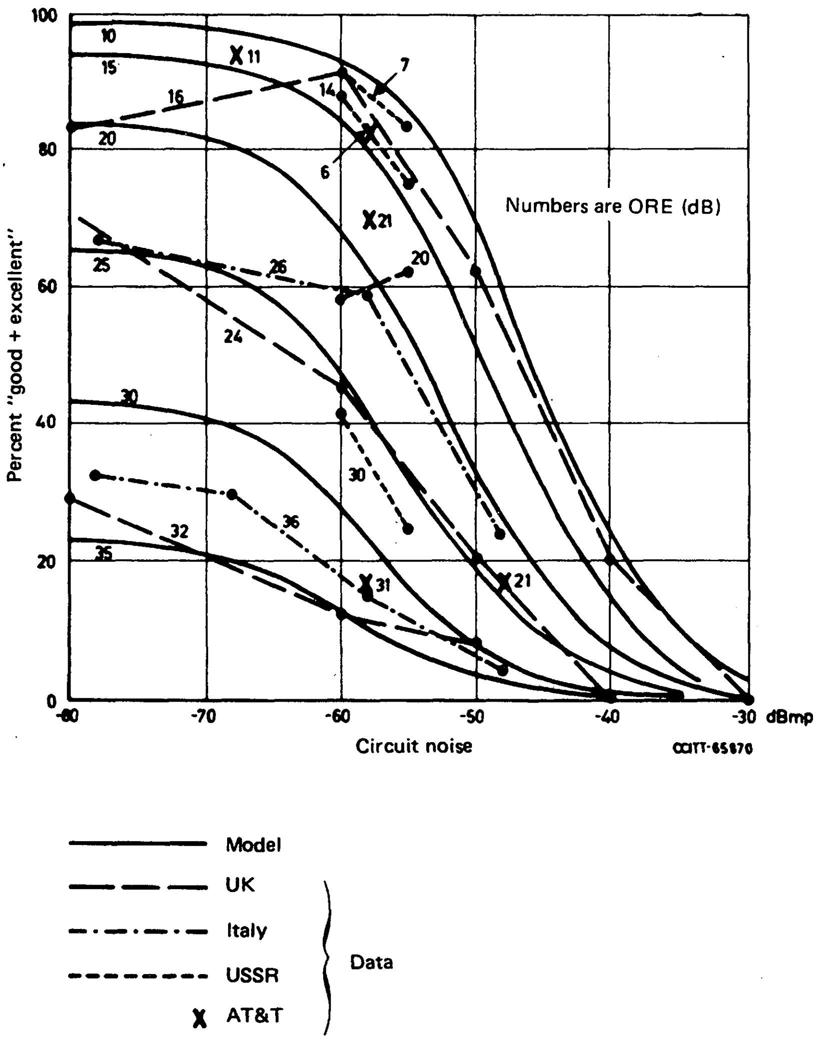


FIGURE A-1/P.11

Comparison of data model in percent "good + excellent"

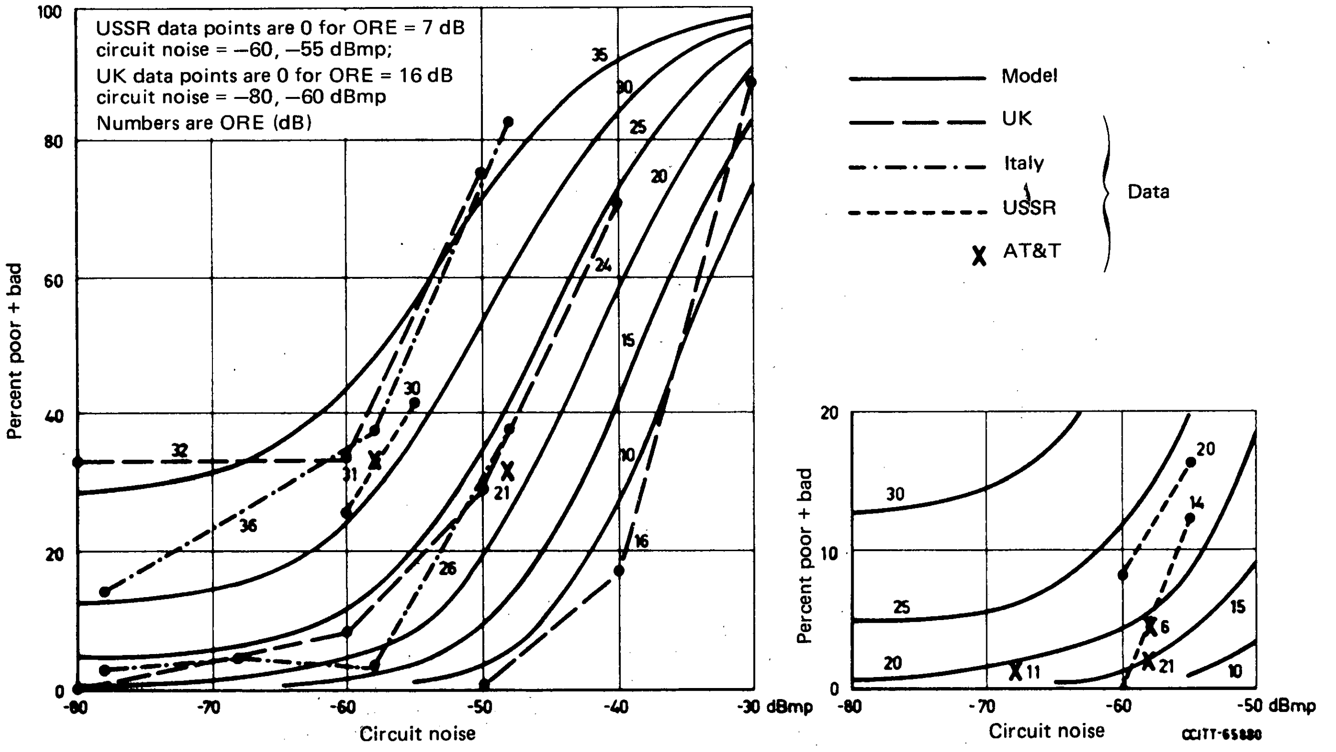


FIGURE A-2/P.11

Comparison of data and model in percent "poor + bad"

The differences between the test data and the composite opinion model are in most cases less than 10 per cent and the trends, as a function of circuit noise, are similar in the data and in the model. Thus, a second conclusion is that the composite opinion model for circuit noise and reference equivalent provides estimates of opinion which are representative of those which can be expected in actual subjective tests. However, this composite model should not be expected to accurately predict the results of any particular test and some systematic differences should be anticipated between the model and test results from a particular test or a particular Administration.

If the opinions predicted by the composite opinion model are considered as representative, rather than absolute, the model is judged to be acceptable and useful for estimating the incremental effects of loudness loss and circuit noise on customer opinion. Therefore, this model can be recommended provisionally for studying transmission performance of telephone connections, subject to the above conclusions and recognizing that the model is in terms of reference equivalents. This Recommendation is not intended to preclude the use of other models or data which may be judged to be more appropriate for specific subject groups or Administrations.

The composite opinion model expresses overall reference equivalent and circuit noise in terms of transmission ratio R_{LN}. The expression is given as Equation (A-1):

$$R_{LN} = -34.88 - 2.257 \sqrt{(L'_e - 8.2)^2 + 1} - 2.0294 N'_F + 1.833 L'_e + 0.02037 L'_e N'_F \qquad \text{(A-1)}$$

where

L'_e is the reference equivalent of an overall telephone connection (in dB)

N'_F is the total effective noise (in dBmp) obtained by the power addition of N' and -62.63 dBmp[2)]

N' circuit noise (in dBmp) at the input to a telephone set with a 0 dB receiving reference equivalent.

The expressions for the proportion of ratings in the categories "Good + Excellent ($G + E$)" and "Poor + Bad ($P + B$)" are:

$$G + E = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{A} e^{-\frac{t^2}{2}} dt \qquad \text{(A-2)}$$

$$P + B = \frac{1}{\sqrt{2\pi}} \int_{B}^{\infty} e^{-\frac{t^2}{2}} dt \qquad \text{(A-3)}$$

For the composite opinion model:

$$A = \frac{R - 62}{15}; \qquad B = \frac{R - 43}{15} \qquad \text{(A-4)}$$

Values of percent $G + E$ and $P + B$ are given in Tables A-1/P.11 and A-2/P.11 respectively. These tables were generated using Equations (A-1) to (A-4) above.

Contours of fixed percent $G + E$ and $P + B$ are given in Figures A-3/P.11 and A-4/P.11.

2) This value was determined from analysis of the conversation test results on which the original model was based. Lower or higher values of this constant may be appropriate for very low or very high values of room noise (see Supplement No. 3 at the the end of this Volume).

TABLE A-1/P.11

Percent "good + excellent" as a function of overall reference equivalent and circuit noise

Overall reference equivalent (dB) \ Circuit noise [a] (dBmp)	−80	−75	−70	−65	−60	−55	−50	−45	−40	−35	−30
0	77.66	76.98	74.94	69.25	56.35	36.15	16.64	5.22	1.09	0.15	0.01
1	82.82	82.25	80.51	75.55	63.77	43.77	22.22	7.88	1.89	0.30	0.03
2	87.15	86.68	85.24	81.07	70.69	51.62	28.73	11.44	3.15	0.59	0.07
3	90.64	90.27	89.13	85.74	76.91	59.38	36.00	16.01	5.02	1.08	0.16
4	93.36	93.08	92.20	89.54	82.29	66.74	43.78	21.60	7.66	1.90	0.32
5	95.41	95.20	94.54	92.52	86.75	73.45	51.72	28.12	11.21	3.18	0.63
6	96.88	96.73	96.25	94.75	90.30	79.27	59.43	35.30	15.71	5.05	1.14
7	97.88	97.77	97.43	96.33	92.93	83.96	66.34	42.57	20.90	7.54	1.95
8	98.42	98.33	98.07	97.20	94.47	86.93	71.16	48.22	25.45	10.02	2.86
9	98.41	98.33	98.06	97.21	94.52	87.12	71.68	49.10	26.38	10.65	3.15
10	98.06	97.96	97.66	96.67	93.65	85.62	69.53	46.87	24.80	9.90	2.90
11	97.54	97.42	97.05	95.88	92.39	83.51	66.57	43.85	22.65	8.84	2.55
12	96.85	96.71	96.26	94.88	90.84	81.03	63.26	40.61	20.43	7.79	2.20
13	96.00	95.83	95.29	93.65	89.03	78.26	59.75	37.34	18.27	6.79	1.88
14	94.95	94.75	94.11	92.20	86.94	75.22	56.11	34.11	16.23	5.89	1.60
15	93.68	93.44	92.70	90.48	84.57	71.95	52.38	30.96	14.33	5.07	1.35
16	92.17	91.90	91.03	88.50	81.92	68.47	48.62	27.94	12.57	4.34	1.13
17	90.40	90.08	89.10	86.23	79.00	64.80	44.85	25.05	10.97	3.69	0.94
18	88.35	87.98	86.87	83.67	75.80	60.98	41.12	22.32	9.51	3.13	0.78
19	85.99	85.58	84.34	80.82	72.36	57.04	37.47	19.77	8.20	2.63	0.65
20	83.33	82.88	81.51	77.68	68.69	53.03	33.92	17.40	7.03	2.21	0.53
21	80.36	79.87	78.38	74.25	64.82	48.99	30.50	15.22	5.99	1.84	0.44
22	77.09	76.56	74.96	70.57	60.78	44.95	27.24	13.22	5.07	1.53	0.36
23	73.54	72.97	71.27	66.67	56.62	40.96	24.17	11.41	4.27	1.26	0.29
24	69.72	69.12	67.34	62.57	52.39	37.06	21.29	9.79	3.58	1.03	0.23
25	65.66	65.04	63.20	58.32	48.13	33.29	18.62	8.34	2.98	0.84	0.19
26	61.42	60.79	58.91	53.97	43.89	29.68	16.17	7.05	2.46	0.69	0.15
27	57.04	56.40	54.50	49.58	39.71	26.25	13.94	5.93	2.02	0.55	0.12
28	52.57	51.93	50.04	45.19	35.65	23.05	11.92	4.95	1.65	0.44	0.10
29	48.06	47.43	45.58	40.85	31.75	20.07	10.12	4.10	1.34	0.36	0.08
30	43.58	42.96	41.17	36.62	28.04	17.34	8.53	3.38	1.08	0.28	0.06
31	39.18	38.59	36.87	32.55	24.55	14.85	7.13	2.76	0.87	0.22	0.05
32	34.91	34.35	32.72	28.68	21.31	12.62	5.92	2.24	0.69	0.18	0.04
33	30.83	30.30	28.78	25.04	18.33	10.63	4.87	1.81	0.55	0.14	0.03
34	26.97	26.48	25.08	21.66	15.63	8.88	3.98	1.45	0.43	0.11	0.02
35	23.36	22.92	21.65	18.56	13.20	7.35	3.23	1.15	0.34	0.08	0.02

[a] At input to set, with 0 dB receiving reference equivalent.

TABLE A-2/P.11

Percent "good + bad" as a function of overall reference equivalent and circuit noise

Overall reference equivalent (dB) \ Circuit noise [a] (dBmp)	−80	−75	−70	−65	−60	−55	−50	−45	−40	−35	−30
0	2.13	2.25	2.62	3.84	7.69	18.08	38.28	63.95	84.77	95.55	99.13
1	1.34	1.42	1.67	2.51	5.27	13.35	30.78	55.83	79.09	93.02	98.41
2	0.82	0.87	1.03	1.59	3.51	9.56	24.03	47.48	72.34	89.50	97.23
3	0.49	0.52	0.62	0.98	2.26	6.63	18.19	39.25	64.68	84.87	95.41
4	0.28	0.30	0.36	0.58	1.42	4.46	13.35	31.52	56.42	79.07	92.75
5	0.16	0.17	0.21	0.34	0.86	2.92	9.51	24.59	47.95	72.21	89.08
6	0.09	0.09	0.12	0.19	0.51	1.87	6.61	18.69	39.74	64.56	84.34
7	0.05	0.05	0.07	0.11	0.31	1.19	4.57	14.02	32.39	56.76	78.77
8	0.03	0.03	0.04	0.07	0.21	0.84	3.40	11.08	27.22	50.55	73.74
9	0.03	0.03	0.04	0.07	0.21	0.82	3.29	10.67	26.27	49.14	72.35
10	0.04	0.05	0.06	0.10	0.26	0.99	3.77	11.74	27.90	50.83	73.53
11	0.06	0.07	0.08	0.13	0.35	1.25	4.51	13.31	30.29	53.34	75.33
12	0.09	0.09	0.11	0.19	0.47	1.60	5.42	15.17	32.99	56.08	77.25
13	0.13	0.14	0.16	0.26	0.63	2.03	6.51	17.26	35.88	58.89	79.17
14	0.18	0.19	0.23	0.36	0.84	2.57	7.77	19.57	38.91	61.71	81.02
15	0.26	0.27	0.33	0.50	1.12	3.23	9.23	22.07	42.03	64.49	82.79
16	0.36	0.39	0.45	0.68	1.47	4.03	10.90	24.76	45.23	67.22	84.45
17	0.51	0.53	0.62	0.92	1.91	4.98	12.77	27.63	48.47	69.87	86.02
18	0.70	0.73	0.85	1.23	2.46	6.11	14.86	30.66	51.72	72.43	87.48
19	0.95	0.99	1.14	1.63	3.14	7.43	17.18	33.84	54.98	74.88	88.83
20	1.27	1.33	1.52	2.13	3.97	8.97	19.71	37.14	58.20	77.21	90.08
21	1.69	1.77	2.01	2.76	4.98	10.73	22.46	40.54	61.37	79.42	91.22
22	2.23	2.32	2.62	3.53	6.17	12.72	25.43	44.01	64.47	81.50	92.26
23	2.90	3.02	3.38	4.48	7.59	14.96	28.58	47.54	67.48	83.45	93.21
24	3.73	3.87	4.31	5.62	9.23	17.45	31.91	51.09	70.39	85.25	94.07
25	4.75	4.91	5.44	6.99	11.13	20.20	35.40	54.63	73.16	86.92	94.83
26	5.97	6.17	6.79	8.59	13.29	23.18	39.01	58.13	75.81	88.46	95.52
27	7.44	7.67	8.38	10.45	15.72	26.40	42.72	61.57	78.30	89.87	96.13
28	9.16	9.43	10.25	12.60	18.43	29.83	46.50	64.92	80.64	91.14	96.68
29	11.16	11.46	12.39	15.03	21.42	33.45	50.31	68.17	82.82	92.29	97.16
30	13.46	13.80	14.84	17.75	24.67	37.23	54.12	71.28	84.84	93.33	97.57
31	16.06	16.44	17.59	20.77	28.16	41.14	57.89	74.23	86.69	94.25	97.94
32	18.97	19.39	20.64	24.07	31.88	45.14	61.60	77.03	88.38	95.07	98.26
33	22.19	22.64	23.98	27.64	35.79	49.20	65.19	79.64	89.91	95.80	98.54
34	25.69	26.17	27.60	31.46	39.86	53.26	68.66	82.07	91.29	96.43	98.77
35	29.47	29.97	31.47	35.48	44.04	57.29	71.97	84.31	92.53	96.99	98.98

[a] At input to set, with 0 dB receiving reference equivalent.

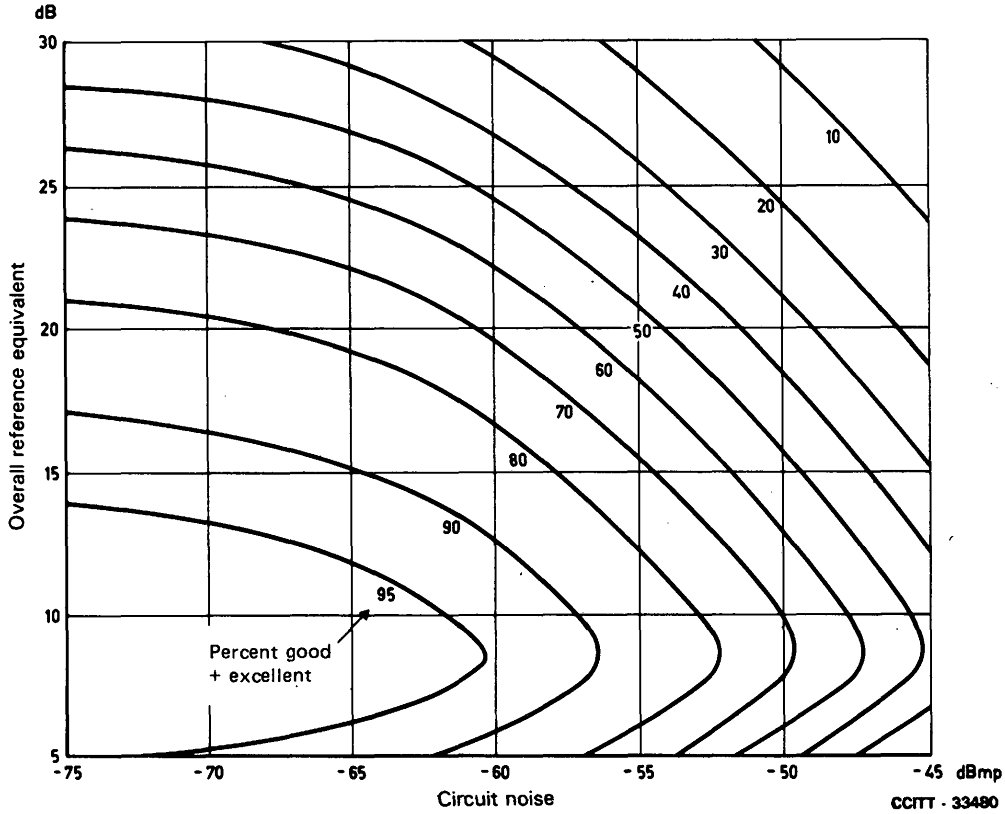


FIGURE A-3/P.11

Composite opinion model with parameter "good + excellent" in percent

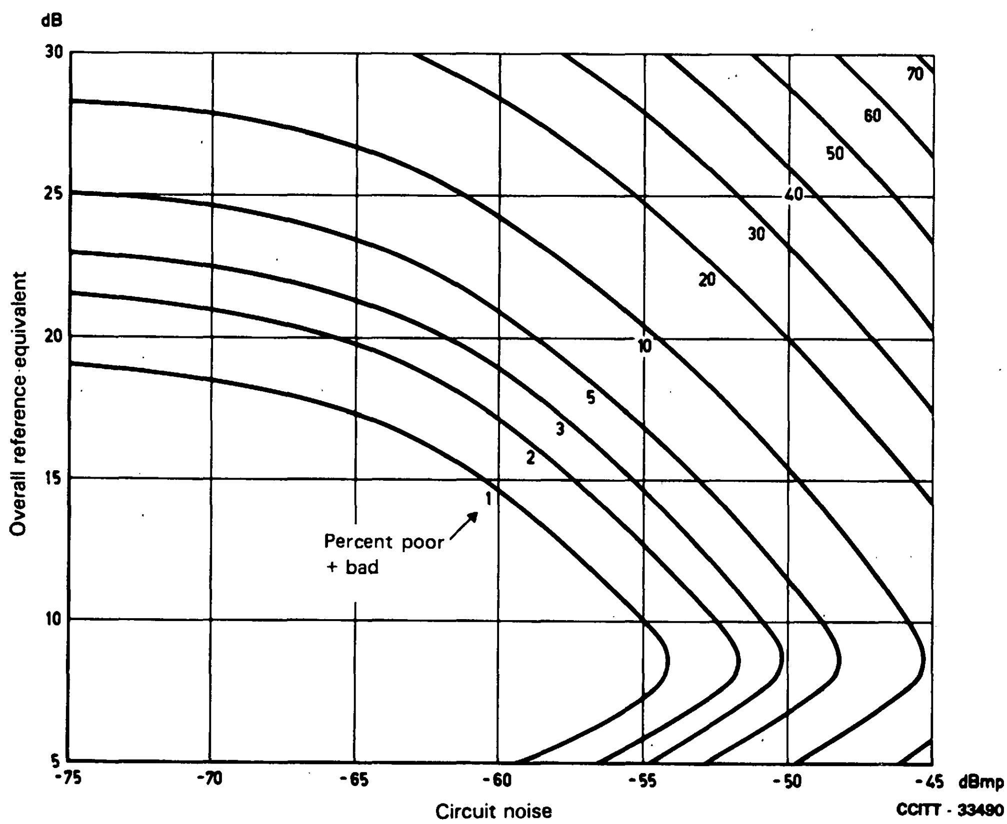


FIGURE A-4/P.11

Composite opinion model with parameter "poor + bad" in percent

ANNEX B

(to Recommendation P.11)

Effects of attenuation distortion on transmission performance

B.1 *Effect of attenuation distortion on loudness and articulation*

The effect of attenuation distortion on loudness is more marked at a lower frequency band than at a higher one.

The effect of attenuation distortion on sound articulation is, on the contrary to loudness, more marked at a higher frequency band than at a lower one. Attenuation distortion equivalent values (I_L) and articulation equivalent loss values (I_A) are equivalent loss difference values referred to a system without frequency band restriction.

For both attenuation distortion equivalent and articulation equivalent loss values due to bandpass characteristics, it can be assumed that an additivity law of impairment values due to highpass and lowpass characteristics holds true, if each attenuation slope is steeper than 15 dB/octave.

These phenomena are induced based on the calculation and subjective test study results as shown in Figures B-1/P.11, B-2/P.11, B-3/P.11 and B-4/P.11.

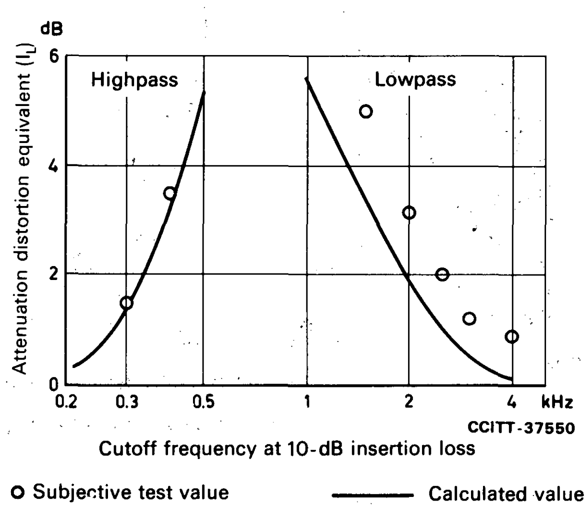
Note – Attenuation distortion equivalent and articulation equivalent loss described here are determined in reference to a complete telephone speech path without attenuation distortion junction.

B.2 *Effect of attenuation distortion on listening and conversation opinion scores*

The effect of attenuation distortion on listening and conversation opinion scores increases noticeably as the overall loudness loss of a connection decreases. This tendency can be more marked when circuit noise exists.

The effect of attenuation distortion on opinion scores is somewhat less than that of loudness loss, which is always dominant at any, particularly high overall loudness loss. However, its effect seems to be comparable to, or even larger than, that of noise under certain conditions, especially in connections of lower overall loudness loss.

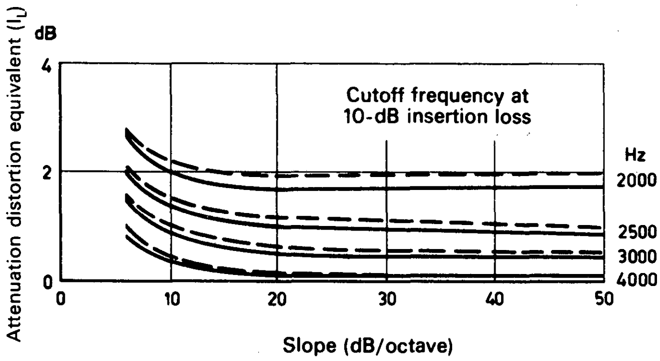
See Figures B-5/P.11, B-6/P.11, B-7/P.11 and Table B-1/P.11.



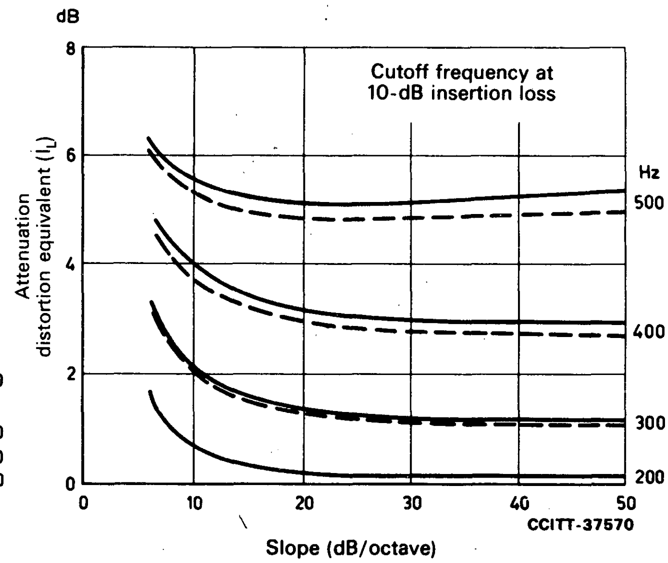
Note – Slope of lowpass and highpass filters is 48 dB/octave.

FIGURE B-1/P.11

Cutoff frequency effect on loudness



a) Lowpass filter effect



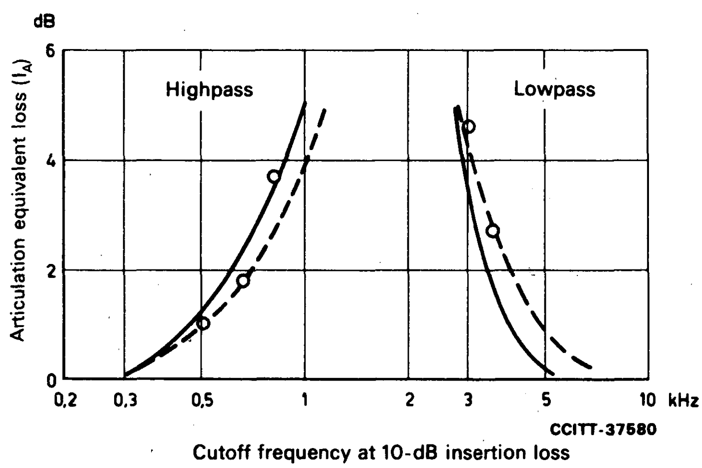
b) Highpass filter effect

——— With 0.4-mm copper subscriber line (7-dB image loss at 1500 Hz)

– – – – Without subscriber line

FIGURE B-2/P.11

Lowpass and highpass filter slope effect on loudness



o Subjective test value

——— Calculated value for the telephone system

– – – – Calculated value for ARAEN system

FIGURE B-3/P.11

Cutoff frequency effect on articulation

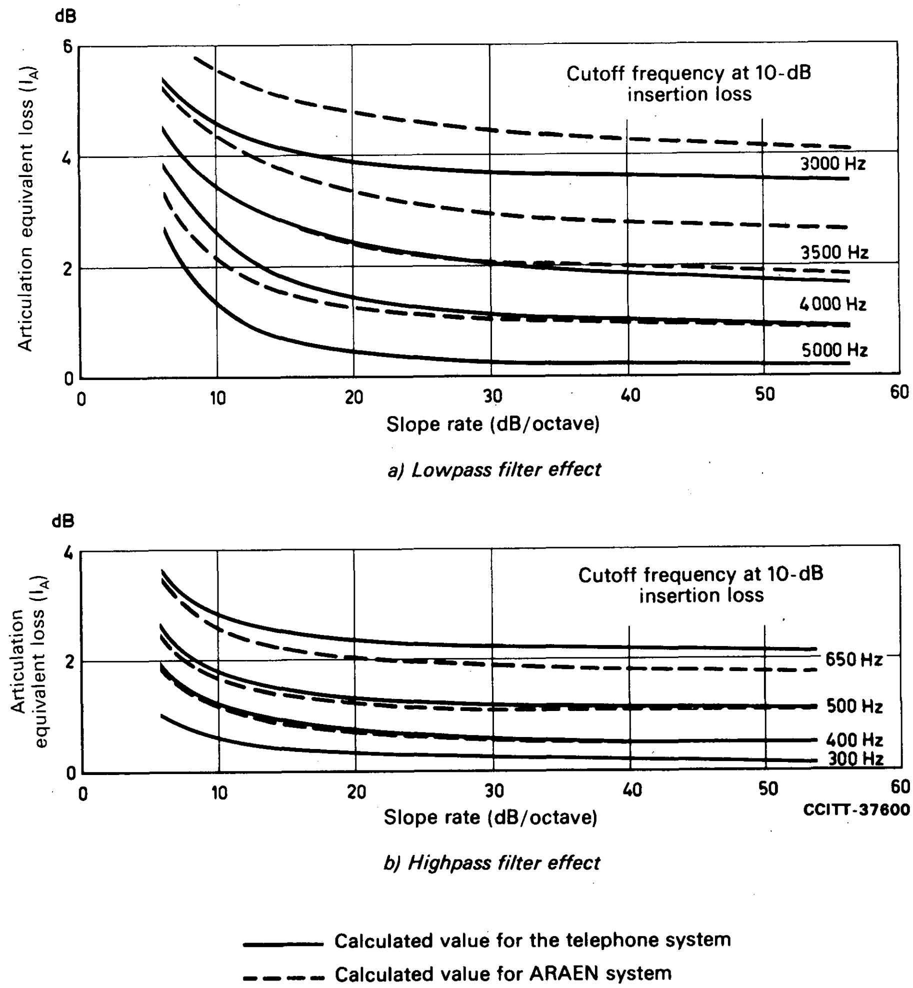


FIGURE B-4/P.11

Lowpass and highpass filter slope effect on articulation

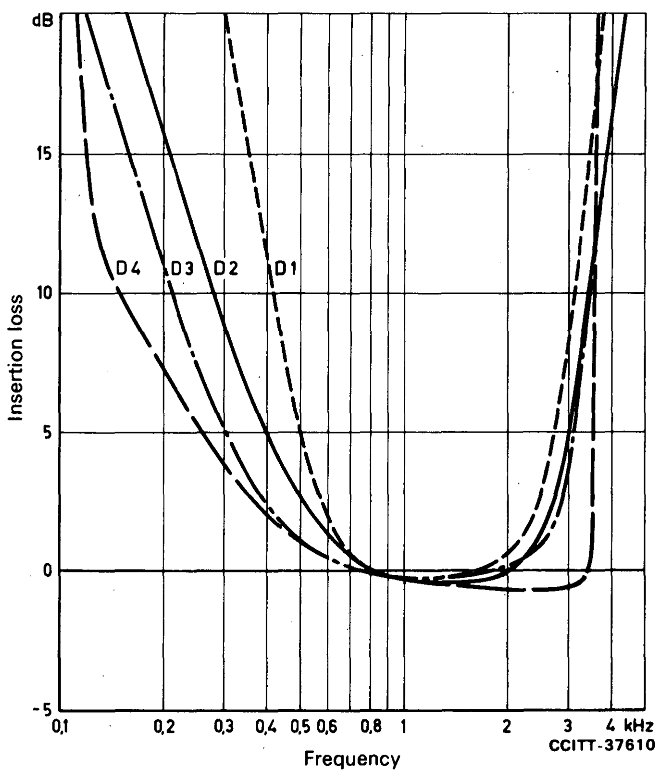
TABLE B-1/P.11

Opinion test conditions

No.	Item	Conditions of conservation opinion test using local telephone circuits	Note
1	Junction loss	3, 13, 23, 29 dB	Measured at 800 Hz
2	Circuit noise level	ICNO [a)] = −48.5 dBmp (14 000 pWp) −54.5 dBmp (3500 pWp) −60.5 dBmp (900 pWp) −78.5 dBmp (14 pWp)	Including exchange noise: −8 dB/octave spectrum characteristics
3	Room noise	50 dBA	
4	Sending and receiving end	Local telephone circuits Telephone: Model 600 Subscriber line: 0,4 mm Ø, 7 dB at 1500 Hz Feeding bridge: XB exchange (220 + 220 Ω) Junction impedance: 600 Ω	SCRE + RCRE = 9,3 dB [b)]
5	Attenuation distortion	D1, D2, D3, D4 (Figure B-5/P.11)	

a) Injected circuit noise referred to the input of a telephone receiving end with 0 dB receive corrected reference equivalent.

b) SCRE = sending corrected reference equivalent, RCRE = receiving corrected reference equivalent.



D 1 12 4-wire circuits chain 95% limit characteristics, based on Figure 1/G.232, Graph No. 2B
D 2 12 4-wire circuits chain characteristics, based on Figure 1/G.132
D 3 Average characteristics of D4 and D2
D 4 SRAEN filter (Recommendations G.111 and P.11)

FIGURE B-5/P.11

Junction attenuation distortion characteristics for test conditions

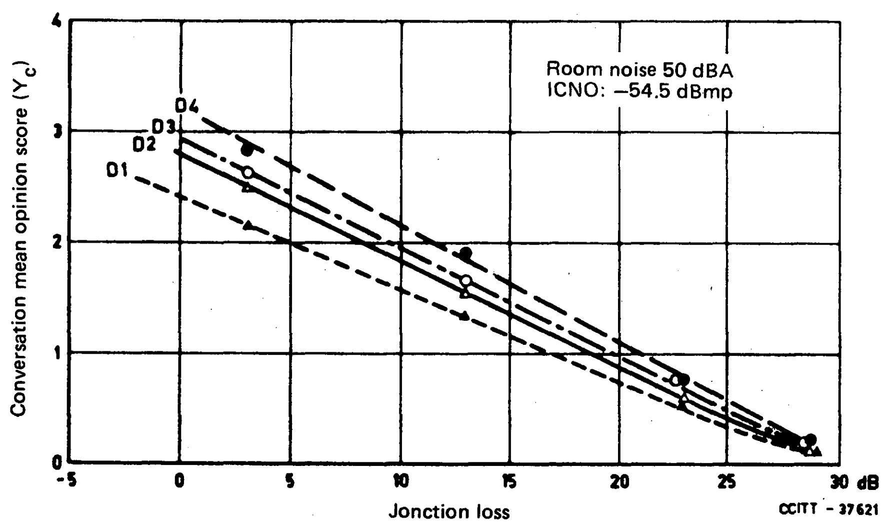


FIGURE B-6/P.11

Attenuation distortion effect on conversation opinion score

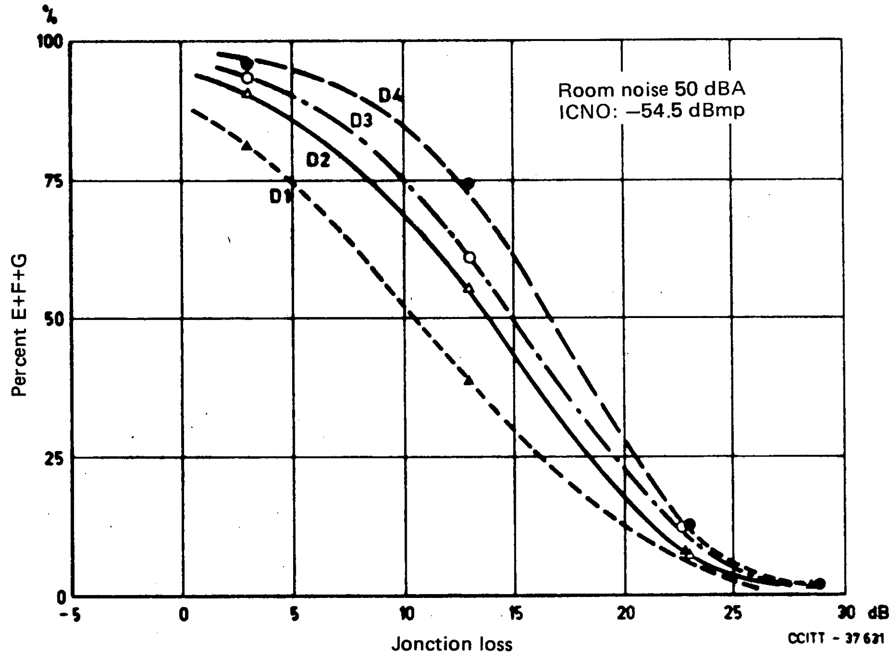


FIGURE B-7/P.11

Attenuation distortion effect on percent F, G and E in conversation test

B.3 *Examples of attenuation distortion characteristics effect*

TABLE B-2/P.11

Example of various method to express attenuation distortion characteristics

Attenuation distortion	Characteristic parameters						Equivalent loss (dB)				
	Cutoff frequency (Hz)		Slope (dB/oct)		Insertion loss (dB)		Aspect 1[a]		Aspect 2[b]	Aspect 3[a]	
	f_{L10}	f_{H10}	f_{L10}	f_{H10}	at 300 Hz	at 3.4 kHz	I_L	I_A	$I_{2.5}$	I_{YC}	$I_{\%FGE}$
D4	150	3500	7.0	300	3.8	0	0	0	0	0	0
D3	210	3400	10.0	31.5	5.2	10	0.8	0.3	-	2.3	1.8
D2	280	3300	10.7	29.1	8.8	10	1.2	0.5	1.8	3.8	2.8
D1	420	3100	22.2	31.1	20.0	15	3.2	2.2	4.2	7.8	6.3

[a] See Reference [31].
[b] Supplemented data to Reference [32].
I_L Attenuation distortion equivalent (calculated value).
I_A Articulation equivalent loss difference at 80% sound articulation (calculated value).
$I_{2.5}$ MOS equivalent loss difference at Y_{LE} = 2.5.
I_{Y_c} MOS equivalent loss difference at Y_c = 2.5.
$I_{\%FGE}$ Accumulated rating equivalent loss difference at 50% F, G and E.

ANNEX C

(to Recommendation P.11)

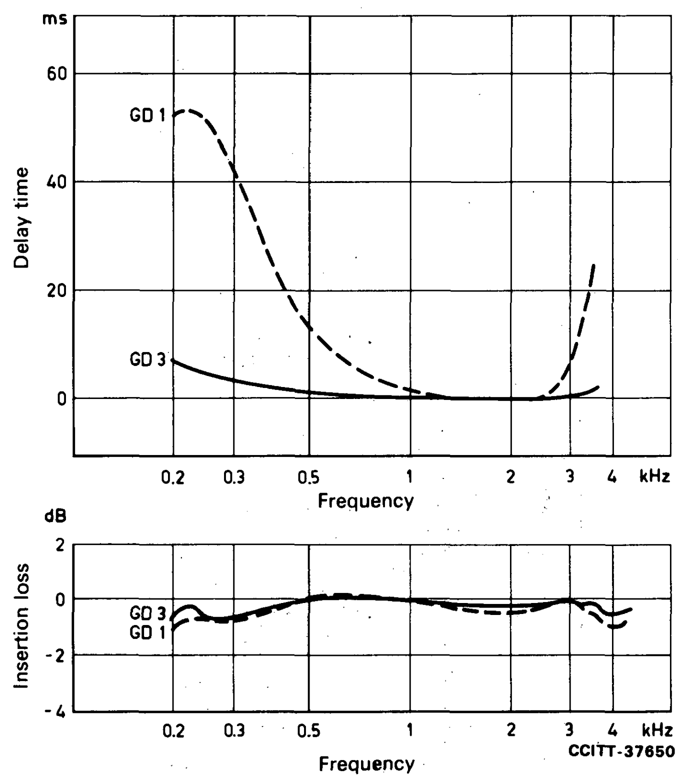
Effects of group-delay distortion on transmission performance

The effect of group-delay distortion at the upper and lower part of a transmitted frequency band is described as “ringing” and “speech blurred”, respectively.

Absence of noise or attenuation distortion has such an influence as to hold the effect conspicuous throughout the possible overall loudness range of a connection.

However, its practical effect in a 4-wire circuit chain does not seem serious, since it is usually accompanied by closely related attenuation distortion.

See Figures C-1/P.11, C-2/P.11 and C-3/P.11.



GD1: Approximated to 12-circuit chain 95% values [19]
GD3: Approximated to typical modern one circuit value

Note - The test conditions are the same as those for the attenuation-distortion opinion test. The circuits modelling junction group-delay distortions used in the test are free from attenuation distortion.

FIGURE C-1/P.11

Junction group-delay distortion of test connection

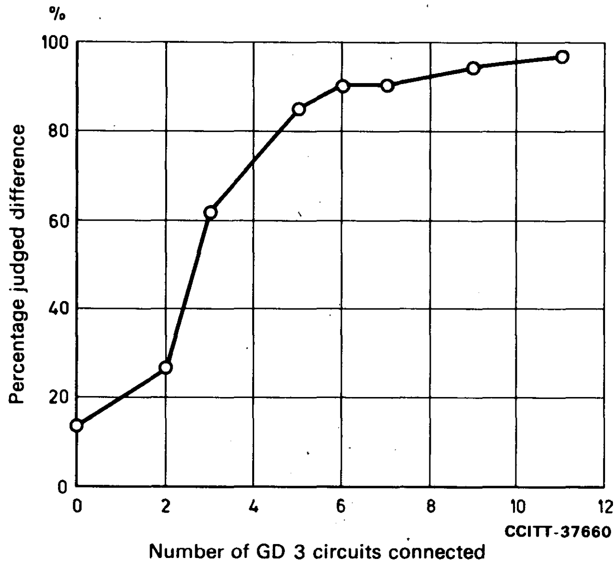


FIGURE C-2/P.11

Group-delay distortion detectability

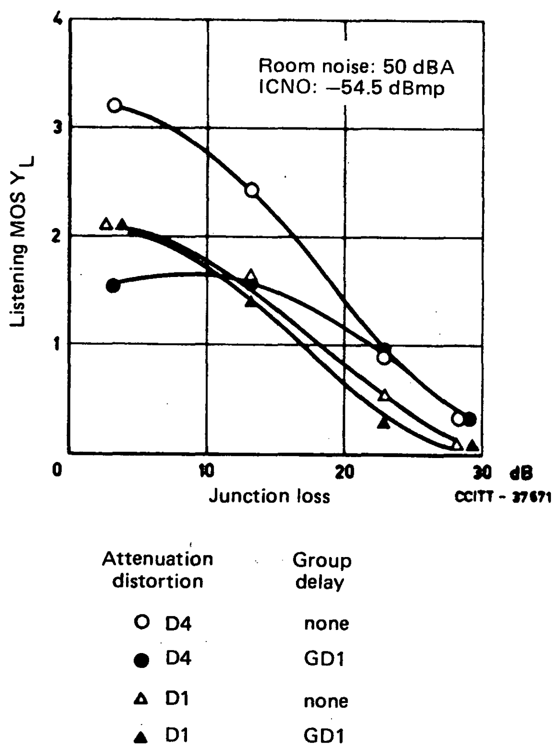


FIGURE C-3/P.11

Effect of group-delay distortion on listening opinion score

ANNEX D

(to Recommendation P.11)

Effects of carbon and linear microphones on transmission performance

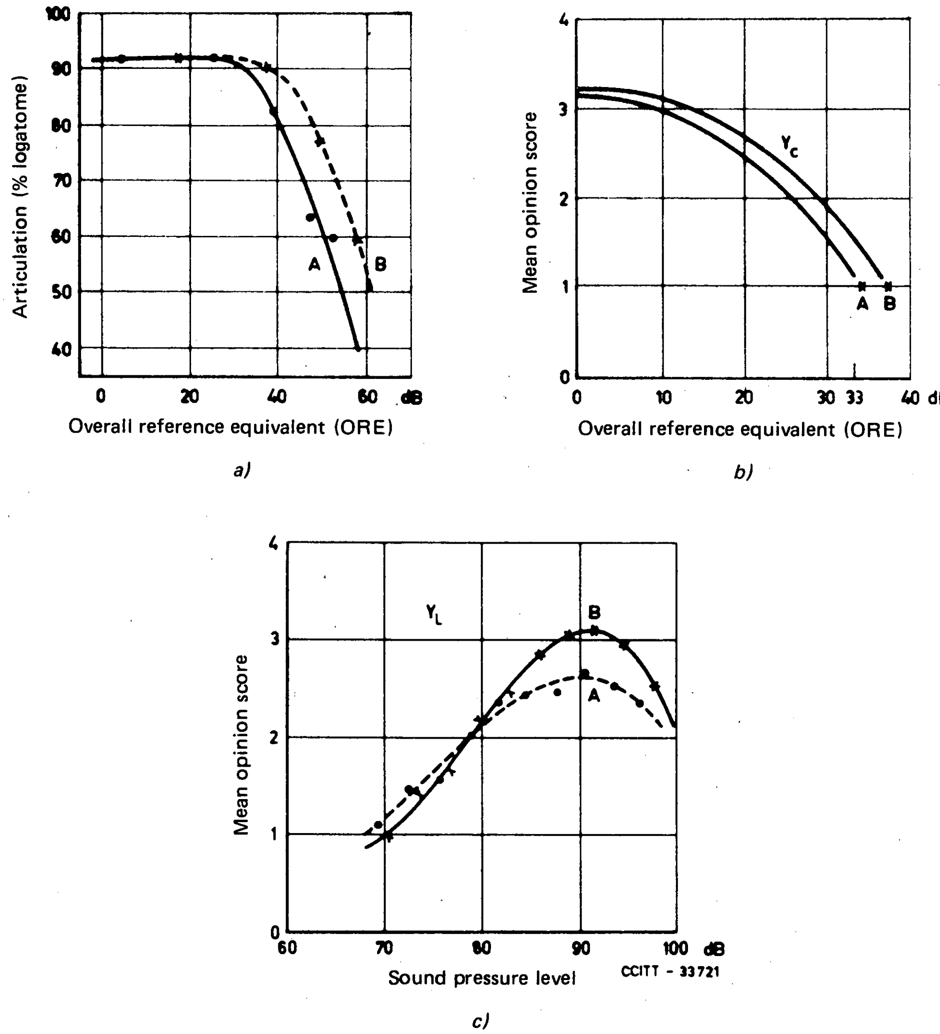
Available information on the difference in performance between carbon and linear (non-carbon) microphones has been collected. The difference depends not only on differences in the content of non-linear distortion due to harmonics and intermodulation products but also on differences in amplitude/frequency distortion ("linear distortion") and amplitude/amplitude distortion (level-dependent sensitivity) between the two types of microphones.

Typical examples of results from comparative tests are given in Figure D-1/P.11. The diagrams show transmission performance measured as articulation or mean opinion score (for conversation or listening only) as functions of reference equivalent or speech level.

No general conclusion can be drawn from such results coming from different sources and dealing with various makes of microphones, because the individual effects of non-linear distortion and of frequency and amplitude-dependent sensitivity cannot be separated. Nevertheless, all three examples indicate some improvement of the transmission performance when a carbon-type microphone is replaced by a linear microphone.

In the particular example *c)* there is a significant improvement at optimum listening level while there is no (or even negative) difference at low listening levels. In that case, with room noise present and insufficient sidetone loss (sidetone reference equivalent 1-4 dB for this test condition) the inferior sensitivity of the specific type of carbon microphone to sound in the acoustic far-field may be an advantage.

For transmission over a larger bandwidth than the conventional telephone band, and in particular for loudspeaker listening, it is likely that there is a more noticeable improvement in sound quality if linear microphones are used instead of carbon microphones.



A Carbon microphone
B Linear microphone

Note – Frequency band: 300-3400 Hz, 50 dB(A) room noise.

FIGURE D-1/P.11

ANNEX E

(to Recommendation P.11)

Quantizing distortion of digital systems

To enable network planning for telephone speech transmission, it is convenient to assign appropriate weights to any nonstandard analogue/digital conversion process, transmultiplex pairs and processes introducing digital loss. An appropriate method is to consider that 1 unit of impairment is assigned to an 8-bit A- or μ-law codec pair to cover quantizing distortion. A planning rule provisionally agreed is to allow 14 units of impairment for an overall international connection, with up to 5 units for each of the national extensions and 4 units for the international chain. Such a rule would allow 14 tandem unintegrated 8-bit processes.

A subjective opinion model (see Supplement No. 3 at the end of this Volume) provides results which indicate that the Q[3)] for an overall connection with 14 unintegrated 8-bit systems in tandem is about 20 dB. The same model shows that one 7-bit system has the same Q as about three 8-bit systems. (This is based on the finding that subjective Q values for digital systems combine on a 15 $\log_{10}$ basis, i.e. 2 digital systems each with a Q = 24.5 dB would yield a Q = 20 dB when connected asynchronously in tandem.) It is recommended that until further information is available, 3 units of impairment (3 qdu) be assigned to a 7-bit system on speech transmission quality.

The provisional values given in Table E-1/P.11 for impairment unit assignment are recommended for planning purposes. These assignments are based on telephone speech considerations.

Note – These preliminary conclusions are based on a limited amount of information and the weights may be revised if new information becomes available.

TABLE E-1/P.11

Impairment unit assignments for telephone speech transmission

Process	Number of impairment units	Notes
One 8-bit A-law or μ-law PCM	1	Note 1
7-bit PCM codec-pair (A-law or μ-law)	3	Note 1
One digital pad realized by manipulating 8-bit PCM code words	1	Note 2
One 32 kbit/s ADPCM-V	3.5	Note 3

Note 1 – For general planning purposes, half the values indicated may be assigned to either of the send or receive parts.

Note 2 – The impairment indicated is about the same for all digital pad values in the range 1-8 dB. One exception is the 6 dB A-law pad which introduces negligible impairment for signals down to about –30 dBm0 and thus attracts 0 units of quantizing distortion.

Note 3 – ADPCM-V = ADPCM with adaptive predictor (Recommendation G.721).

3) Q is the ratio of speech power to speech-correlated noise power determined subjectively by using the MNRU (Modulated Noise Reference Unit) (see the Recommendation P.70). Methods used for subjective assessment of codecs using the MNRU are outlined in Supplement No. 14.

References

[1] CCITT – Question 19/XII, Contribution COM XII-No. 1, Study Period 1985-1988, Geneva, 1985.

[2] CCITT – Question 4/XII, Contribution COM XII-No.1, Study Period 1985-1988, Geneva, 1985.

[3] CCITT – Question 4/XII, Contribution COM XII-No.1, Study Period 1985-1988, Geneva, 1985.

[4] CCITT – Question 9/XII, Contribution COM XII-No.1, Study Period 1985-1988, Geneva, 1985.

[5] CCITT – Question 24/XII, Contribution COM XII-No.1, Study Period 1985-1988, Geneva, 1985.

[6] CCITT – Question 14/XII, Contribution COM XII-No.1, Study Period 1985-1988, Geneva, 1985.

[7] CCITT – Questions 6/XII and 11/XII, Contribution COM XII-No.1, Study Period 1985-1988, Geneva, 1985.

[8] CCITT – Question 5/XII, Contribution COM XII-No.1, Study Period 1985-1988, Geneva, 1985.

[9] CCITT – Question 13/XII, Contribution COM XII-No.1, Study Period 1985-1988, Geneva, 1985.

[10] CCITT – Question 18/XII, Contribution COM XII-No.1, Study Period 1985-1988, Geneva, 1985.

[11] CCITT – Question 7/XII, Contribution COM XII-No.1, Study Period 1985-1988, Geneva, 1985.

[12] CCITT – Contribution COM XII-No. 29 (UKPO), Study Period 1973-1976, Geneva, 1974.

[13] CCITT – Contribution COM XII-No. 60 (Italy), Study Period 1973-1976, Geneva, 1974.

[14] CCITT – Contribution COM XII-No. 144, Study Period 1973-1976, Geneva, 1976.

[15] CCITT – Contribution COM XII-No. 1, Study Period 1977-1980, Geneva, 1976.

[16] CCITT – Contribution COM XII-No. 1, Annexes 2 and 3, Study Period 1981-1984, Geneva, 1981.

[17] CCITT – Contribution COM XII-No. 179 (NTT), Study Period 1977-1980, Geneva, 1979.

[18] CCITT – Contribution COM XII-No. 33 (UKPO), Study Period 1973-1976, Geneva, 1974.

[19] CCITT – Question 1/XVI, Annex 1, Figure 8, Contribution COM XVI-No. 1, Study Period 1977-1980, Geneva, 1976.

Recommendation P.16

SUBJECTIVE EFFECTS OF DIRECT CROSSTALK; THRESHOLDS OF AUDIBILITY AND INTELLIGIBILITY

(Geneva, 1972; amended at Geneva, 1976, 1980; Malaga-Torremolinos, 1984)

1 Factors which affect the crosstalk thresholds

The degree of audibility and intelligibility of a crosstalk signal depends on a large number of factors.

A simple and generally applicable method for estimating the required loss in the crosstalk path as a function of the factors affecting the audibility or the intelligibility of the speech crosstalk signal can be obtained if certain simplifications are made.

The main factors influencing the intelligibility of the vocal crosstalk signal are listed below.

1.1 *Quality of transmission of telephone apparatus* [1]

The sending and receiving corrected reference equivalents are decisive factors. The same is true of the reference equivalent of sidetone when room noise is present. The use of modern telephone apparatus with smooth frequency curves is assumed.

1.2 *Circuit noise*

The circuit noise on the connection of the disturbed call must be taken into account. This is measured by a psophometer equipped with a weighting network for telephone circuits.

1.3 *Room noise*

Room noise affects the ear directly through ear-cap leakage between the ear and the receiver and indirectly by sidetone. Sidetone also depends on operating conditions. Unlike circuit noise, the effect of room noise can be reduced to some extent by the use of the telephone. For this reason and to allow for unfavourable cases, the measurements were made with slight [40 dB (A)] room noise as well as with negligible room noise.

1.4 *Conversation on the disturbed connection*

While there is active speech on the disturbed connection, practical levels of crosstalk are inaudible. However, before the conversation starts or during long pauses in the conversation, it is possible for crosstalk to be heard and perhaps understood. In general, it would be unwise to plan on the basis that the disturbed connection is always active and accordingly the information given in this Recommendation assumes no conversation on the disturbed connection.

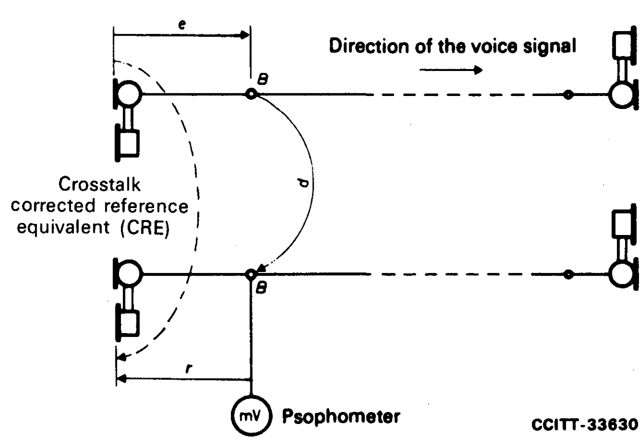
1.5 *Microphone noise*

The noise produced by the carbon microphone of the disturbed telephone may slightly reduce the intelligibility of the vocal crosstalk signal owing to sidetone. Good quality modern microphones have been assumed in this Recommendation.

1.6 *Crosstalk coupling*

The intelligibility of a crosstalk signal also depends on the nature of the crosstalk coupling which is generally a function of frequency. The corrected reference equivalent (CRE) of the crosstalk transmission path can be conventionally divided into the sending CRE of the subscriber's set causing the disturbance, the receiving CRE of the subscriber's set subject to disturbance, and the transmission loss of the crosstalk transmission path. Figure 1/P.16 illustrates this conventional subdivision.

In the absence of further information, the corrected reference equivalent of the crosstalk coupling may be taken to be the attenuation measured or calculated at a frequency of 1100 Hz, as advocated in Recommendation G.134 [2] for telephone exchanges.



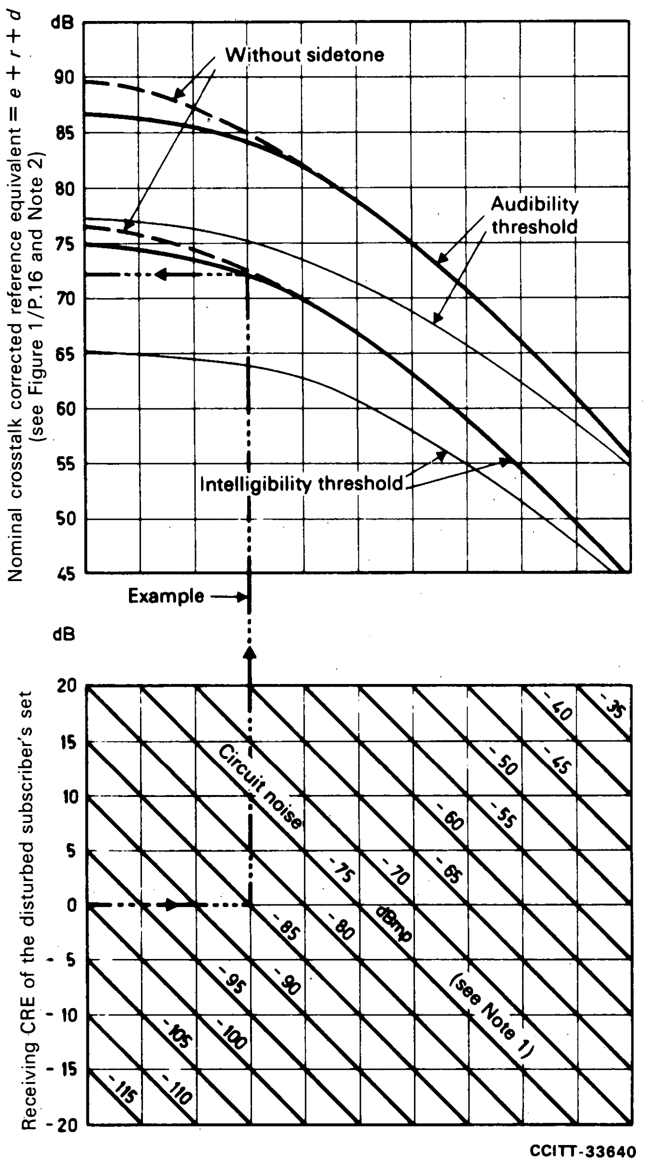
e = sending CRE of the disturbing subscriber's set;
r = receiving CRE of the disturbed subscriber's set;
d = crosstalk path attenuation, so that: crosstalk CRE = $e + d + r$;
B = terminals of subscriber's set.
Disturbing and disturbed subscriber's sets at the same end: near-end crosstalk.
Disturbing and disturbed subscriber's sets at opposite ends: far-end crosstalk.

FIGURE 1/P.16

Conventional subdivision of crosstalk CRE

2 Median listener threshold of the audibility and intelligibility of vocal crosstalk

The curves in Figure 2/P.16 represent the nominal overall reference equivalents of the crosstalk transmission path corresponding to the thresholds of audibility and intelligibility as a function of the receiving CRE. Their parameter is the circuit noise; room noise is negligible or equal to 40 dB (having Hoth spectrum and measured with A weighting). For planning purposes, it is recommended that room noise be regarded as negligible.



The sidetone reference equivalent of the disturbed subscriber's set appropriate to these curves is +13 dB.

Note 1 — The circuit noise is referred to the terminals of the subscriber's set having the CRE indice.

Note 2 — The sending CRE of 0 dB corresponds to a vocal level of −10 VU.

FIGURE 2/P.16

Crosstalk corrected reference equivalent as a function of the receiving corrected reference equivalent and of circuit noise

These curves represent median values for the various conditions such that in each case 50% of subscribers' opinions are respectively above and below the particular curve. The standard deviation for listeners has been observed to lie in the range 4 to 6 dB and a value of 5 dB is recommended for planning purposes.

The results of the original experiments (which form the basis of the curves in Figure 2/P.16) were expressed in terms of speech level (e.g. in volume units) and on that basis showed a satisfactory degree of agreement among themselves.

The thresholds are based on the assumption that a subscriber set with a sending CRE of 0 dB corresponds in practice to a speech level of −10 VU at the subscriber set terminals with a load of 600 ohms.

However, in order for the results to be directly useful for planning purposes for networks designed and characterized on the basis of CREs, it is necessary to introduce a factor (c) which effectively establishes the relationship between speech level and sending CRE.

The correction factor c has been defined in the following manner:

$$c = V_c - V_L \text{ dB}$$

where

V_c = speech level in decibels under normal conversational conditions at a particular point on the disturbing connection;

V_L = speech level in decibels at the same point on the disturbing connection under conditions corresponding to a speech level of −10 VU at the output of a subscriber set with a sending CRE of 0 dB (i.e. it is assumed that the listening tests have been carried out at this speech level);

V_L = $-(10 + a)$, if the actual CRE is a dB. Thus, $c = V_c + 10 + a$.

Thus, the correction factor c is positive for conditions in which the speech level on the disturbing circuit is greater than that corresponding to −10 VU at the output of a subscriber set with 0 dB sending CRE. This correction factor must be added to the value of nominal crosstalk CRE given in Figure 2/P.16. Thus, the threshold more closely corresponding to working conditions becomes: $e + r + d + c = t + c$.

In general, the values of c will be a function of the overall CRE and to some extent of the circuit noise and sidetone reference equivalent on the disturbing circuit. Typical values have been estimated from speech level measurements made by several Administrations and are given in Table 1/P.16 together with standard deviations.

TABLE 1/P.16

Mean values and standard deviations of the factor c for various Administrations

Administration	Nominal overall corrected reference equivalent of the disturbing connection (dB)	Speech level (VU)	Corrected reference equivalent (sending) (dB)	Estimated mean value $\bar{C}$ of the factor c (dB)	Estimated standard deviation σ_c of the factor c (dB)
AT&T	13 24 35	−21 −17 −14	+11 +11 +11	0 +4 +7	4 4 4
Switzerland	40	− 8	+ 2	+4	4
Sweden	6 20	−16 −15	+ 2 + 2	−4 −3	5.3 6.1
United Kingdom Post Office	13 24 35	−17 −16 −14	+ 8 + 8 + 8	+1 +2 +4	4.8 4.8 4.8

As indicated in Table 1/P.16, the value of $\overline{C}$ lies within the range −4 to +7 dB and σ_c = 4 dB. The value of 4 dB for the mean speech level is high. Thus, it is assumed that the disturbing talker is speaking over a connection with a high overall CRE.

3 Crosstalk probability

The curves in Figure 2/P.16 represent median values for the various conditions. The probability of crosstalk in percent can be determined for any crosstalk attenuation with the aid of the method shown in Annex A.

Although the maintenance of telephone secrecy is primordial, the subscriber is more likely to make a severe judgement on crosstalk in a local call taking place in his immediate environment and in which indiscretion due to crosstalk may have unfortunate social consequences.

In practice, simultaneity of speaking on the interfering line and listening on the impaired line (during conversation pauses) is not present in all cases. Information concerning this topic and showing how to calculate the probabilities concerned will be found in [3].

Provisionally, it is recommended that the probabilities of subscribers encountering potentially intelligible crosstalk should not be worse (more) than the following:

Own exchange calls: 1 in 1000

Other calls: 1 in 100.

ANNEX A

(to Recommendation P.16)

This Annex comprises:

1) an example illustrating the method of calculation;
2) a graph showing the probability of crosstalk;
3) an example of a local connection.

A.1 *Example illustrating the method of calculation*

In order to demonstrate the method of using the information given in this Recommendation to calculate the probability of encountering (for example) intelligible crosstalk, a hypothetical reference connection given in Recommendation G.105 [4] is needed. Figure A-1/P.16, based on Figure 3/G.105 [5], illustrates two connections with crosstalk between them introduced by the international circuit.

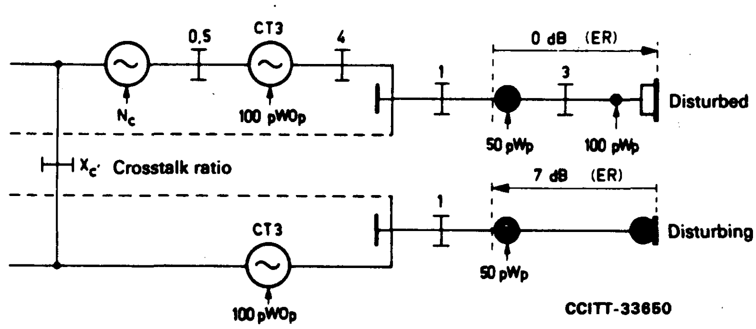


FIGURE A-1/P.16

The crosstalk path of interest may be redrawn as shown in Figure A-2/P.16.

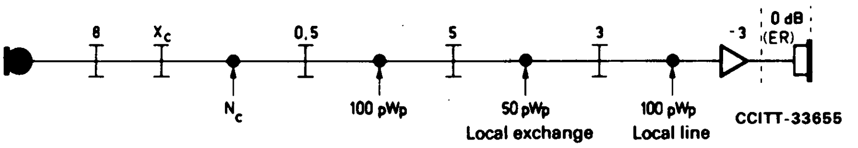


FIGURE A-2/P.16

(The 50 pWp and 100 pWp sources prior to the X_c-pad are ignored, for after transversing the X_c-pad, the resultant noise power contributions will be negligibly small.)

The diagram may be further simplified by referring all the given noise powers to the input of a local system having a corrected reference equivalent of 0 dB and summing (as far as possible) the various losses (see Figure A-3/P.16).

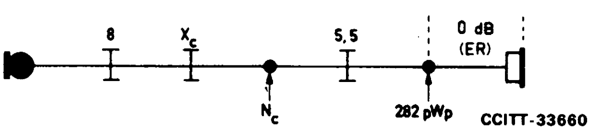


FIGURE A-3/P.16

Considering for the sake of example two specific cases, namely X_c = 58 dB; N_c = 500 pW0p and X_c = 62 dB; N_c = 200 pW0p, the corresponding values of overall X and total N are:

Examples studied	Corresponding values X	N
58 dB; 500 pW0p	71.5	−63.7
62 dB; 200 pW0p	75.5	−64.7

which are associated with the arrangement in Figure A-4/P.16.

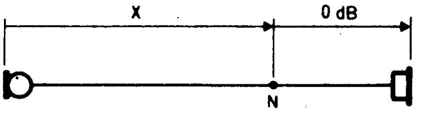
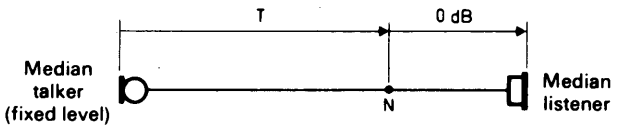


FIGURE A-4/P.16

Table A-1/P.16 records the values of the median threshold of intelligible crosstalk between an active talker and a silent listener. The values have been taken from the curves given in Figure 2/P.16.

TABLE A-1/P.16

Median listener thresholds if intelligible crosstalk as a function of the noise power level at the input to a 0 dB corrected reference equivalent receiving end, for a variety of listening conditions (based on Figure 2/P.16)



CCITT-33680

<table>
<tr><th rowspan="3">N dBmp,
noise power
level at
input to
0 dB RRE</th><th colspan="3">T dB, nominal overall corrected reference equivalent of the crosstalk path</th></tr>
<tr><th colspan="2">Negligible room noise</th><th rowspan="2">+40 dB(A) room noise with sidetone</th></tr>
<tr><th>Without sidetone</th><th>With sidetone</th></tr>
<tr><td>−100</td><td>76.5</td><td>75.0</td><td>65.1</td></tr>
<tr><td>−95</td><td>75.7</td><td>74.5</td><td>64.9</td></tr>
<tr><td>−90</td><td>74.0</td><td>73.0</td><td>64.2</td></tr>
<tr><td>−85</td><td>72.5</td><td>72.0</td><td>64.0</td></tr>
<tr><td>−80</td><td colspan="2">70.0</td><td>62.5</td></tr>
<tr><td>−75</td><td colspan="2">67.0</td><td>60.5</td></tr>
<tr><td>−70</td><td colspan="2">63.0</td><td>58.0</td></tr>
<tr><td>−65</td><td colspan="2">59.0</td><td>55.0</td></tr>
<tr><td>−60</td><td colspan="2">54.5</td><td>51.5</td></tr>
<tr><td>−55</td><td colspan="2">49.5</td><td>47.5</td></tr>
<tr><td>−50</td><td colspan="2">44.0</td><td>43.0</td></tr>
</table>

Note — The sidetone reference equivalent is +13 dB. For intermediate values, use linear interpolation. It has been assumed that the talker delivers −10 VU from a 0 dB corrected reference equivalent end.

In order to take account of the distribution of real talker volumes a correction factor, c, is needed which, being characteristic of national networks, at this present time must be supplied by the user. As indicated in Table 1/P.16 the value of $\overline{C}$ lies in the range −6 to +5 dB and σ_c in the range 4 to 5 dB.

For this example we will use $\overline{C} = 4$ dB, and $\sigma_c = 4$ dB.

We do not have a distribution of crosstalk corrected reference equivalents to take account of at this time; just two specific values.

The standard deviation of the listeners' threshold about the median value is in the range 4 to 6 dB. We will take the value $\sigma_t = 5$ dB in this example.

If t is the threshold value for a particular listener, c the speech volume correction factor for a particular talker and x the actual value of the corrected reference equivalent of the crosstalk path between them, then when x is less than $t + c$, intelligible overhearing occurs. Denoting the difference $x - (t + c)$ by z, intelligible overhearing for this particular pair arises when z is zero or less.

If x, t, and c are each normally distributed (or may fairly be assumed to be so) with mean values $\overline{X}$, $\overline{T}$, and $\overline{C}$ and standard deviations σ_x, σ_t, and σ_c, then z is also normally distributed with mean value, $\overline{Z} = \overline{X} - (\overline{T} + \overline{C})$ and standard deviation $\sigma_z = \sqrt{(\sigma_x^2 + \sigma_t^2 + \sigma_c^2)}$.

The normal deviate at $z = 0$ is given by $\overline{Z}/\sigma_z$ and the probability of $z \leqslant 0$ can be found from tables of the cumulated normal distribution (single upper tail).

The percentage figures of probability can be taken from the usual tables. The graph in Figure A-5/P.16 shows the relation between $\overline{Z}$, σ_z and crosstalk probability.

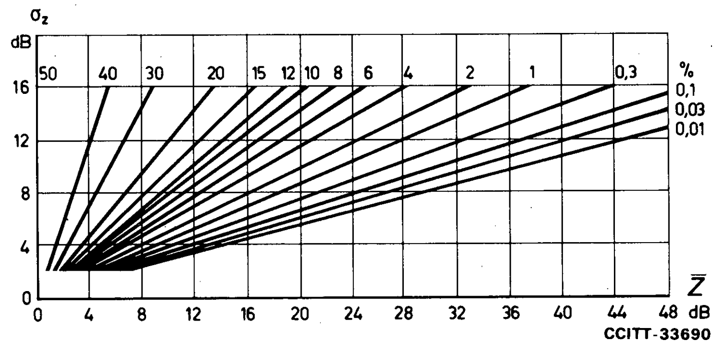


FIGURE A-5/P.16

Relation between σ_z, $\overline{Z}$ and the probability in %

Taking the particular case of 58 dB; 500 pW0p and considering +40 dB (A) room noise with +13 dB sidetone then $N = -63.7$ gives $\overline{T} = 54.1$ (by interpolation of Table A-2/P.16) so that

$$\overline{Z} = \overline{X} - (\overline{T} + \overline{C}) = 71.5 - (54.1 + 4{,}0) = 13.4$$

and

$$\sigma_z = \sqrt{(\sigma_x^2 + \sigma_t^2 + \sigma_c^2)} = \sqrt{(0 + 25 + 16)} = \sqrt{41} = 6.4.$$

Hence $\overline{Z}/\sigma_z = 13.4/6.4 = 2.10$ corresponding to 1.8% risk of intelligible overheading.

Table A-2/P.16 displays the results for each combination used in this example.

TABLE A-2/P.16

Probabilities of intelligible overhearing between active talkers and silent listeners

($\sigma_x = 0$; $\sigma_t = 5$; $\sigma_c = 4$; $\overline{C} = 4$)

Example studied	+40 dB (A) room noise; 13 dB sidetone	Negligible room noise; 13 dB sidetone (or no sidetone)[a)]
58 dB; 500 pW0p 62 dB; 200 pW0p	1.8 % 0.5 %	6.7 % 2.4 %

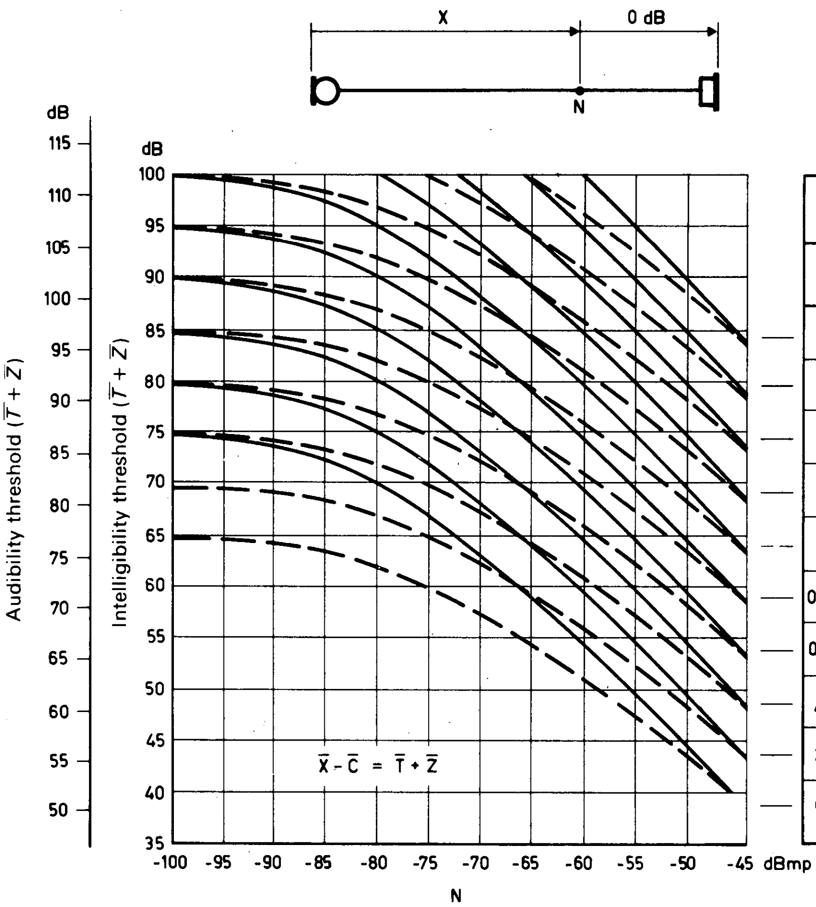
a) With the values used in the examples, the presence or absence of sidetone has no effect.

The values indicated in Table A-2/P.16 for probabilities of intelligible overhearing between active talkers and silent listeners concern crosstalk couplings with negligible standard deviation. Such values can be applied to determine the limits for circuits (channel transfer equipment, for example).

Another method of calculating the probability of intelligible crosstalk using Monte Carlo methods is described in the CCITT manual cited in [6].

A.2 *General diagram concerning the probability of crosstalk*

The graph in Figure A-6/P.16 is based on similar calculations as the example in § A.1, i.e. on the assumption that the noise sources are concentrated in a single point from which the receive corrected reference equivalent is 0 dB. It is further assumed that there is a Gaussian distribution of crosstalk attenuation. The thresholds of audibility- and intelligibility-curves are similar in shape and have been combined into a single pair of curves with different ordinates.



Probability in %				
σ_z				
6	7	8	9	10
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0,02
0	0	0,01	0,04	0,13
0	0,02	0,09	0,28	0,62
0,05	0,22	0,62	1,3	2,3
0,62	1,6	3,0	4,8	6,7
4,8	7,7	10	13	16
20	24	26	29	31
50	50	50	50	50

CCITT-33701

——— Negligible room noise   – – – 40 dB (A) room noise

Note – Each curve corresponds to a specific value of probability (%) indicated on the right-hand side of the diagram as a function of σ_z.

FIGURE A-6/P.16

Probability of (potentially) audible or intelligible crosstalk as a function of $\overline{T} + \overline{Z} = \overline{X} - \overline{C}$, N and σ for room noise of 40 dB (A), and negligible room noise

A.3 *Example of a local connection (based on the hypothetical reference connection given in Recommendation G.105* [4]*)* (see Figure A-7/P.16)

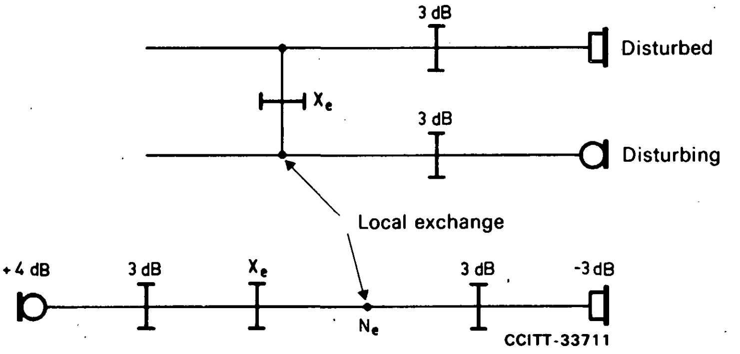


FIGURE A-7/P.16

For this type of connection, the sending corrected reference equivalent of the telephone set is taken to be +4 dB and the receiving corrected reference equivalent −3 dB. The standard deviation of the telephone set at the sending and receiving ends is $\sigma_p = 2$ dB (total). See Table A-3/P.16.

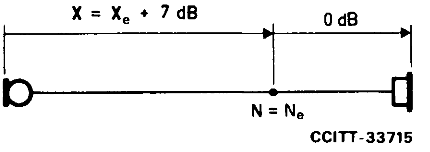


TABLE A-3/P.16

Probability of intelligible overhearing crosstalk (in % for specific values of X_e) between active talkers and silent listeners for a local connection

Room noise		Negligible					
Noise power	N_e pW0p N_e dBm0p	100 −70		1000 −60		10 000 −50	
$\sigma_{\bar{}}$ (total) $\sigma_{\bar{}}$ (diaphonie)		7 2	10 7.4	7 2	10 7.4	7 2	10 7.4
X_e:	60	84	76	50	50	10	18
	62	71	69	39	42	5.8	14
	64	66	62	28	34	3.2	9.7
	66	56	54	20	27	1.6	6.7
	68	44	46	13	21	0.8	4.5
	70	34	38	7.6	16	0.34	2.9
	72	24	31	4.4	11	0.13	1.8
	74	16	24	2.2	7.9	0.06	1.1
	76	10	18	1.1	5.5	0.02	0.6
	78	6.2	13.6	0.5	3.6	0	0.35
	80	3.1	9.5	0.2	2.3	0	0.2
	82	1.6	6.7	0.1	1.4	0	0.1
	84	0.8	4.5	0.03	0.8	0	0.05
	86	0.3	2.8	0.01	0.5	0	0.02
	88	0.1	1.8	0	0.3	0	0.01
	90	0.05	1.1	0	0.1	0	0

References

[1] *Justification for the values of CRE appearing in Recommendations G.111 and G.121,* Volume III, Appendix I to Section 1, § I.7.

[2] CCITT Recommendation *Linear crosstalk,* Vol. III, Rec. G.134.

[3] LAPSA (P. M.): Calculation of multidisturber crosstalk probabilities, *B.S.T.J.,* Vol. 55, No. 7, September 1976.

[4] CCITT Recommendation *Hypothetical reference connection for crosstalk studies,* Vol. III, Rec. G.105.

[5] *Ibid.,* Figure 3/G.105.

[6] CCITT manual *Transmission planning of switched telephone networks,* ITU, Geneva, 1976.

SECTION 2

SUBSCRIBERS' LINES AND SETS

Recommendation P.33

SUBSCRIBER TELEPHONE SETS CONTAINING EITHER LOUDSPEAKING RECEIVERS OR MICROPHONES ASSOCIATED WITH AMPLIFIERS

(Mar del Plata, 1968; amended at Geneva, 1972 and 1980)

The CCITT,

considering

(a) that since an increasing number of loudspeaker sets is being used in the telephone network,

(b) that in view of the complex nature of the effect of factors introduced by these equipments on telephone transmission performance,

(c) that in order to help Administrations to determine the conditions in which the use of such equipment may be authorized in telephone networks,

makes the following provisional recommendation:

(1) In order to avoid overload of carrier systems, the mean long-term power of speech currents should not exceed the mean absolute power level assumed for system design. In Recommendation G.223 [1] the value adopted for this mean power level is −15 dBm0 (mean power = 31.6 microwatts). Loudspeaker telephones having a sending sensitivity that complies with Recommendation P.34 can be assumed to fulfil this Recommendation. Furthermore, in order to avoid excessive crosstalk from high-level speech currents and/or inadequate received volume from low-level speech currents, care should be taken to ensure that the variation of speech currents is not substantially greater than that from modern handset telephones.

(2) Administrations should take the necessary precautions so that the person listening may be able to break the sending circuit if oscillations occur or devise suitable methods so that a device controlled by the voice may prevent oscillations.

Reference

[1] CCITT Recommendation *Assumptions for the calculation of noise on hypothetical reference circuits for telephony*, Vol. III, Rec. G.223.

TRANSMISSION CHARACTERISTICS OF LOUDSPEAKER TELEPHONES

(Geneva, 1980; amended at Malaga-Torremolinos, 1984)

1 Introduction

The sending and receiving sensitivities of handset telephones, normally expressed as Corrected Reference Equivalent (CRE) or Loudness Rating (LR) values, are used in most countries in connection with their national transmission plan for the design of the national network.

However, since it is possible to fulfil Recommendations such as G.121 by distributing CRE or LR, values between the telephone sets and the network in different ways, it is not possible to issue an international Recommendation stating CRE and LR values of telephone sets alone, regardless of whether they are handset or loudspeaker telephones.

On the other hand, it is possible to recommend sensitivity values for Loud Speaker Telephones (LSTs) *relative to* the standard handset telephone used nationally. The aim of such Recommendations should be to obtain equivalent performance with both types of telephones, at least concerning send and receive loudness. That means that the average user's behaviour and preferences while talking and listening have to be taken into account. The relative sensitivities defined in §§ 2 and 3 are derived from performance tests in order to fulfil this requirement.

Other important features contributing to the quality of telephone calls made from loudspeaker telephones cannot presently be dealt with by Recommendations and are studied within the scope of Question 17/XII [1].

2 Sending sensitivity

The sending R25 equivalent (SR25E) or SCRE of a loudspeaker telephone should be about 9 dB worse than the SR25E or SCRE of the corresponding handset telephone (the actual value will depend on the type of handset used). The loudness rating should be about 5 dB worse for the LST than for the handset.

Note – Conversation tests in several countries have shown that comparable speech voltages are obtained on the line when the sending reference equivalent of the LST is 8-11 dB higher than that of the handset telephone used.

The difference of 8-11 dB is composed of several components:

a) the average talking level for LSTs, which is about 3 dB higher than for handsets;

b) the output level from a handset telephone in conversational use, which is about 3-7 dB lower than what is obtained in the speaking position specified for reference equivalent measurements. This difference is reduced to 1-2 dB when the loudness rating guard ring position is used;

c) other minor differences to be considered, such as different frequency response curves.

If the sending sensitivity is controlled by the room noise level, this control should be designed to compensate the expected rise of the talking level with room noise.

It should not be possible for the user to adjust the sending sensitivity.

3 Receiving sensitivity

The receiving sensitivity of a loudspeaker telephone without automatic gain control should be adjustable within a range of 15-30 dB. This range should span the value of the receiving R25 equivalent (RR25E) or receiving loudness rating (RLR) which is equal to that of the corresponding handset telephone, as well as a RR25E or RLR value about 10 dB better. The same values apply to loudness ratings.

Note 1 – In principle the receiving R25 equivalent of the LST should be equal to the RCRE of the corresponding handset telephone in a quiet room. The range of room noise levels met in normal office use necessitates, however, an additional gain of at least 10 dB.

For loudspeaker telephones equipped with an automatic gain control for the receive level (the gain being controlled by the incoming speech voltage), corrected reference equivalents or loudness ratings may not be applicable. In this case the LST should be designed so that the listening level at the maximum line length for which the LST is intended to be used can be preset to a value that may be considered as the best compromise between the levels required for listening in quiet and noisy rooms.

Note 2 – The preferred listening level depends on the room noise level as well as on other external conditions. There is furthermore a great variance between individual listeners.

The average preferred level for listening only appears to be a sound pressure level of about 65 dB for 45 dB(A) room noise, or 70 dB for 55 dB(A) room noise. However, to obtain maximum Mean Opinion Scores in conversation tests, about 5-10 dB higher listening levels may be required.

4 Frequency response curves

4.1 *Sending*

Available information indicates that the optimum slope of the sending response curve when measured with the LST on a table lies between 0 and +3 dB/octave, if the receiving response curve is flat.

Only under highly reverberant conditions a somewhat higher preemphasis may increase the intelligibility. Therefore, if a frequency compensation for the probable cable attenuation is included, the sending curve should not rise with frequency by more than 2-3 dB/octave.

Below 300 Hz there should be a gradual roll-off. The slope may be steeper below 200 Hz.

Note – The interval 200-300 Hz makes a significant contribution to the naturalness of the transmitted speech and should therefore be included in the transmission band of the LST.

Above 4000 Hz, a roll-off by at least −6 dB/octave, preferably −12 dB/octave, is appropriate in order not to cause interference by crosstalk to adjacent channels in certain types of long-distance circuits.

4.2 *Receiving*

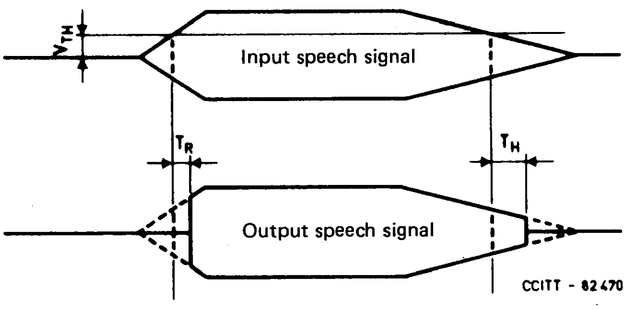
The receiving response curve should be substantially flat in the frequency range of 200-4000 Hz.

The requirement refers to the sound pressure in the undisturbed field at the listener's position with a set-up including the table as described in § 6.

5 Switching characteristics

Most loudspeaker telephones contain a voice-switched circuit which inserts a loss in either the sending or receiving direction in order to permit sufficient gain in the active direction without risk of singing through acoustic feedback. Switching from one direction to the other occurs when a signal above a given threshold is applied from the opposite direction.

Fundamental voice-switching parameters are threshold level, build-up time and hang-over time, as defined in Figure 1/P.34. By a suitable choice of parameter values the degradation of speech quality that is introduced by voice-switching can be made negligible, while an inadequate choice of parameter values, in particular, switching times, may lead to serious clipping effects and loss of initial or final consonants in the transmitted speech.



Threshold level V_{TH}: minimum necessary signal level for removing insertion loss, relative to the active speech level.

Build-up time (or initial delay time) T_R: time from crossing the threshold level up to 50 % of the complete removal of the insertion loss.

Hang-over time (or final delay time) T_H: time from crossing the threshold level up to 50 % of the complete insertion of the switched loss.

Note – The 50 % limit refers to a representation of the signal amplitude on a linear scale. With logarithmic amplitude representation, T_R and T_H are counted to within –6 dB of the complete removal or insertion of the switched loss.

FIGURE 1/P.34

Definition of fundamental voice-switching parameters

In general, it is desirable to keep the threshold value low, build-up time short and hand-over time long. On the other hand, in practical applications extremely short build-up times (about a few milliseconds) may cause the voice-switching circuit to be operated by impulsive noises, while very long hang-over times are likely to impede the natural switch-over in conversation. Furthermore, if the threshold level is more than 25 dB below the active speech, the voice-switching circuit will too easily be operated by ambient noise.

Therefore, the following switching characteristics are recommended for loudspeaker telephones:

a) The build-up time T_R should be less than 15 ms, preferably below 10 ms.

b) the hang-over time T_H should be greater than 100 ms. If the threshold level is in the preferred range, values of T_H between 150 and 250 ms are recommended. Hang-over times greater than 400 ms do not improve the performance noticeably.

c) The threshold level V_{TH} should be at least 20 dB below the active speech level. Levels between -20 and -15 dB may be used if the hang-over time is greater than 300 ms. Levels above -15 dB should not be used.

6 Conditions of measurement

For both subjective and objective measurements, physical test arrangements as described in this section should be used.

6.1 *Test table*

During the measurements the LST is placed on a table defined as follows:

The surface of the table should be hard (e.g. polished marine plywood or suitable hardwood), flat, rigid and horizontal to provide a sound-reflecting surface on which the LST being tested rests. The dimensions of the table should be such that the surface area is about 1 m^2 but not less than 0.96 m^2 and the width not less than 800 mm [2].

Note – This arrangement should be used for all measurements, including for recording of frequency responses, although diffraction effects due to the table are likely to cause severe dips or peaks in the response curve (see § 6.5.2).

6.2 *Test arrangements*

The physical test arrangements of one and two piece LSTs [3] for subjective and objective measurements is shown in Figure 2/P.34.

If the projections of the housing are not rectangular, the point B is positioned at the crossing of the centre line through the housing and the outline of the vertical projection of the housing.

The edge of the front of the box should be perpendicular to the line A-B.

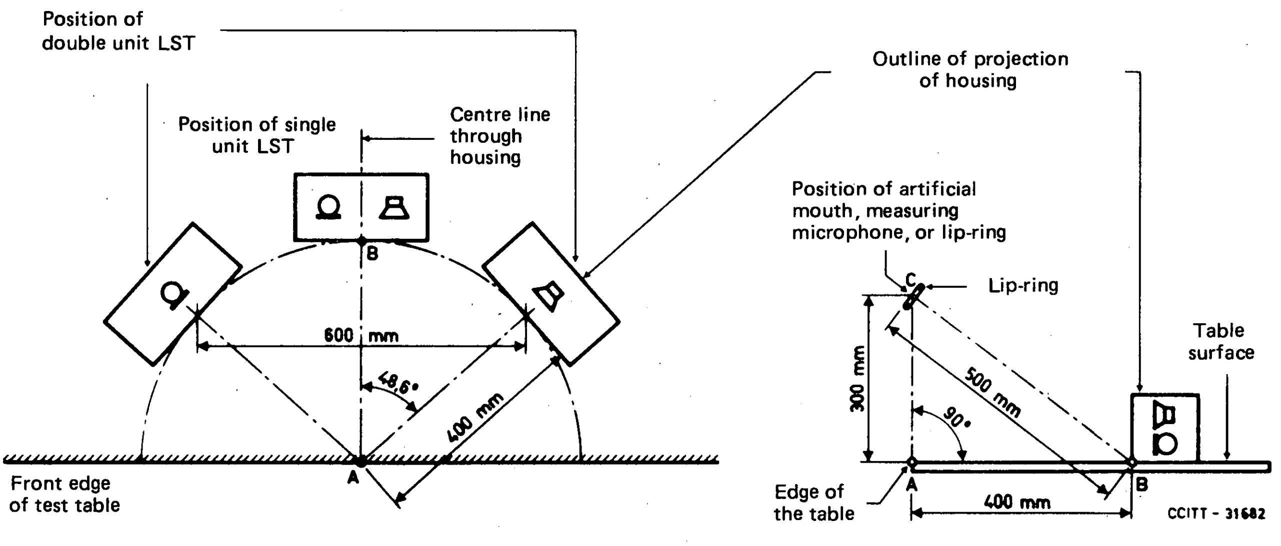


FIGURE 2/P.34

Physical test arrangements for subjective and objective measurements

6.3 *Test environment*

When performing tests the room acoustics must not have a dominating influence. It is recommended for objective measurements that the test environment be practically free-field (anechoic) down to a lowest frequency of 175 Hz, and be such that the test arrangement lies totally within the free-field volume.

Note – Satisfactory free-field conditions may be considered to exist where errors due to the departure from ideal conditions do not exceed ± 1 dB.

The tests should be made in an environment where the ambient noise level is negligible. For objective measurements this is achieved if the Noise Rating (NR) of the Noise Criterion (NC) is lower than 15 [4, 5]. For subjective tests, it may be sufficient to keep the sound level of ambient noise below 35 dB(A).

6.4 *Subjective determinations*

If reference equivalent determinations for LSTs are carried out on an LST they should be in accordance with Recommendation P.72 using the R25 equivalent method. Loudness rating determinations should be determined in accordance with Recommendation P.78.

6.4.1 *Sending*

The talking level for the measurement of SR25E or SLR of an LST should normally be the same as specified for measurements on handset telephones.

It is not necessary for the talker during the test to shift between the reference microphone guard-ring and the guard-ring positioned relative to the LST if the obstacle effect of the reference microphone can be assumed to be negligible.

Normally the specified talking level and the use of a conventional test phrase or sentence should be sufficient to ensure that a voice switched LST will be in the sending condition during the determination of loudness related ratings (SRE or SLR). If this is not the case the talking level may be increased by up to 5 dB.

Note – To preserve the same listening level at balance the NOSFER attenuator may be increased from 25 dB to 30 dB.

If the sending sensitivity is controlled by the room noise level the subjective measurement should be done in a quiet environment [< 35 dB(A)]. Further information about the LST performance may then be estimated by repeating the sending measurements with increasing levels of room noise, up to a maximum of 60 dB(A).

6.4.2 *Receiving*

The talking level at the reference microphone for the measurement of RR25E or RLR should normally be the same as specified for the measurement of handset telephones. This should normally ensure that when loudness balance is achieved between NOSFER and the test system path, a signal of sufficient magnitude is present at the LST to switch it into the receive condition. Problems can sometimes occur when approaching the balance condition from the condition of high attenuation in the balance attenuators, when the low level input signal may fail to switch the LST into the receiving condition. If this does occur the talking level may be increased by up to 5 dB.

Note – This will increase the listening level at balance, but in this case it is not possible to correct this by changing the NOSFER attenuator.

Obtaining the loudness balance for the receiving condition may be facilitated by use of a loudspeaking intermediate reference system. The specification of such a system is, however, outside the scope of this Recommendation.

6.5 *Objective determinations*

6.5.1 *Sending*

The sending sensitivity of a local telephone system (LTS) with the handset telephone replaced by the LST is consistent with Recommendation P.64 and is expressed as follows:

$$S_{mJ} = 20 \log_{10} \frac{V_s}{p_s} \quad \text{dB rel 1 V/Pa}$$

where V_s is the voltage across a 600 ohm termination and p_s is the sound pressure at 500 mm from the lip position of the artificial mouth.

The sound pressure level (p_s) is -20 dBPa and is measured in the absence of the LST and table.

Note 1 – The sound pressure level is consistent with a) in the Note of § 2.

Note 2 – The method of free-field measurement means that the diffraction effects of the table and LST will be the same for the artificial mouth in the measuring position as for a real mouth in the same position (if both sound sources have the same radiation pattern).

Note 3 – Some LSTs use "noise guard" circuits and therefore the source signal has to be modified. A suitable method is to pulse the source signal at an adequate rate, e.g. 250 ms "ON" and 150 ms "OFF". Sending sensitivities determined for LSTs in this way are not suitable for use in calculating send loudness ratings (SLR). For this purpose the reference sound pressure should have a level at the MRP, which is (on average, over the frequency range of interest) 24.2 dB higher. Thus for SLR calculations the send sensitivity required is given by the expression:

$$S_{mJ} = 20 \log_{10} \frac{V_s}{1.62} \qquad \text{dB rel 1 V/Pa at MRP}$$

the sound pressure being 1.62 Pa (corresponding to 4.2 dBPa) in this case.

6.5.2 *Receiving*

The receiving sensitivity of a local telephone system (LTS) with the handset telephone replaced by the LST is consistent with Recommendation P.64 and is expressed as follows:

$$S_{Je} = 20 \log_{10} \frac{p_R}{\frac{1}{2} E_J} \qquad \text{dB rel Pa/V}$$

where p_R is the sound pressure measured at point C in Figure 2/P.34 and ½ E_J is half the e.m.f. in the 600 ohms source.

Depending on the mounting of the loudspeaker in the housing and its position relative to the table surface, interference patterns may occur at the position of the measuring microphone. In order to investigate the receiving performance of a loudspeaker telephone for a real listening situation, it may be advantageous to measure the response at a number of microphone positions in a range ± 10 cm from the recommended position. The average of these measurements at any frequency can be used to represent the receiving response instead of a single measurement at one point.

An alternative method is to use a sufficiently wideband pink noise signal and analyze in ⅓ octave bands.

Note – Receiving sensitivities determined for LSTs in this way may not be used for the calculation of receive loudness ratings (RLR). The subject of RLR calculations from objective measurements or LSTs is complex and involves the difference between loudness judgements made monaurally and binaurally. This item is still being studied under Question 17/XII.

References

[1] CCITT – Question 17/XII. Contribution COM XII-No. 1. Study period 1981-1984. Geneva, 1981.

[2] European Committee for Standardization (CEN) *Office chair/desk working position – dimensions and design requirements*, CEN: prEN91/August 1981.

[3] CCITT *A method for measuring the sensitivity of a loudspeaking telephone set*, Annex 2 to Question 17/XII, White Book, Vol. V, ITU, Geneva, 1969.

[4] ISO *Assessment of noise with respect to community response*, ISO Recommendation 1996 – 1971.

[5] BERANEK (L. L): Noise and Vibration Control, *McGraw Hill*, pp. 564-566, New York, 1971.

Recommendation P.35

HANDSET TELEPHONES

(Malaga-Torremolinos, 1984)

The transmission plan for international telephone connections is given in Recommendation G.101.

Recommendations on the transmission quality, i.e. connected reference equivalents and loudness ratings for international and national telephone connections are given in Recommendations G.111 and G.121 respectively.

These Recommendations permit Administrations to split the requirements between analogue telephone sets and the other parts of the network as long as the above-mentioned Recommendations are fulfilled.

Therefore no precise Recommendations can be given for analogue telephone sets. However, some design considerations can be given. These considerations can be found in Supplement No. 10.

Recommendation P.36

EFFICIENCY OF DEVICES FOR PREVENTING THE OCCURRENCE OF EXCESSIVE ACOUSTIC PRESSURE BY TELEPHONE RECEIVERS

(Malaga-Torremolinos, 1984)

The purpose of this Recommendation is to define methods for checking the efficiency of the devices for protection against acoustic shocks, the use of which is recommended in CCITT Recommendation K.7.

The CCITT is also pursuing studies of methods more generally applicable to the qualification of devices for preventing any dangerous or annoying acoustic pressure from telephone receivers.

1 Efficiency of protection against acoustic shocks

1.1 *Electroacoustic measurements using impulses*

Preliminary Note – On the basis of the findings of scientific studies, several authors or organizations have proposed ear-damage risk criteria based on variations in acoustic pressure, under impulse conditions for which, incidentally, there is no single definition. However, these criteria cannot be directly transposed to the test conditions and measurements described below. Nor could the results be cross-checked without introducing certain hypotheses that are not specified in this Recommendation, the purpose of which is merely to describe a method simple both in its application and in the analysis of the results obtained. The criteria recommended are based on experience gained in several countries in the telephone receiver quality necessary to ensure the safety of users and operators.

In order to check whether a telephone set affords satisfactory protection against the risk of acoustic shocks, it is recommended that its characteristics be examined as follows:

a) the entire telephone set, including the protective device, is placed in normal operating conditions as regards current supply, and its position for the exchange of a call (e.g. with the handset raised);

b) the earpiece of the handset earphone is applied in the normal way to an artificial ear conforming to Recommendation P.51 (which corresponds to IEC Publication 318);

c) the artificial ear is electrically connected to a precision sound level meter conforming to IEC Publication 651, correctly calibrated and having the necessary circuits for measuring peak acoustic pressure values. This equipment must be of class 2 for prototype testing, and maybe of class 3 for checking mass-produced sets;

d) electrical impulses are applied to the telephone set by a suitable assembly which enables them to be superimposed on the d.c. supply without the latter short-circuiting them. These impulses are produced by a generator which conforms with Figure 1/K.17, and whose components are those described for symmetric-pair repeater tests (R_3 = 25 ohms, C_2 = 0.2 μF, see Table 1/K.17). The test voltage is between 0 and 1.5 kV;

e) for the application of such an electric impulse, the peak acoustic pressure level observed (maximum instantaneous value) should be below 140 dB [1]. In the long term, Administrations are recommended to limit this value to 135 dB [1] for sets in common use.

Note – Administrations may deam it appropriate to use different limits for specific cases, for instance for the headsets used by operators.

1.2 *Electroacoustic measurements using steady-state sinewave signals*

It is recommended to check whether the strong-signal attenuation obtained by protective devices does not cause deterioration of the normal speech signals, e.g. by non-linear distortion. This may be done by conducting a series of measurements using steady-state sine-wave signals at a frequency of 800 Hz and relating to the following magnitudes:

N (dB): electric voltage level at the terminals of the set

$$N = 20 \log_{10} \frac{V_{\text{r.m.s}}}{0.775}$$

where $V_{\text{r.m.s.}}$ represents the r.m.s. value of the voltage across the terminals.

$P(N)$: acoustic pressure produced by the telephone receiver under given conditions. (This may be the pressure measured on an artificial ear in accordance with Recommendation P.51), corresponding to the application of voltage level N across the terminals of the set.

A(dB): Attenuation of electroacoustic efficiency in relation to its reference value determined for $N = -20$ dB.

$A(N)$ is determined by the relation:

$$A(N) \text{ dB} = 20 \log_{10} \frac{P(-20)}{P(N)} + N + 20$$

$[A(N) = 0$ when $N = -20$ dB].

The values obtained for $A(N)$ must match those in Table 1/P.36 which have been obtained from measurements carried out on several types of set fitted with various protective devices.

Note 1 – It may be useful to make a few additional measurements to ensure that, at frequencies between 200 Hz and 4000 Hz, the values observed for $A(N)$ are of the same order.

Note 2 – Some sets of recent design have special features, such as electroacoustic sensitivity which depends on the conditions of d.c. current supply or on the level of the speech signals received, quite apart from the effect of the protective devices. In that case, Administrations intending to use such sets may have to adapt the above conditions, taking care nevertheless to comply with their principles.

1) Relative to μPa.

TABLE 1/P.36

N (dB)	$A(N)$ (dB)
−20	0
−10	< 0.5
0	⩽ 2

2 Efficiency of protection against other acoustical annoyances (hazards or discomfort)

The CCITT is at present studying methods for checking this type of efficiency.

Recommendation P.37

MAGNETIC FIELD STRENGTH AROUND THE EARCAP OF TELEPHONE HANDSETS WHICH PROVIDE FOR COUPLING TO HEARING AIDS

(Malaga-Torremolinos, 1984)

1 Introduction

Magnetic induction systems incorporated in telephone handsets generate an alternating magnetic field with spatial characteristics which make the field detectable by hearing aids equipped with induction pick-up coils.

Reception of an audio-frequency signal via an induction pick-up coil can often allow an acceptable signal-to-noise ratio to be achieved in cases where the acoustical reception would otherwise be degraded by reverberation and background noise.

The magnetic field strength which enables induction pick-up coils in hearing aids to function effectively must be high enough to produce an acceptable signal-to-noise ratio but not so high as to cause overloading of the hearing aid.

The value of magnetic field strength recommended in this standard has been chosen so that these requirements are met as far as possible.

2 Scope

This Recommendation applies to telephone handsets which provide a magnetic field for coupling to hearing aids. It specifies the level linearity and frequency dependence of the magnetic field strength produced by the handset and characteristics for a calibrated probe coil.

3 Explanation of terms

3.1 *Level of magnetic field strength*

The maximum value of the magnetic field strength is measured in accordance with § 4.2.

3.2 *Plane of measurement*

A plane parallel to the earcap plane at a distance of 10 mm.

4 Magnetic field strength measurements and recommended values

4.1 *Calibration of acoustic receive level*

Using the measurement configuration shown in Figure 3/P.64, the drive level of the oscillator shall be adjusted to produce a sound pressure level of 80 dB at 1000 Hz. This drive level shall be used for measuring the level and frequency characteristics of the magnetic field strength.

4.2 *Magnetic field strength level*

Place (per § 5) the centre of the calibrated probe coil in the plane of measurement and circuit orientate it for maximum coupling. Determine the magnetic field strength at 1000 Hz using the drive level as per § 4.1.

Recommended range of values for the magnetic field strength is:

-17 to -30 dB relative to 1 A/m.

Note – Hearing aids with magnetic pick-up coils primarily intended for coupling to magnetic loops in auditoria in accordance with IEC Publication 118-4 are likely to have a sensitivity that corresponds to a field strength in the upper end of the range recommended for coupling to telephones.

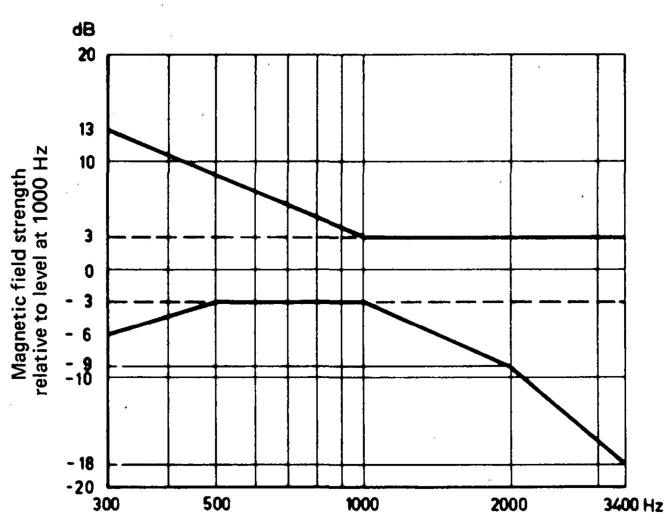
4.3 *Linearity of the magnetic field strength*

With the probe coil positioned as in § 4.2, increase the 1000 Hz drive level specified in § 4.1 by 20 dB and measure the resulting magnetic field strength.

The field strength should increase by 20 dB ± 1 dB, or if the telephone set has a higher linearity the linearity of the magnetic field shall be equally as good.

4.4 *Measurement of frequency characteristics*

With the probe coil positioned as in § 4.2 and the drive level as in § 4.1 vary the frequency from 300 Hz to 3400 Hz and measure the resulting field strength. The magnetic field strength frequency characteristics shall fit within the template shown in Figure 1/P.37.



Note – Preferred frequency characteristics are within the dotted lines (± 3 dB). Range of acceptable characteristics is within the solid lines.

FIGURE 1/P.37

Magnetic field strength frequency characteristics

5 Probe coil

5.1 *Dimensions*

For measuring the magnetic field strength a calibrated probe coil having the following dimensions is recommended:

Core: length (12.5 ± 1 mm)
cross section (1 mm ± 0.5 mm) × (2 mm ± 0.5 mm)

Winding: length (10 mm ± 1 mm)
cross section (2 mm ± 0.5 mm) × 3 mm ± 0.5 mm)

The winding shall be shorter than the core.

Note 1 – The magnetic field may be inhomogeneous within distances comparable to the length of the probe coil. The introduction of a magnetic core material may also redirect the magnetic field contours, therefore the magnetic material of the core may be of importance.

Note 2 – The probe coil may be combined with frequency correcting elements to obtain a flat frequency response in the range of 300 Hz to 3400 Hz.

5.2 *Calibration of the probe coil*

The sensitivity as a function of frequency of the probe coil shall be measured with an accuracy of ± 0.5 dB.

A method of producing a homogeneous magnetic field of known intensity is given in IEC Publication 118-1. The harmonic distortion of the magnetic field used for the calibration shall be less than 1%.

5.3 *Distortion*

The distortion of the probe coil shall be less than 2%, when measuring field strength up to +2 dB relative to 1 A/m.

Bibliography

Methods of measurement of electro-acoustical characteristics of hearing aids. Part 4: Magnetic field strength in audio-frequency induction loops for hearing aid purposes, IEC Publication 118-4, 1981.

AHLBORG (H.): Speech levels in the Swedish telephone network. *TELE Engl. Ed. No. 1*, 1978.

DAHLGAARD (T.) and NIELSEN (A. K.): A statistical analysis of speech signals in a local exchange, and a calculation of the line impedance from the natural speech signals. *Teleteknik No. 2*, 1974.

GLEISS (N.): Preferred listening levels in telephony. *TELE Engl. Ed. No. 2*, 1974.

SECTION 3

TRANSMISSION STANDARDS

Recommendation P.42

SYSTEMS FOR THE DETERMINATION OF REFERENCE EQUIVALENTS

(amended at Mar del Plata, 1968; Geneva, 1980
and Malaga-Torremolinos, 1984)

Three systems are in existence for the determination of reference equivalents. These three systems should comply with the conditions shown below and are designated as follows:

1) the new fundamental system for the determination of reference equivalents (NOSFER) (French abbreviation of: Nouveau Système Fondamental pour la détermination des Equivalents de Réference);
2) primary systems for the determination of reference equivalents: (replicas or NOSFER);
3) working standard systems.

The NOSFER is the system used in the CCITT Laboratory. Formerly, reference equivalents were determined by comparison with the European master reference system for telephone transmission (SFERT), defined in [1].

Values of reference equivalents determined by comparison, directly or indirectly, with the SFERT remain valid.

In the past, other telephone transmission reference systems were also used; these are described in [1].

1 Fundamental system for the determination of reference equivalents (NOSFER)

Two systems are currently in use, both designed to fulfill the same performance specification as detailed in the following sections. In essence this is to achieve sending and receiving sensitivity/frequency characteristics for NOSFER, substantially similar to those of SFERT and thus ensure the validity of reference equivalents previously obtained by comparisons with SFERT.

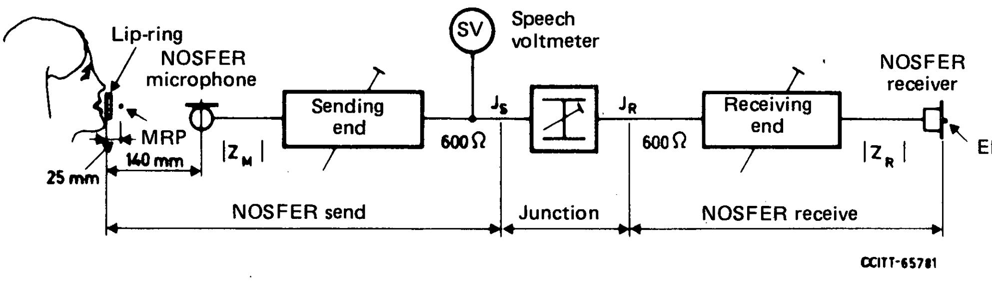
These systems are:

a) the original NOSFER which was based on the components of the high quality transmission apparatus (ARAEN) which is fully described in Recommendation P.41 (*Yellow Book*) [2] and [3], [4] and [5];
b) the NOSFER-1984, in which components readily obtainable as current production items are associated with integrated circuits and active equalizers. A particular advantage of this system is the use of a condenser microphone which greatly simplifies calibration processes.

In the following sections the Recommendation gives only the essential performance which is applicable to both. Details of composition and setting-up procedures, including the necessary modifications to ARAEN to achieve the original NOSFER, are given in Supplement No. 12.

1.1 *Basic representation of the complete NOSFER*

The basic representation of the complete NOSFER is shown in Figure 1/P.42. A general description and other essentials including sensitivity/frequency characteristics are given under separate headings for sending end, junction and receiving end.



Note 1 – MRP and ERP: mouth and ear reference points (see Recommendation P.64).

Note 2 – Junction input and output impedance is 600 ohm non-reactive.

FIGURE 1/P.42

Basic representation of complete NOSFER

1.2 *NOSFER sending end*

Basically this is specified as a highly stable and linear assembly of electro-acoustical equipment which gives a sensitivity/frequency characteristic as shown in Table 1/P.42. In addition it requires facilities for coarse and fine gain control, as given in Supplement No. 12.

TABLE 1/P.42

Sensitivity of the NOSFER sending system

Frequency (Hz)	Sensitivity [dB (V/Pa)]	Frequency (Hz)	Sensitivity [dB (V/Pa)]
100	6.80 (–6.6)	1000	12.00 (–1.4)
125	7.90 (–5.5)	1250	12.30 (–1.1)
160	8.40 (–5.0)	1600	12.10 (–1.3)
200	8.60 (–4.8)	2000	12.50 (–0.9)
250	8.20 (–5.2)	2500	12.50 (–0.9)
315	7.80 (–5.6)	3150	13.90 (+0.5)
400	8.00 (–5.4)	4000	13.30 (–0.1)
500	8.80 (–4.6)	5000	9.20 (–4.2)
630	10.00 (–3.4)	6300	1.60 (–11.8)
800	11.00 (–2.4)	8000	–2.50 (–15.9)

Note 1 – Sensitivity dB (V/Pa) with voltage measured across 600 ohm junction input (J_S) and free-field sound pressure at the microphone.

Note 2 – Sensitivities shown in parenthesis are those related to sound pressure at the mouth reference point (see Annex A to Recommendation P.64), required for the calculation of reference equivalents, using an appropriate algorithm.

Note 3 – Tolerances are ± 1 dB (0.1-4.0 kHz), 1 ± 1.5 dB (4.0-6.3 kHz) and ± 2 dB (6.3-8.0 kHz).

Note 4 – The values presented in the table were taken from the original United Kingdom data and interpolated to ISO mid-frequencies of 1/3 octave bands; therefore they may differ slightly from values shown in previous versions of Recommendation P.42 (e.g. the *Yellow Book* version [3]).

When in use for reference equivalent determinations the talking distance for NOSFER is 140 mm, measured between the plane tangent to the lip-ring nearest to the talker's lips and the centre of the acoustic grid of the microphone. The acoustic grid lies in a horizontal plane which is tangential to the central axis of the lip-ring.

A volume meter, of the type used previously in ARAEN, is connected across the output of the NOSFER sending system and the input of the junction (J_S). This volume meter (speech voltmeter) was manufactured originally for ARAEN in the United Kingdom as type SV 3 with principal characteristics as given in Table 1/P.52. The "standard" speech volume used with NOSFER is defined in Recommendation P.72, § 4.2.1.

Figure 2/P.42 shows the sensitivity/frequency characteristic of the NOSFER sending end. Allowable tolerances on the shape of the curve are as shown in Note 4 to Table 1/P.42.

The composition of the sending ends of both systems is shown for the original NOSFER based on ARAEN in Figure 4/P.42 and for the NOSFER-1984 version in Figure 5/P.42.

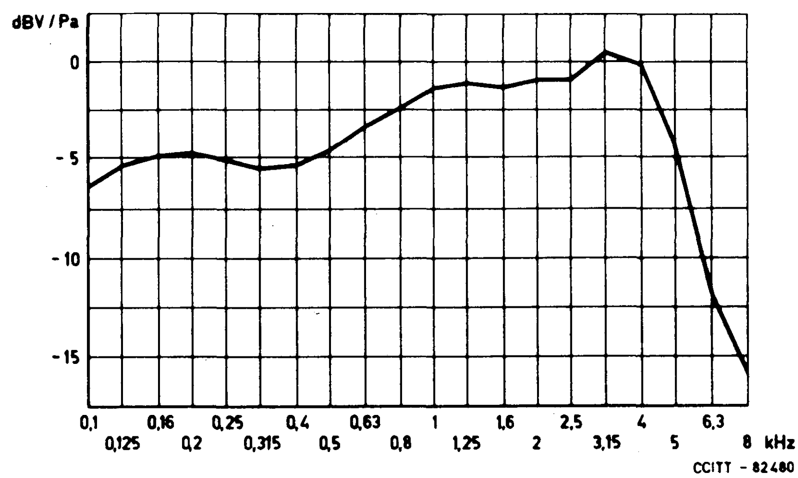


FIGURE 2/P.42

Nominal sensitivity/frequency of the NOSFER sending system

1.3 *NOSFER junction*

The variable loss element connecting the sending and receiving ends of NOSFER is essentially composed of two variable attenuators, both calibrated in 1 dB steps, with input/output non-reactive impedances of 600 ohms. Two attenuators are used in tandem as there is a need for a "hidden loss" element remote from the operator adjusted attenuator. Occasionally there is also a need for an amplification in the adjustable loss path and for this purpose a 600 ohm input/output amplifier with gain being set to 0, 10 or 20 dB is included in series with the 600 ohm balancing attenuator.

1.4 *NOSFER receiving end*

As with the sending end, this is again specified in basic terms as a highly stable and linear set of electro-acoustical equipment which gives a sensitivity/frequency characteristic as shown in Table 2/P.42. Facilities for coarse and fine gain control are again detailed in the practical description of the NOSFER receiving end, which is given in Supplement No. 12.

The sensitivity/frequency characteristic of the NOSFER receiving end is shown in Figure 3/P.42. Allowable tolerances on the shape of the curve are as shown in Note 2 to Table 2/P.42.

TABLE 2/P.42

Sensitivity of the NOSFER receiving system

Frequency (Hz)	Sensitivity [dB (Pa/V)]	Frequency (Hz)	Sensitivity [dB (Pa/V)]
100	1.30	1000	6.60
125	2.10	1250	6.40
160	2.40	1600	6.70
200	3.30	2000	7.50
250	4.00	2500	7.70
315	4.80	3150	6.10
400	5.70	4000	4.10
500	7.30	5000	7.40
630	6.70	6300	12.5
800	6.50	8000	2.0

Note 1 – Sensitivity dB (Pa/V) – sound pressure measured at ear reference point (see Recommendation P.64) with constant voltage level applied across NOSFER junction output (J_R). For the sound pressure measurement the NOSFER receiver is positioned on the IEC 318 artificial ear with flat plate (see Annex C).

Note 2 – Tolerances are ± 1 dB (0.1-4.0 kHz), 1 ± 1.5 dB (4.0-6.3 kHz) and ± 2 dB (6.3-8.0 kHz).

Note 3 – The values presented in Table 2/P.42 were taken from the original United Kingdom data and were interpolated to ISO mid-frequencies of 1/3 octave bands; therefore they may differ slightly from values shown in previous versions of Recommendation P.42 (e.g. *Yellow Book* version [3]).

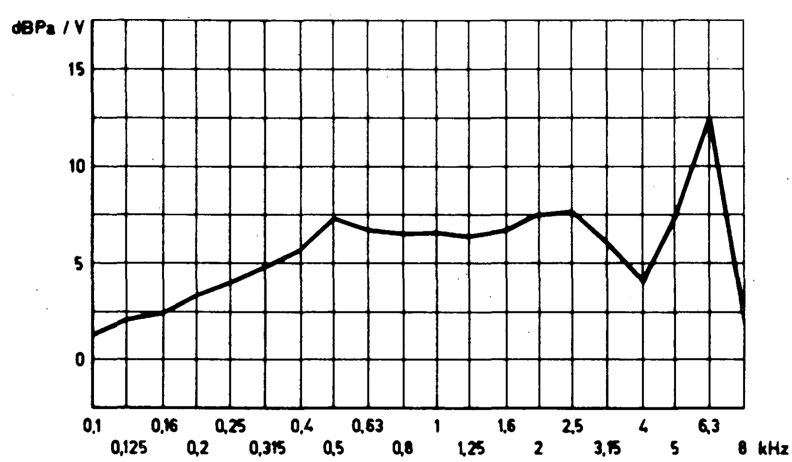


FIGURE 3/P.42

Nominal sensitivity/frequency of the NOSFER receiving system

The composition of the receiving ends of both systems is shown for the original NOSFER in Figure 4/P.42 and for the NOSFER-1984 version in Figure 5/P.42.

2 Physical representation of NOSFER

Figures 4/P.42 and 5/P.42 show the composition of the two NOSFER systems. Component details are given in Supplement No. 12 together with setting-up procedures for voice-ear tests.

Note – In addition to the change from moving coil to capacitor microphone previously mentioned, it has become necessary in recent years to specify a replacement receiver. CCITT agreement has been reached to use a receiver type DR 701 (IWATSU, Japan) as an alternative to the type 4026A (STC, UK) on either system. Previously it had been shown that both receivers when fitted with the specified rubber cushion had the same impedance and substantially similar performance. It has been found useful to present in more detail the internal construction of the NOSFER receiver and this is included for the receiver type 4026A as Annex D to this Recommendation. The internal construction of the receiver type DR 701 is different from the type 4026A.

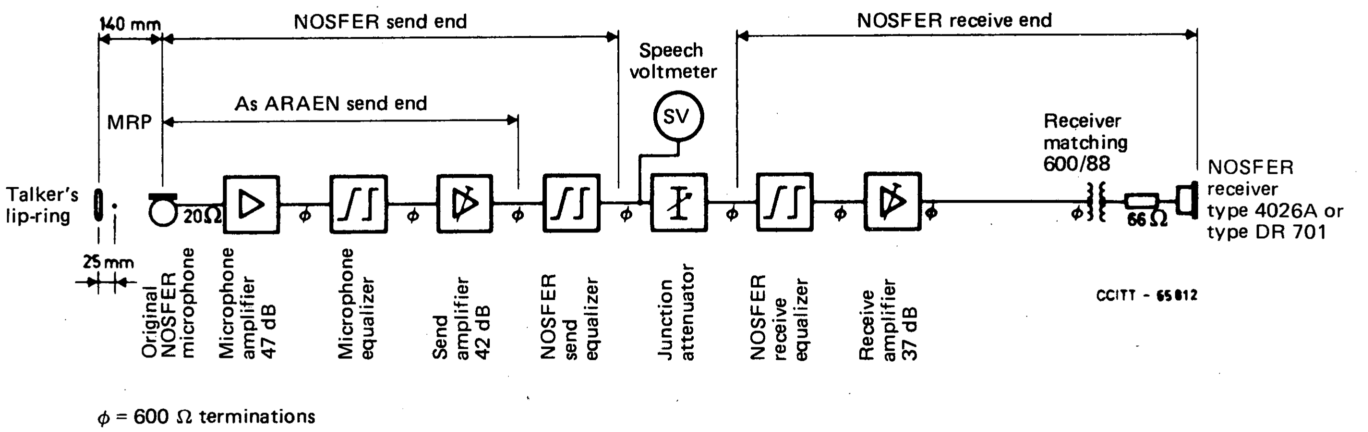
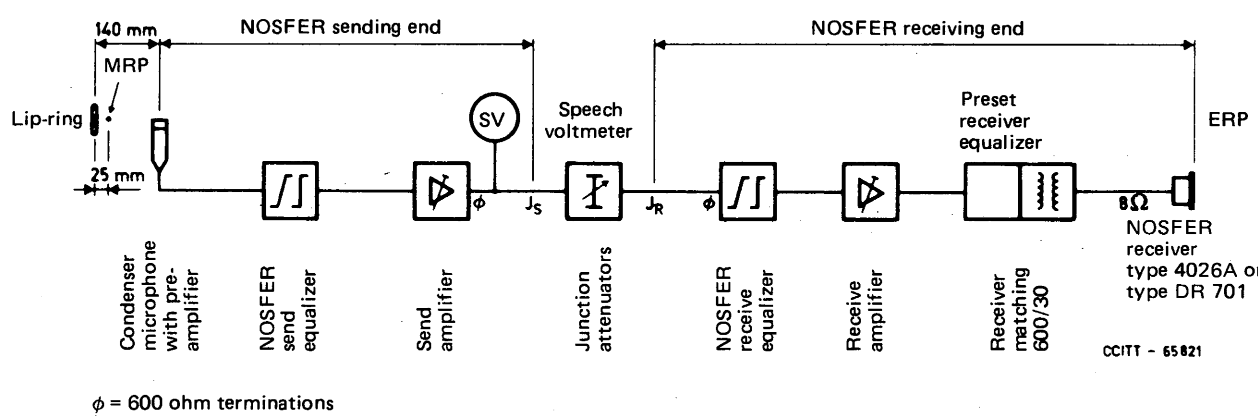


FIGURE 4/P.42

Physical representation of the original NOSFER (ARAEN based model)



Note – MRP and ERP: mouth and ear reference points (Recommendation P.64).

FIGURE 5/P.42

Physical representation of NOSFER-1984

3 Normal adjustment of NOSFER

The normal adjustment of NOSFER is, in principle, a process covering:

a) periodic test of the long-term stability and

b) simple every-day check that the system is aligned for use and that the specified performance is being maintained.

The procedure previously applied to the original NOSFER is described in the *Yellow Book* version of this Recommendation [3].

For both NOSFERs, although periodic checks of the microphone and receiver are still necessary, it is nowadays considered preferable to calibrate the sending end and receiving end as complete units in the manner detailed in Annex B.

Other relevant information on the normal adjustment of NOSFER as applied to the setting-up procedure is given in Supplement No. 12.

Because the calibration of NOSFER earphones forms an important part of the above processes, the method of measurement using the IEC-318 artificial ear is given as Annex C.

4 Primary systems for the determination of reference equivalents

"Primary system for the determination of reference equivalents" is the name given to:

a) a system consisting of a replica of NOSFER,

b) a system conforming to the description given in [6].

It is assumed that such a system:

1) has its sending and receiving sensitivities identical with NOSFER, conforming to Table 1/P.42 and Table 2/P.42 values;

2) is defined by a detailed description, including the relevant method of objective calibration of the physical parameters of the system;

3) has been compared directly with NOSFER.

5 Working standard systems for the determination of reference equivalents

It is admitted for the purposes of the application of Recommendations that the reference equivalent of a commercial system may be determined by taking the sum of the relative equivalent of this commercial system obtained by comparison with a working standard and the reference equivalent of the working standard system (see Recommendation P.72).

By way of information, brief descriptions of the working standard systems are reproduced in Annex A.

Before being officially put in service, any working standard should be compared with the NOSFER or with a primary system for determining reference equivalents.

This comparison is intended to define the transmission qualities of sending and receiving systems of the working standard as compared with the corresponding systems of NOSFER or a primary system for the determination of reference equivalents. It indicates in decibels the amount by which the respective sending or receiving system of the working standard is worse or better relative to the sending or receiving system of NOSFER (or a primary system for the determination of reference equivalents).

The measuring method used in the CCITT Laboratory is the so-called "two-operator, hidden-loss method" (see Recommendation P.72).

The circuit diagrams showing the method of evaluation of sending and receiving systems of the working standard in comparison with NOSFER are shown in Figures 1/P.72 and 2/P.72 respectively.

5.1 *Periodic calibration of working standard systems*

Working standard systems must be periodically compared against NOSFER or a primary system for the determination of reference equivalents. Recommendations for forwarding such apparatus are contained in Recommendation P.43.

ANNEX A

(to Recommendation P.42)

Brief description of the working standard systems

A.1 *Rules concerning the composition of working standards with subscribers' equipment (SETAB)*

Working standards with subscribers' equipment consist of a sending system, an attenuator and a receiving system. The sending and receiving systems consist respectively of subscribers' sets of a commercial type with a linear (non-carbon) microphone, associated with a subscriber's line and a feeding bridge.

The attenuator connected between the sending and receiving systems should have a minimum loss of 40 dB, variable in 1 dB steps, and input/output impedances of 600 ohms.

The system should be completed with a volume meter to enable the vocal power used during telephonometric tests to be maintained.

A.2 *Description of a working standard having an electro-dynamic microphone and receiver (SETED)*

The SETED working standard was originally designed for use as a reference system for ratings based on loudness and for articulation rating (AEN). A detailed description of this working standard is given in [3]. Basically it consists of a calibrated speech path, having a frequency characteristic similar to that which would be given by a 1-metre air path including the obstacle effect of the human head, but band-limited from 300 to 3400 Hz. It uses a moving coil microphone of a special type designed for close talking which is substantially protected against the effects of breath moisture. To standardize the lip position a guard-ring is fitted to the microphone at a distance of 25 mm.

SETED is provided with means for absolute calibration of its microphones and receivers and for this purpose makes use of a calibrated quartz crystal microphone. In recent years it has become possible to confirm calibrations with modern capacitor microphones.

The complete description of SETED can be found in the *Yellow Book* version of Recommendation P.42 (Annex B) [3].

ANNEX B

(to Recommendation P.42)

Calibration of NOSFER using Recommendation P.79

Using an adaptation of the method of calculating loudness ratings (Recommendation P.79) it has been shown that calculated send and receive loudness ratings of NOSFER relative to those of a system having the specified sending and receiving sensitivities of Tables 1/P.42 and 2/P.42 yield reliable criteria for any departure from specification which can be used to make fine adjustment to the sensitivities to give the desired NOSFER performance.

For this purpose it was necessary to calculate a set of "*W*" weights incorporating the specified NOSFER sensitivities rather than those used in Recommendation P.79 (which are based on the IRS 2 of Recommendation P.48). The "*W*" weights are shown in Table B-1/P.42; for W_S and W_O they apply to the condition where the reference sound pressure is setup at the 140 mm point, i.e. the position used for the NOSFER sending system calibration.

The calculation method, as used here, also assumes that the measured frequency characteristics of §§ 1.2 and 1.4, are within the required tolerances and that no allowance will be made for L_E as defined in Recommendation P.79.

TABLE B-1/P.42

"W" weights for NOSFER calibration

Frequency (Hz)	Send W_S	Receive W_R	Overall W_O	Frequency (Hz)	Send W_S	Receive W_R	Overall W_O
100	123.0	117.5	124.3	1000	76.0	70.6	82.6
125	95.8	90.0	97.9	1250	84.0	78.1	90.4
160	80.1	74.1	82.5	1600	80.7	75.3	87.4
200	80.0	74.7	83.3	2000	91.8	86.8	99.3
250	74.0	69.8	78.0	2500	88.2	83.4	95.9
315	81.0	78.0	85.8	3150	106.6	98.8	112.7
400	67.2	64.9	72.9	4000	90.1	80.9	94.2
500	76.9	75.4	84.2	5000	102.3	100.5	109.7
630	73.9	70.6	80.6	6300	98.3	109.2	110.8
800	74.2	69.7	80.7	8000	93.8	98.3	95.8

Note 1 – Values based on reference sound pressure at the 140 mm point.

Note 2 – "*W*" weights only valid for very similar frequency characteristics; they should not be used to calculate reference or R25 equivalents of telephone sets.

To apply the method, the sending and receiving sensitivities of NOSFER are first measured in the manner of Tables 1/P.42 and 2/P.42. Calculations are then made to determine SLR, RLR and OLR (the latter, solely when required, as a check) using the equations:

$$\mathrm{SLR} = -57.1 \log_{10} \sum_{I=1}^{20} 10^{\left(\frac{1}{57.1}\right)\left(S_{RMJ}(I) - W_S(I)\right)} \tag{B-1}$$

$$\mathrm{RLR} = -57.1 \log_{10} \sum_{I=1}^{20} 10^{\left(\frac{1}{57.1}\right)\left(S_{RJe}(I) - W_R(I)\right)} \tag{B-2}$$

$$\mathrm{OLR} = -57.1 \log_{10} \sum_{I=1}^{20} 10^{\left(\frac{1}{57,1}\right)\left(S_{RMJ}(I) + S_{RJe}(I) - W_o(I)\right)} \tag{B-3}$$

Any deviations from zero in SLR, RLR or OLR are then taken up by appropriate adjustments to the gains of send or/and receive amplifiers to restore all loudness ratings to within 0 ± 0.2 dB.

ANNEX C

(to Recommendation P.42)

Calibration of NOSFER earphones with the IEC-318 model artificial ear

The measurement results contained in [7] which coincide, moreover, with those given in [8] demonstrate that the sensitivity of the NOSFER and former ARAEN receiver type 4026A with rubber cushion earpad may be measured on either the IEC-318 model [9] or the formerly used ARAEN artificial ear, to yield substantially the same result, providing the receiver is seated in each case on a flat plate flush with the rim of the artificial ear (see Figure C-1/P.42).

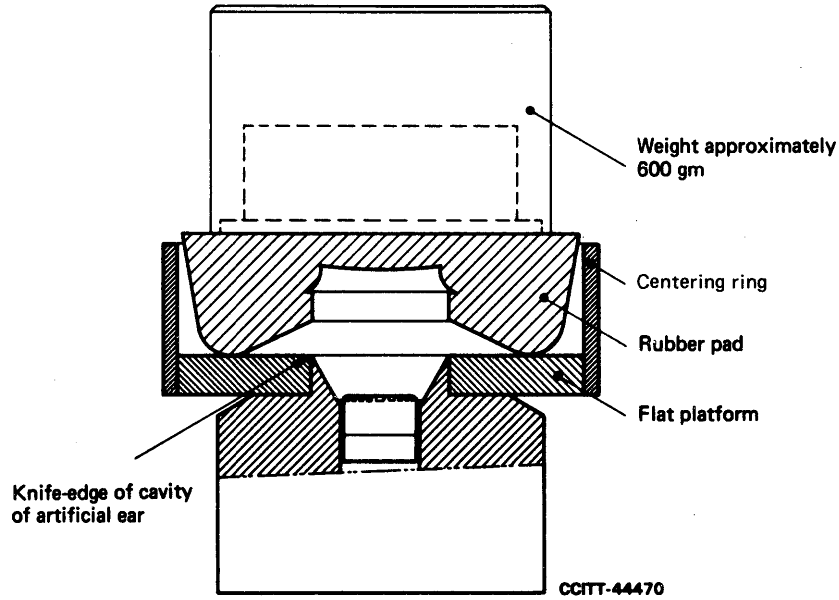


FIGURE C-1/P.42

NOSFER receiver applied to a platform mounted flush with the top of the input cavity of the IEC-318 model artificial ear

Furthermore, it is known that the setting-up of the ARAEN receiving end was originally based on the agreement between real ear calibrations of the receiver type 4026A with rubber cushion and the calibration method used above with the ARAEN artificial ear. Also, that recent artificial ear measurements of the receiver IWATSU type DR 701 with the same rubber cushion have been shown to be in sufficient agreement with those of the type 4026A for approval to be given for its use as a replacement NOSFER receiver.

The CCITT therefore recommends that, for future objective measurements of the NOSFER receive end, for exploring correlation between subjectively measured ratings and calculated ratings based on objective measurements, the IEC-318 model artificial ear be used with a flat plate as described. The receiver should be seated on the artificial ear with a mass of 600 grams (excluding the receiver mass). The pressure so exerted corresponds broadly to that of the former headband used with ARAEN or that of a hand-held receiver for careful listening as with NOSFER.

Note – The Recommendation contained in this annex is made solely in connection with the calibration of NOSFER type receivers with rubber earpad. It is assumed that receivers of conventionally shaped telephone handsets will be seated directly on the artificial ear as prescribed in [9] and Recommendation P.51.

ANNEX D

(to Recommendation P.42)

NOSFER Receiver – type 4026A

D.1 *Introduction*

Although the CCITT Laboratory has in recent years been fortunate in securing a replacement receiver (IWATSU type DR 701) for the STC type 4026A, interest has been expressed in having some detail of the construction of the initial receiver type 4026A. This has been prompted to some extent by the exceptionally high degree of stability over a period of more than 30 years of use with both ARAEN and NOSFER. Furthermore, a great deal of information exists on all aspects of both objective and subjective performance (see [4], [5], [7] and [8]).

D.2 *Construction*

A cross-sectional sketch of the receiver, approximately 1.5 times normal size, which covers the essential features, is shown in Figure D-1/P.42.

The domed diaphragm is made of a light aluminium alloy with tangential corrugations to obtain good amplitude linearity. The diaphragm carries an equally light speech coil which is formed of edge wound aluminium tape, thus dispensing with the need for an armature and giving an excellent space factor in relation to the magnetic gap.

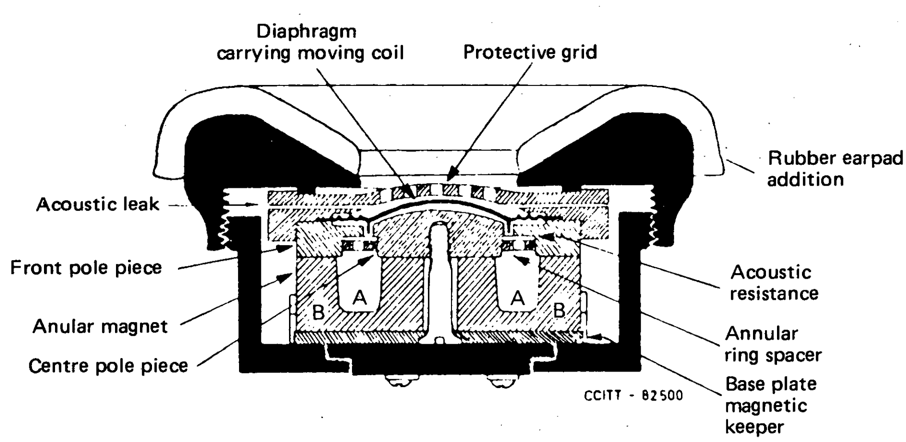
The pole pieces are precision-made so that in conjunction with a perforated annular ring spacer they provide a magnetic gap which cannot become eccentric. The ring perforations, which couple the rear of the diaphragm to a major cavity within and around the largely cylindrical ring magnet, are covered with silk to provide a specific acoustical resistance between these coupled regions. It is these features, together with acoustical equalization provided by lamb's-wool partially filling the rear magnet cavity, which enable a wide and level frequency response to be obtained.

To safeguard the receiver against acoustic surges and changes of barometric pressure, two "acoustic leak" drillings are provided in the diaphragm cover, thus coupling the front and rear diaphragm cavities. These drillings are partially filled with lamb's-wool and are again used to exercise some control over the frequency response.

The protective diaphragm cover, with its array of holes providing the acoustic radiation, is covered with a cloth mesh to prevent ingress of dust. It is securely clamped by the screw thread of the earpiece against a circular washer in order to prevent unwanted sound transmission to the rear cavity.

The rubber earpad, which is fixed to the hard surfaced earpiece, ensures a reasonably repeatable seal to the human ear and gives some attenuation to unwanted signals at frequencies at least above 300 Hz. More particularly, it ensures reasonable agreement between real ear and artificial ear calibrations which is required from loudness tests and calculations.

Sensitivity/frequency characteristics and other performance details can be obtained from [3], [4] and [5], and other information from [10]. (In this respect, the section on moving coil microphones should be studied, as the construction of the NOSFER microphone type 4021E included is almost identical to that of the receiver type 4026A apart from diaphragm thickness which is increased for the receiver.)



Note — The outer walls BB of the annular air space AA are cut away to some extent along the diameter at 90° with respect to that shown in the figure, to allow acoustic coupling of the main annular cavity with the air space surrounding the magnet. It also allows space for screws to fasten the front circular pole piece to the base plate in a similar manner to the fastening of the central pole.

FIGURE D-1/P.42

References

[1] CCIF *Measuring methods and apparatus*, Green Book, Vol. IV, Part 3, pp. 27-43, ITU, Geneva, 1956.

[2] CCITT Recommendation *Description of the ARAEN*, Yellow Book, Vol. V, Rec. P.41, ITU, Geneva, 1981.

[3] CCITT Recommendation, *Systems for the determination of reference equivalent*, Yellow Book, Vol. V, Recommendation P.42, ITU, Geneva, 1981.

[4] *Research Report No. 13200*, United Kingdom Post Office, April 1950.

[5] *Absolute calibration of the ARAEN at the CCITT Laboratory*, White Book, Vol. V, Supplement No. 9, ITU, Geneva, 1969.

[6] CCIF *Measuring methods and apparatus*, Green Book, Vol. IV, Part 3, § 3.1.1, II, pp. 27-34, ITU, Geneva, 1956.

[7] CCITT – Contribution COM XII-125, Study Period 1968-1972, Geneva, 1971.

[8] CCITT – Contribution COM XII-12, Technical Report No. 355 of the CCITT Laboratory, Study Period 1966-1968, Geneva, 1967.

[9] *An artificial ear of the wideband type, for the calibration of earphones used in audiometry*, IEC publication 318, Section 4, Geneva, 1970.

[10] GAYFORD (M. L.): Electroacoustics – Microphones, earphones and loudspeakers, *STC Monograph*, 1970.

Recommendation P.43 [1)]

INSTRUCTIONS FOR FORWARDING TO THE CCITT LABORATORY PRIMARY AND WORKING STANDARD SYSTEMS FOR CALIBRATION, AND COMMERCIAL TELEPHONE SYSTEMS FOR DETERMINING EQUIVALENT VALUES (AND OTHER PARAMETERS, IF REQUESTED)

(Malaga-Torremolinos, 1984)

Administrations, organizations and operating agencies are requested to follow the instructions given below when they forward to the CCITT Laboratory their standard systems for calibration, or commercial telephone systems for subjective or objective determination of their ratings in one of the following units: reference equivalents, relative equivalents (see Recommendation P.72) or loudness ratings (see Recommendation P.76).

Other objective parameters, e.g. sensitivity/frequency measurements of the systems or their particular parts can be measured on request as well by the CCITT Laboratory. From such measurements, loudness ratings can be derived by calculation in accordance with Recommendation P.79.

1 Standard reference systems for calibration

If an Administration, organization or operating agency wishes to have its standard system calibrated or recalibrated, assuming that the system concerned can be transported without risk of damage, it should be supplied to the CCITT Laboratory with the necessary documentation and, if necessary, instructions for checking the various parts of the system (e.g. amplifiers, attenuators, microphone, receiver etc.).

If the volume meter associated with the system does not possess the basic characteristics of the speech voltmeter of NOSFER (the same as of ARAEN [1]), the volume meter must be sent to the CCITT Laboratory at the same time as the system itself.

The method of reading the volume meter should be indicated. During the measurements, the CCITT Laboratory will thus be able to calibrate the volume meter using subjective and objective methods to determine the adjustment corresponding to the "standard" speech power for telephonometric measurements.

1) This Recommendation replaces former Recommendations P.43 and P.47.

The standard system can be calibrated for evaluation of one of the following units: reference equivalents, R25 equivalents (values theoretically equal to corrected reference equivalents) or loudness ratings. In the last case, the national Intermediate Reference System (see Recommendation P.48) has to be calibrated as well. The calibration of any standard system sent to the CCITT Laboratory is obtained through subjective tests and objective measurements.

2 Commercial telephone systems

Determinations of ratings of any commercial telephone system are not calibration measurements; their aim is to determine these values by direct comparison with the master standard reference system, which by definition has the rating equal to zero. As for national standard system calibration, commercial telephone systems can be evaluated in three different units, namely: reference equivalents, R25 equivalents and loudness ratings.

Reference equivalents are normally determined subjectively for sending, receiving and side-tone paths, whereas R25 equivalents only for sending and receiving paths of the telephone set under test, by comparison with corresponding paths of NOSFER.

Loudness ratings of a telephone system are evaluated subjectively using as a reference system NOSFER and Intermediate Reference System (IRS).

Loudness ratings can be evaluated for send and receive paths, as well as for the junction and overall system. The side-tone path evaluation will be possible in the near future as well.

Loudness ratings of different paths of a telephone system can also be evaluated in the CCITT Laboratory through objective measurement, including side-tone path (STMR).

To evaluate any of the ratings, as above, of a commercial telephone system, in addition to telephone sets to be tested, some accessories should also be dispatched to the CCITT Laboratory, as described below.

2.1 *Items to be dispatched to the CCITT Laboratory*

2.1.1 *Telephone sets*

Preferably three samples of one type of a telephone system should be sent if a reasonably reliable evaluation of the telephone system is to be made. Less than three samples could be evaluated, but if more than three samples of the same type were required to be tested, it should be agreed in advance.

2.1.2 *Feeding bridge*

At least one d.c., feeding bridge, suitably constructed for laboratory use, should be sent. A circuit diagram giving the value of the d.c. voltage and its polarity and the values of the resistive and capacitive components should be shown on the bridge.

A list of d.c. feeding bridges which are already available in the CCITT Laboratory is given in Annex A. If possible, the party may indicate its choice and avoid sending its own bridge. Attention is drawn to the fact that this list does not contain inductance (L) information of the bridges.

2.1.3 *Artificial subscriber lines*

Artificial subscriber lines, if required to be used, should be sent. They should be constructed in such a way that they can readily be used during the laboratory tests. The CCITT Laboratory has no artificial line at its disposal.

2.1.4 *Reference equivalent guard-ring*

If the reference equivalent or R25 equivalent values are to be evaluated, the reference equivalent guard-ring (see Recommendation P.72, § 4), appropriate to the handsets of the tested sets should be dispatched as well. It is advisable that a special gauge (as shown in Annex B) be constructed and used to verify the correctness of a reference equivalent guard-ring for a particular type of handset.

2.1.5 *Ear-cap centering ring*

If the loudness ratings have to be evaluated subjectively, or any objective measuring with a telephone set has to be carried out, then the earcap centering ring compatible with the shape and dimensions of the earcap piece of the handset of the telephone set should accompany the telephone sets. The centering ring can be manufactured using a bakelite, perspex, or any other rigid material. It is essential that its external shape be circular with a diameter of 85 mm and the internal diameter be compatible with the shape of the earcap of the handset in such a manner that the geometrical center of the opening is the same as the ear reference point (ERP) of given receiver (see Annex C).

2.1.6 *Spare handset*

An extra handset (of the same type but without capsules and not connected to any set but fitted with appropriate cord and the capsule holders) is to be sent as well, if the side-tone path equivalent value has to be evaluated.

2.1.7 *Spare parts*

It is advisable that some spare parts, e.g. two of each type of transmitter and receiver capsules, be sent along with the telephone sets. If during the test any of the capsules fitted in the handset were found to be faulty or damaged (during transport), the faulty capsule could then easily be replaced.

2.1.8 *Circuit diagram*

A circuit diagram of the telephone system under test should accompany the equipment. It is imperative that the subscriber line terminals be clearly marked on the diagram.

2.1.9 *Line current (d.c.)*

All d.c. currents relating to different artificial subscriber line conditions to be employed during the measurements should be indicated either in the correspondence or within the dispatch. This will enable the CCITT Laboratory to check whether the circuit is correct before starting to carry out tests or measurements.

2.2 *Subjective test team*

Normally six operators, divided into two sub-groups of three each, carry out the tests giving rise to twelve talker-listener pair ratings for one test condition. This method of evaluation is known as the 3/6 method. In the absence of one operator of the team, five operators carry out the tests, each one talking to the other four, giving rise to twenty talker-listener pair ratings. This method of evaluation is known as the 5/5 method. The mean of twelve or twenty particular ratings respectively, give the mean value of reference equivalent, R25 equivalent or loudness rating. Both the 3/6 and 5/5 methods are taken as comparable (i.e., the standard deviation in both cases is similar). For details, see Recommendations P.72 and P.76.

3 Technical report

At the end of the tests and/or the objective measurements, a confidential technical report is written giving all the details of the tests, e.g. the method used, results obtained and any other special points. Seven copies of this technical report are normally sent to the customer concerned or on his written request to any other address.

4 Returning measured items and accessories

Once the tests are carried out, the standard system or the telephone sets along with all other accessories (including spare parts if any) are returned to the customer as soon as possible. Transport is to be paid by the customer.

5 Charges for subjective tests and objective measurements

All subjective tests and objective measurements carried out on the request of Administrations, organizations or acting agencies are charged for fees which take into account the cost per hour of work of the CCITT Laboratory team. The units of charges are fixed by the Administrative Council of ITU in Swiss francs and may be revised from time to time.

The total cost depends on hours of work, or in case of subjective tests, on the quantity of the test conditions. Additional fees can be added to the total cost, such as the fees of the transport agency, some repairs or corrections needed so that the work with given systems can be correctly carried out, or the cost of the editing of the technical report with results. There is also a distinction between the charges applied for "Members" and "Non-members" of the ITU.

Members in this respect are defined as the PTT or Telecommunications Administration of the Member countries of ITU or any manufacturing concern or scientific or engineering organizations which subscribe annually to participate in the meetings of CCITT or those of any other permanent organ of ITU.

As the charges can vary according to the decision of the Administrative Council of ITU, they are presented in a separate document issued by CCITT each time the new tariff is applied.

6 General comment

The object of the general instructions given above is to guide the customer interested in the subject. When the customer concerned wishes to have the work carried out in the CCITT Laboratory, he should get in touch with the CCITT Laboratory before sending his devices and accessories, so that the technical and experimental conditions of the work as well as the date for it may be fixed in advance.

ANNEX A

(to Recommendation P.43)

D.c. feed bridges available in the CCITT Laboratory

TABLE A-1/P.43

D.C. feed bridge No.	R (Ω)	C (μF)	Nominal voltage (V)	Provenance of the bridge
1	2 × 150	2 × 2.0	60	France
2	2 × 200	2 × 2.0	48	Japan
3	2 × 200	2 × 2.2	48	ATT (USA)
4	2 × 200	2 × 2.0	50	BT (UK)
5	2 × 220	2 × 2.0	48	ITT (UK)
6	2 × 250	2 × 2.2	48	Chile
7	2 × 250	2 × 110.0	48	Norway
8	2 × 360	2 × 2.0	48	Switzerland
9	2 × 400	2 × 2.2	48	ITT (UK)
10	2 × 490	2 × 2.0	36	Sweden
11	2 × 500	2 × 3.8	60	Poland
12	2 × 500	2 × 4.0	60	Italy

Note 1 – R values measured using a d.c. ohmeter.

Note 2 – C values measured directly, using the impedance analyzer type HP 4192A, with f = 1 kHz, voltage output = 0.1 V, set to the serial mode.

ANNEX B

(to Recommendation P.43)

Gauge for circular earcaps (as used in the CCITT Laboratory) for verifying the position of reference equivalent rings

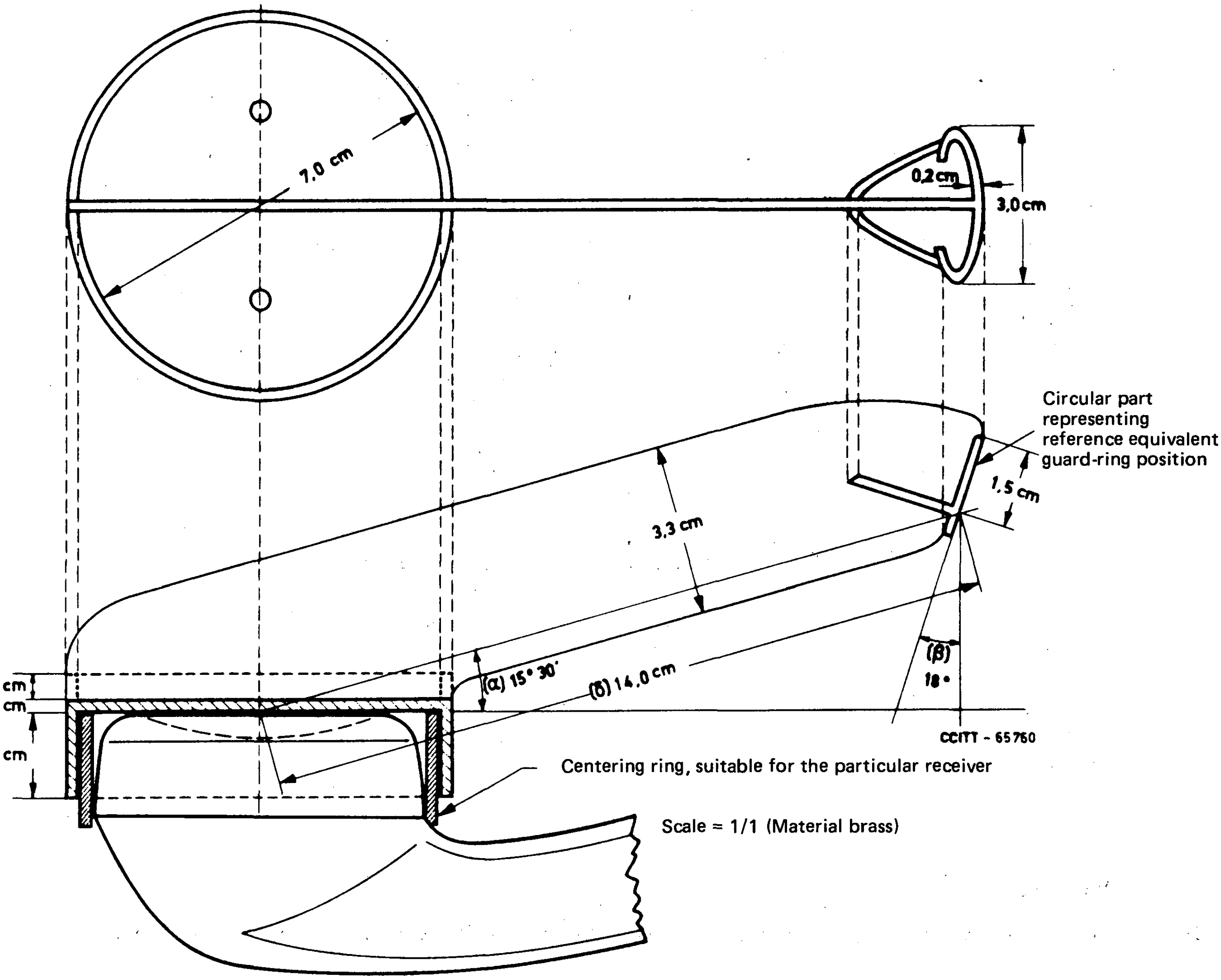


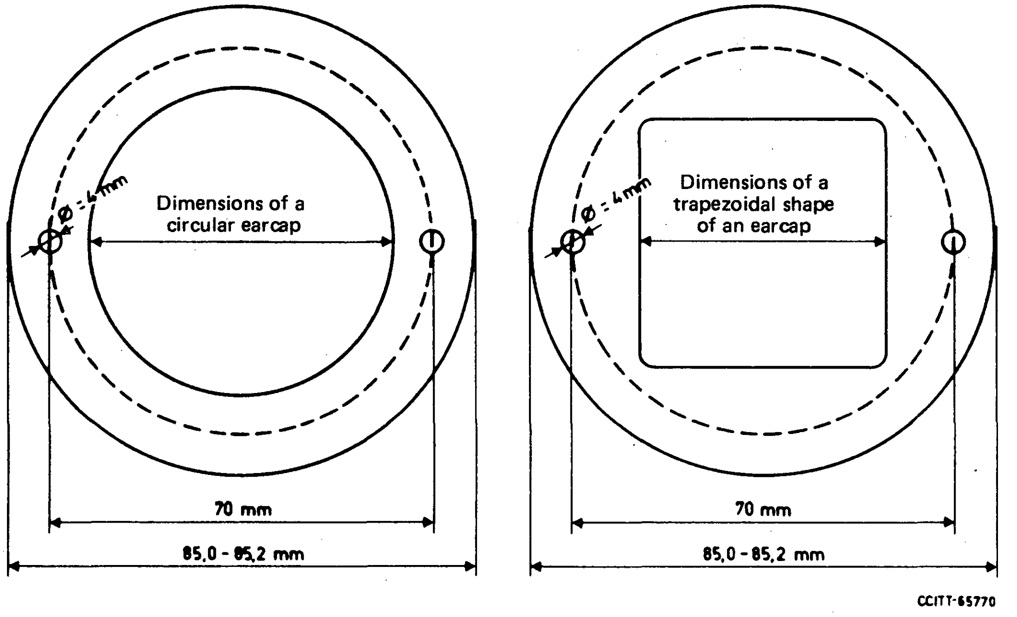
FIGURE B-1/P.43

ANNEX C

(to Recommendation P.43)

Earcap centering ring

Two examples of earcap centering rings to be fitted to the artificial ear support of the artificial head (*jig*), centering the receiver of a handset of measured telephone set are given in Figure C-1/P.43.



Note – Material: bakelite, perspex or any other similar material.

FIGURE C-1/P.43

Reference

[1] *ARAEN volume meter or speech voltmeter*, White Book, Vol. V, Supplement No. 10, ITU, Geneva, 1969.

Recommendation P.48

SPECIFICATION FOR AN INTERMEDIATE REFERENCE SYSTEM

(Geneva, 1976; amended at Geneva, 1980 and at Malaga-Torremolinos, 1984)

Summary

This Recommendation intends to specify the intermediate reference system (IRS) to be used for defining loudness ratings. The description should be sufficient to enable equipment having the required characteristics to be reproduced in different laboratories and maintained to standardized performance.

1 Design objectives

The chief requirements to be satisfied for an intermediate reference system to be used for tests carried out on handset telephones [1] are as follows:

a) the circuit must be stable and specifiable in its electrical and electro-acoustic performance. The calibration of the equipment should be traceable to national standards;

b) the circuit components that are seen and touched by the subjects should be similar in appearance and "feel" to normal types of subscribers' equipment;

c) the sending and receiving parts should have frequency bandwidths and response shapes standardized to represent commercial telephone circuits;

d) the system should include a junction which should provide facilities for the insertion of loss, and other circuit elements such as filters or equalizers;

e) the system should be capable of being set up and maintained with relatively simple test equipment.

Note – The requirements of a) to d) have been met in the initial design of the IRS by basing the sending and receiving frequency responses on the mean characteristics of a large number of commercial telephone circuits and confining the bandwidths to the nominal range 300-3400 Hz.

Since the detailed design of an IRS may vary between different Administrations, the following specification defines only those essential characteristics required to ensure standardization of the performance of the IRS.

The principles of the IRS are described and its nominal sensitivities are given in §§ 2, 3, 4 and 5 below; requirements concerning stability, tolerances, noise limits, crosstalk and distortion are dealt with in §§ 6 to 9 below. Some information concerning secondary characteristics is given in § 10 below.

Certain information concerning installation and maintenance are given in [1].

2 Use of the IRS

The basic elements of the IRS comprise:

a) the sending part,

b) the receiving part,

c) the junction.

When one example each of a), b) and c) are assembled, calibrated and interconnected, a reference (unidirectional) speech path is formed, as shown in Figure 1/P.48. For performing loudness rating determinations, suitable switching facilities are also required to allow the reference sending and receiving parts to be interchanged with their commercial counterparts.

---

1) For other types of telephone, e.g. headset or loudspeaking telephone, a different IRS will be required. The IRS is specified for the range 100-5000 Hz. The nominal range 300-3400 Hz specified is intended to be consistent with the nominal 4 kHz spacing of FDM systems, and should not be interpreted as restricting improvements in transmission quality which might be obtained by extending the transmitted frequency bandwidth.

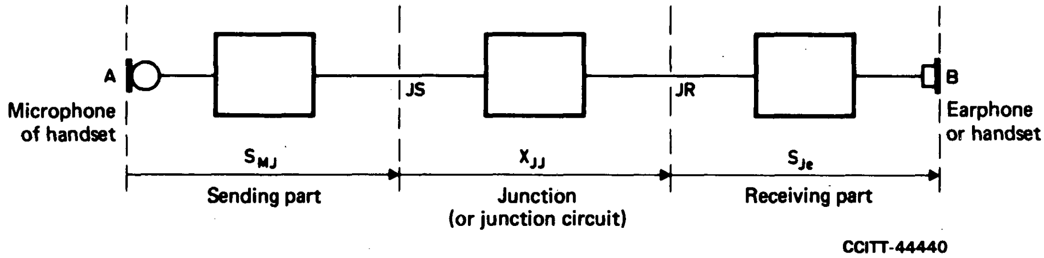
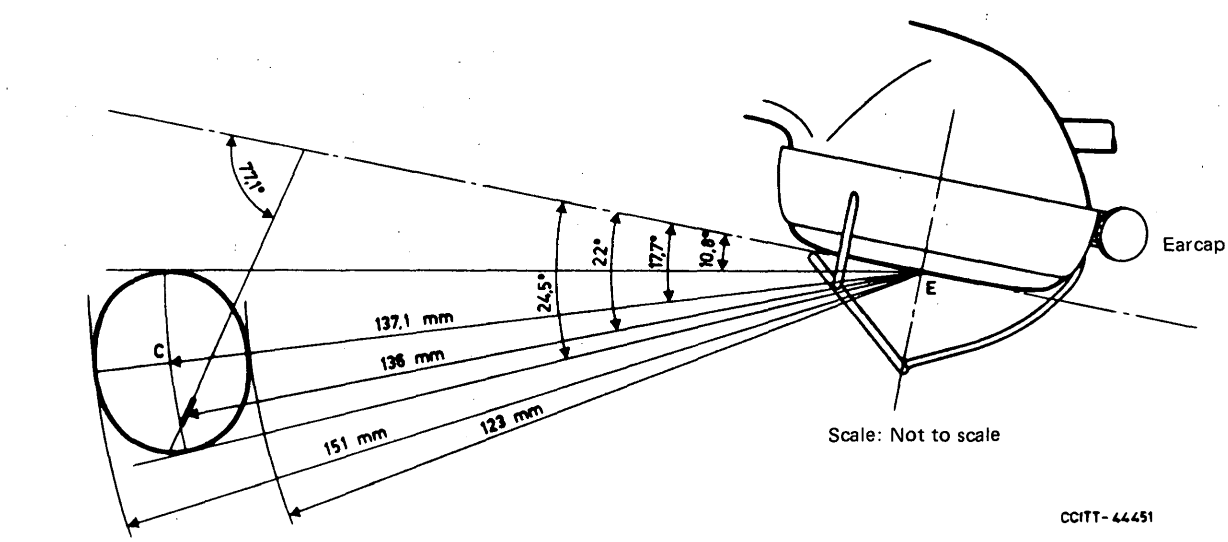


FIGURE 1/P.48

Composition of the complete intermediate reference system

3 Physical characteristics of handsets

The sending and receiving parts of an IRS shall each include a handset symmetrical about its longitudinal place and the profile produced by a section through this plane should, for the sake of standardization, conform to the dimensions indicated in Figure 2/P.48. In practice, any convenient form may be considered use being made, for example, of handsets of the same type as those used by an Administration in its own network. The general shape of the complete handset shall be such that, in normal use, the position of the earcap on the ear shall be as definite as possible, and not subject to excessive variation.



Note 1 – The guard-ring is shown in the 'loudness rating guard-ring' position.

Note 2 – The boundary of the ellipse shown encloses 80 % of observations on a sample of heads.

Note 3 – The centre C of the ellipse is located as shown above.

Note 4 – The minor axis of the ellipse is 28 mm and is colinear with the line joining the centre C of the ellipse with the ear reference point E.

Note 5 – The major axis of the ellipse is 33 mm.

Note 6 – The contour of the mouthpiece shall preferably just touch the ellipse. In any case, it should not overlap, or be separated from the ellipse by more than 5 mm.

FIGURE 2/P.48

Location of ellipse defining certain preferred dimensions for the IRS handset

The microphone capsule, when placed in the handset, shall be capable of calibration in accordance with the method described in Recommendation P.64. The earcap shall be such that it can be sealed on the circular knife-edge of the IEC/CCITT artificial ear for calibration in accordance with IEC 318, and the contour of the earcap shall be suitable for defining the ear reference point as described in Annex A to Recommendation P.64.

Transducers shall be stable and linear, and their physical design shall be such that they can be fitted in the handset chosen. A handset shall always contain both microphone and earphone capsules, irrespective of whether either is inactive during tests. The weight of a handset, so equipped, shall not exceed 350 g.

4 Subdivision of the complete IRS and impedances at the interfaces

Figure 1/P.48 shows the composition of the complete IRS, subdivided as specified in § 2 above. The principal features of the separate parts are considered below.

4.1 *Sending part*

The sending part of the IRS is defined as the portion A-JS extending from the handset microphone A to the interface with the junction at JS. The sending part shall include such amplification and equalization as necessary to ensure that the requirements of §§ 5.1 and 7 below are satisfied.

The return loss of the impedance at JS, towards A, against $600\underline{/0^\circ}$ ohms, when the sending part is correctly set up and calibrated, shall be not less than 20 dB over a frequency range 200-4000 Hz, and not less than 15 dB over a frequency range 125-6300 Hz.

4.2 *Receiving part*

The receiving part of the IRS is defined as the portion JR-B extending from the interface with the junction at JR to the handset earphone at B. The receiving part shall include such amplification and equalization as necessary to ensure that the requirements of §§ 5.2 and 7 below are satisfied.

The return loss of the impedance at JR, towards B, against $600\underline{/0^\circ}$ ohms, when the receiving part is correctly set up and calibrated, shall be not less than 20 dB over a frequency range 200-4000 Hz, and not less than 15 dB over a frequency range 125-6300 Hz.

4.3 *Junction*

For loudness balance and sidetone tests, the junction of the IRS shall comprise means of introducing known values of attenuation between the sending and receiving parts, and shall consist of a calibrated 600 ohm attenuator having a maximum value of not less than 100 dB

(e.g. 10×10 dB + 10×1 dB + 10×0.1 dB)

and having a tolerance, when permanently fitted and wired in position in the equipment, of not more than ± 1% of the dial reading or 0.1 dB, whichever is numerically greater. Provision shall be made for the inclusion of additional circuit elements (e.g. attenuation/frequency distortion) in the junction. The circuit configuration of such additional elements shall be compatible both with that of the attenuator and the junction interfaces. The return loss of the junction against $600\underline{/0^\circ}$ ohms, both with and without any additional circuit elements, shall be not less than 20 dB over a frequency range 200-4000 Hz, and not less than 15 dB over a frequency range 125-6300 Hz. For these tests, the port other than that being measured shall be closed with $600\underline{/0^\circ}$ ohms.

5 Nominal sensitivities of sending and receiving parts

The absolute values given below are provisional and may require changes to some extent as a result of the study of Question 19/XII [2].

5.1 *Sending part*

The sending sensitivity, S_{MJ} is given in Table 1/P.48, column (2) (see [3]).

5.2 *Receiving part*

The receiving sensitivity, S_{Je}, on a CCITT/IEC measured artificial ear (see Recommendation P.64) is given in Table 1/P.48, column (3) (see [3]).

6 Stability

The stability should be maintained, under reasonable ranges of ambient temperature and humidity, at least during the period between routine recalibrations. (See also [1].)

7 Tolerances on sensitivities of sending and receiving parts

This section specifies tolerances on:

a) the shape of the sensitivity/frequency characteristics of the send and receive parts of the IRS, and

b) the loudness weighted mean sensitivities.

7.1 *Shapes of sensitivity/frequency characteristics*

The shape of the sensitivity/frequency characteristics of the sending and receiving parts of the IRS shall be such that the limits specified in Table 2/P.48 are satisfied. In checking the shape, the mean values of sensitivity may be adjusted to best advantage.

TABLE 1/P.48

Nominal sending sensitivities and receiving sensitivities of the IRS

(These values were adopted provisionally)

Frequency (Hz)	S_{MJ}	S_{Je}
	dB V/Pa	dB Pa/V
(1)	(2)	(3)
100	−45.8	−27.5
125	−36.1	−18.8
160	−25.6	−10.8
200	−19.2	−2.7
250	−14.3	2.7
300	−11.3	6.4
315	−10.8	7.2
400	−8.4	9.9
500	−6.9	11.3
600	−6.3	11.8
630	−6.1	11.9
800	−4.9	12.3
1000	−3.7	12.6
1250	−2.3	12.5
1600	−0.6	13.0
2000	0.3	13.1
2500	1.8	13.1
3000	1.5	12.5
3150	1.8	12.6
3500	−7.3	3.9
4000	−37.2	−31.6
5000	−52.2	−54.9
6300	−73.6	−67.5
8000	−90.0	−90.0

7.2 *Tolerances on mean values of sensitivity*

The gain setting in the send and receive parts of the IRS shall be such that the loudness weighted mean sensitivities shall be within ± 0.2 dB of the loudness weighted mean of the sensitivities given in Table 1/P.48. The loudness weighted means should be determined in accordance with the principles laid down in Recommendation P.79.

TABLE 2/P.48

Tolerances on shapes of sending and receiving sensitivities

Frequency (Hz)	Relative sensitivity (dB)	
	Sending part	Receiving part
180- 225	±2.0	−13.0, +2.0
225- 280	±2.0	−7.5, +2.0
280-2800	±2.0	±2.0
2800-3550	±2.5	±3.0
3550-4500	±6.7	±8.2

8 Noise limits

It is important that the noise level in the system be well controlled. See [4].

9 Nonlinear distortion

In order to ensure that nonlinear distortion will be negligible with the vocal levels normally used for loudness rating, requirements in respect of distortion shall be met.

10 Complete specifications

Certain secondary characteristics of an IRS may be included in Administrations' specifications. Particularly, special care must be given to adjustable components, stability and tolerances, crosstalk, installation and maintenance operations, etc. Reference [1] gives some guidance on these points.

References

[1] *Precautions to be taken for correct installation and maintenance of an IRS*, Orange Book, Vol. V, Supplement No. 1, ITU, Geneva, 1977.

[2] CCITT – Question 19/XII, Contribution COM XII-No. 1, Study Period 1985-1988, ITU, Geneva, 1985.

[3] *Precautions to be taken for correct installation and maintenance of an IRS*, Orange Book, Vol. V, Supplement No. 1, § 9.2, ITU, Geneva, 1977.

[4] *Ibid.*, § 5.

PAGE INTENTIONALLY LEFT BLANK

PAGE LAISSEE EN BLANC INTENTIONNELLEMENT

SECTION 4

OBJECTIVE MEASURING APPARATUS

Recommendation P.51

ARTIFICIAL VOICES, ARTIFICIAL MOUTHS, ARTIFICIAL EARS

(amended at Mar del Plata, 1968, Geneva, 1972, 1976, 1980 and Malaga-Torremolinos, 1984)

The CCITT,

considering

(a) that it is highly desirable to design an apparatus for telephonometric measurements such that in future all these measurements may be made with it, without using the human mouth and ear;

(b) that the standardization of the artificial voices, mouths and ears used in the construction of such apparatus is a subject for general study by the CCITT;

(c) that the standardization of an accurate artificial mouth can only be obtained after conclusion of the studies undertaken by various Administrations, comparison of their results and study of the models to check their characteristics;

(d) that in the meantime it would be useful to issue a provisional Recommendation regarding a "sound source" designed in accordance with the sensitivity-frequency characteristics,

recommends

the use of the artificial ears described in § 1 of this Recommendation, and

recommends provisionally

a) the use of the sound source described in § 2;

b) the use of the artificial voice described in § 3.

Note 1 – The above is still on the understanding that it is considered essential that all reference equivalent measurements at the CCITT Laboratory should continue to be made with the human mouth and ear.

Note 2 – Administrations may, if they wish, use in the future, devices which they have been able to construct for large-scale testing of telephone apparatus supplied by manufacturers, provided that the results obtained with these devices are in satisfactory agreement with results obtained by real voice-ear methods.

Note 3 – The Plenary Assembly at Copenhagen in 1936 considered that it would be of interest to deal separately with the design, on the one hand, of an artificial speech source and, on the other hand, of apparatus for producing a defined acoustic field according to certain specified conditions which will reproduce artificially a human mouth. The term "artificial voice" may be used for the former and "artificial mouth" for the latter.

1 Artificial ears recommended by the CCITT

Five types of artificial ears were defined by the International Electrotechnical Commission (IEC, TC29 Meeting at Liège 1960):

1) simple conventional type,
2) simple type used for reference equivalent measurements,
3) wideband type for audiometric measurements,
4) special type for calibrating insert earphones,
5) a type which faithfully reproduces the characteristics of the average human ear, for use in laboratory.

Artificial ear type 1 (or reference coupler) is the subject of the publication cited in [1]; this coupler is different from the "CCITT reference coupler".

Type 3 is covered by IEC Recommendation 318 [2] and type 4 by IEC Recommendation 711 [3].

The study of type 2 artificial ear and the study of acoustic leaks have been deleted from the programme of work of the IEC and are carried on by the CCITT.

(Artificial ear of type 5 is the object of further study in the IEC.)

The CCITT recommends

(a) the use of an artificial ear conforming to IEC 318 [2] for measurements on supra-aural earphones, e.g. handsets;

(b) the use of an insert ear simulator conforming to IEC 711 [3] for measurements on insert earphones, e.g. some headsets.

Note – For the calibration of NOSFER earphones with rubber earpads (types 4026A and DR 701) the method detailed in Annex B to Recommendation P.42 should be used.

2 Sound source provisionally recommended by the CCITT

2.1 *Introduction*

Before recommending a particular type of artificial mouth as suitable for objective telephonometric measurements, it is proposed that, as a first stage, experience should be gained in the use of one form of sound source to determine the shape of the sensitivity/frequency characteristic of a commercial sending system to be obtained whatever type of microphone inset is used, this sound source can only be used for handsets.

Such a sound source is required to permit useful comparisons to be made between the results obtained in various laboratories. This advantage already exists for comparison of the sensitivity/frequency curves of earphones since the adoption by the CCITT at Mar del Plata of the IEC-318 model artificial ear.

It would be desirable to supplement existing documentation on the human mouth.

Note – It is not proposed that the choice of sound source should prejudice the definition of a more precise artificial mouth that can be used universally for measuring objective ratings.

2.2 *Acoustic characteristics of the sound source*

2.2.1 The sound source must permit calibration of microphone at short distances.

2.2.2 At the measuring distances normally used, the acoustic properties should be close to those of the average human mouth; in particular, the law of decrease in sound pressure on the axis should be close to that of the average human mouth from a distance of about 10 mm onwards from a plane called the lip plane of the source.

Table 1/P.51 shows the sound pressures measured by some Administrations at points along the axis and expressed in relation to the sound pressure at 25 mm from the lip plane. The sound pressures should be measured with a very small (say about 6 mm diameter) microphone or a probe microphone.

2.2.3 The directivity, in a region of space around the axis, should be close to that of the average human mouth.

TABLE 1/P.51

Distance from lip plane (mm)	Relative sound pressure level (dB relative to the sound pressure 25 mm from the lip plane)		
	U.K. Post Office	Chile Telephone Co.	L. M. Ericsson
10	+4.8	+5.5	+4.6
20	+1.5	+1.5	+1.3
25	0	0	0
40	−3.3	−3.3	−3.4
60	−6.5	(See Note)	

Note – **Beyond 40 mm, the sound pressure can be assumed to be inversely proportional to the distance from an equivalent point source lying 6 mm behind the lip plane.**

2.2.4 For the measurements obtained with different specimens of the source to be comparable, it is necessary to define a reference point on the main axis at which the characteristics of the source will be checked and which will serve as reference in inter-laboratory tests. It is suggested that a point on the axis 25 mm from the lip plane would be suitable.

2.2.5 As a preliminary indication, the sound source should be able to deliver to the above reference point acoustic pressure levels of not less than 90 dB [relative to $2 \cdot 10^{-5}$ Pa (Pascal)] in a frequency range comprising at least the 200-4000 Hz band. (Sound pressure levels up to 100 dB over the frequency range 100-8000 Hz would be desirable.)

2.2.6 The source should be stable and reproducible.

2.3 *Choice of a model*

The result of measurements made with the B & K 4219 source and the United Kingdom Post Office artificial mouth have shown good agreement between the two models. These results are not very different from the values measured for the human mouth as far as the distribution of acoustic pressure in a free field along the axis is concerned. These two models also meet the other specifications of § 2.2 above.

3 **Artificial voice**

The signal here described reproduces the characteristics of human speech, relevant to the characterization of linear and non-linear telecommunication systems and devices, intended for the transduction or transmission of speech. It is known that for some purposes, such as objective loudness rating measurements, less sophisticated signals can be used as well. Examples of such signals are pink noise or spectrum shaped gaussian noise, which nonetheless cannot be referred to as "artificial voices" for the purpose of this Recommendation.

3.1 *Introduction*

An artificial voice is a signal that can be mathematically defined and that reproduces time and spectral characteristics of speech which significantly affect the performances of telecommunication systems [4].

The following characteristics have been considered:

a) long-term average spectrum,

b) instantaneous amplitude distribution,

c) voiced and unvoiced structure of speech waveform,

d) syllabic envelope.

An example of the generation process for the artificial voice is described in Supplement No. 7.

3.2 *Scope, purpose and definition*

3.2.1 *Scope and purpose*

The artificial voice is aimed at reproducing the characteristics of real speech over the bandwidth 100 Hz-8 kHz. It can be utilized for characterizing many devices, e.g. microphones, loudspeaking telephone sets, non-linear coders, echo controlling devices, syllabic compandors, non-linear systems in general.

Of course when a particular system is tested, the characteristics of the transmission path preceding it are to be considered. The actual test signal has then to be produced as the convolution between the artificial voice and the path response.

3.2.2 *Definition*

The **artificial voice** is a signal, mathematically defined, which reproduces all human speech characteristics, relevant to the characterization of linear and non-linear telecommunication systems. It is then intended to give a satisfactory correlation between objective measurements and real speech tests.

3.3 *Terminology*

The artificial voice can be produced both as an electric or as an acoustic signal, according to the system or device under test (e.g. communication channels, coders, microphones). The following definitions apply with reference to Figure 1/P.51.

3.3.1 **electrical artificial voice**

The artificial voice produced as an electric signal, used for testing transmission channels or other electric devices.

3.3.2 **artificial mouth excitation signal**

A signal applied to the artificial mouth in order to produce the acoustic artificial voice. It is obtained by equalizing the electrical artificial voice for compensating the sensitivity/frequency characteristic of the mouth.

Note – The equalization depends on the particular artificial mouth employed and can be accomplished electrically or mathematically within the signal generation process.

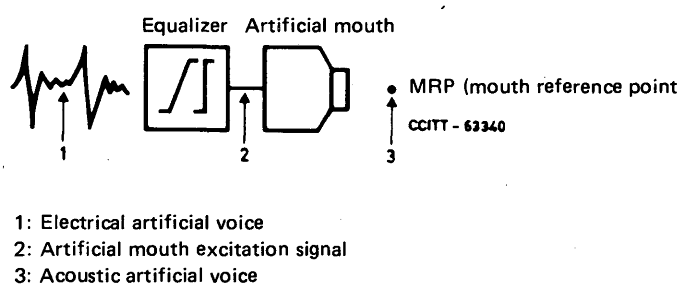


FIGURE 1/P.51

3.3.3 **acoustic artificial voice**

It is the acoustic signal at the MRP (mouth reference point) of the artificial mouth and has to comply with the same time and spectral requirements of the electrical artificial voice.

3.4 *Characteristics*

3.4.1 *Long-term average spectrum*

The third octave filtered long-term average spectrum of the artificial voice is given in Figure 2/P.51 and Table 2/P.51, normalized for a wideband sound pressure level of −4.7 dB Pa. The Table is calculated from the theoretical equation reported in [5].

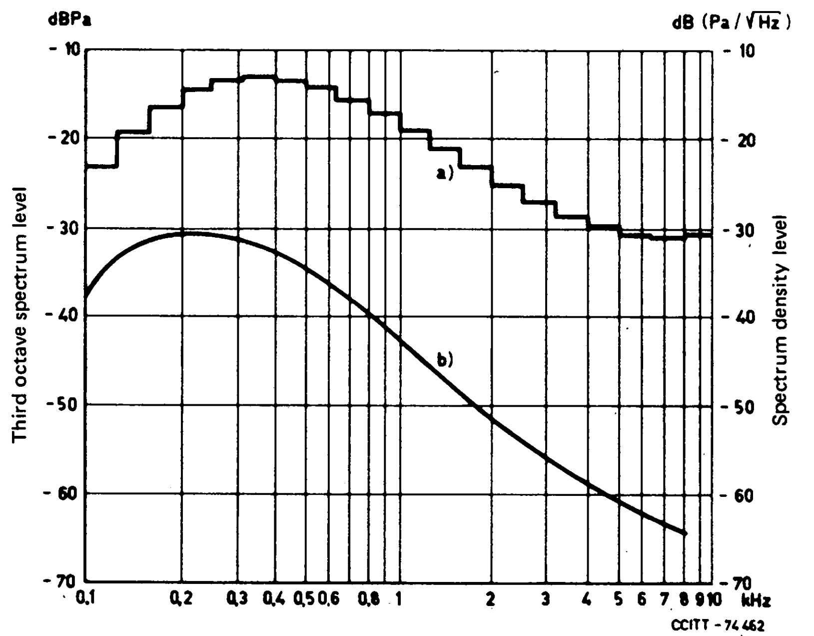
Note – The values of the long-term spectrum of the artificial voice at the mouth reference point can be derived from the equation:

$$S(f) = -376.44 + 465.439 (\log_{10} f) - 157.745 (\log_{10} f)^2 + 16.7124 (\log_{10} f)^3 \qquad (3\text{-}1)$$

where $S(f)$ is the spectrum density in dB relative to 1 pW/m^2 sound intensity per Hz at frequency f. The definitive frequency range is from 100 to 8000 Hz.

The curve of the spectrum is shown in Figure 2/P.51. The values of $S(f)$ at 1/3 octave ISO frequencies are given in the fourth column of Table 2/P.51. The tolerances are given in the fifth column of Table 2/P.51.

The total sound pressure level of the spectrum defined in Equation (3-1) is −4.7 dB Pa. However, this spectrum is also applicable for the levels from −19.7 to +10.3 dB Pa. In oher words the first term of Equation (3-1) may be from −391.44 to −361.44.



a) Third octave spectrum [Column (3), Table 2/P.51]
b) Spectrum density [Column (3)–(2), Table 2/P.51]

FIGURE 2/P.51

Long-term spectrum of artificial voice

TABLE 2/P.51

Long-term spectrum of the artificial voice

1/3 octave center frequency (Hz) (1)	Bandwidth correction factor 10 $\log_{10} \Delta f$ (dB) (2)	Sound pressure level (third octave) (dB Pa) (3)	Spectrum density (dB) (3) – (2)	Tolerance (dB)
100	13.6	−23.1	−36.7	–
125	14.6	−19.2	−33.8	+3, −6
160	15.6	−16.4	−32	+3, −6
200	16.6	−14.4	−31	+3, −6
250	17.6	−13.4	−31	±3.0
315	18.6	−13.0	−31.6	±3.0
400	19.6	−13.3	−32.9	±3.0
500	20.6	−14.1	−34.7	±3.0
630	21.6	−15.4	−37	±3.0
800	22.6	−17.0	−39.6	±3.0
1000	23.6	−18.9	−42.5	±3.0
1250	24.6	−21.0	−45.6	±3.0
1600	25.6	−23.0	−48.6	±3.0
2000	26.6	−25.1	−51.7	±3.0
2500	27.6	−26.9	−54.5	±3.0
3150	28.6	−28.6	−57.2	±3.0
4000	29.6	−29.8	−59.4	±6.0
5000	30.6	−30.6	−61.2	±6.0
6300	31.6	−30.9	−62.5	±6.0
8000	32.6	−30.5	−63.1	–

3.4.2 *Instantaneous amplitude distribution*

The probability density distribution of the instantaneous amplitude of artificial voice is shown in Figure 3/P.51 [6].

3.4.3 *Segmental power level distribution*

The segmental power level distribution of the artificial signal, measured on time windows of 16 ms is shown in Figure 4/P.51. The upper and lower tolerance limits are reported as well.

Note – The upper tolerance limit represents the typical segmental power level distribution of normal conversation, while the lower limit represents continuous speech (telephonometric phrases) [7], [8].

3.4.4 *Spectrum of the modulation envelope*

The spectrum of the modulation envelope waveform is shown in Figure 5/P.51 and should be reproduced with a tolerance of ± 5 dB on the whole frequency range.

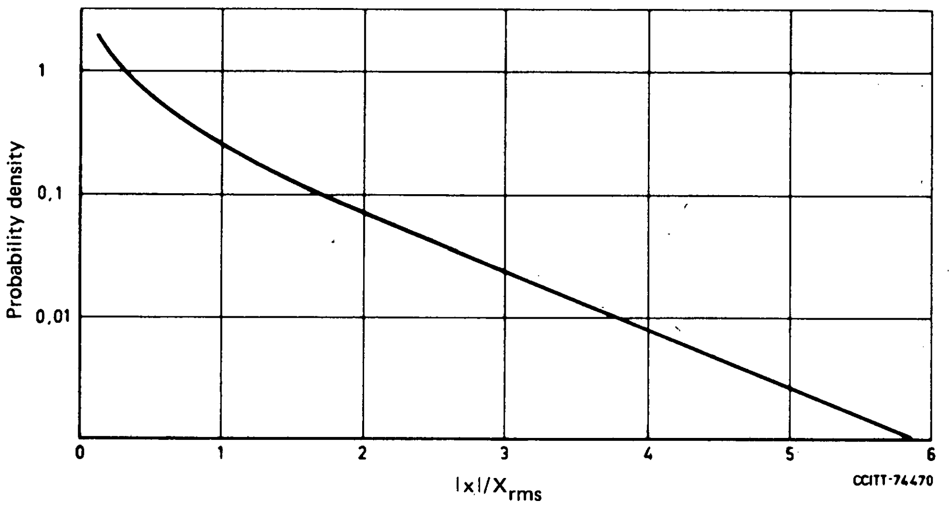
3.4.5 *Spectrum of unvoiced sounds*

The power density spectrum of unvoiced sounds inserted into the artificial signal is shown in Figure 6/P.51.

Note – The recommended spectrum is referred to the fricative sound of substained "s" [7].

3.4.6 *Time convergence*

Artificial voice must exhibit characteristics as close as possible to real speech. Particularly it should be possible to obtain the long-term spectrum and amplitude distribution characteristics in 10 seconds.



|x| : absolute value of the instantaneous amplitude
X_{rms}: root mean square of the signal

FIGURE 3/P.51

Instantaneous amplitude distribution

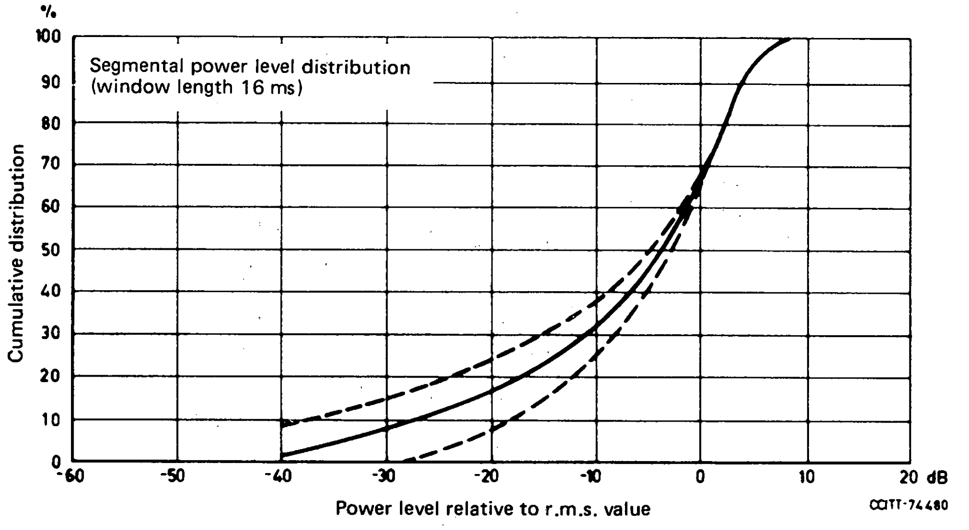


FIGURE 4/P.51

Segmental power level distribution

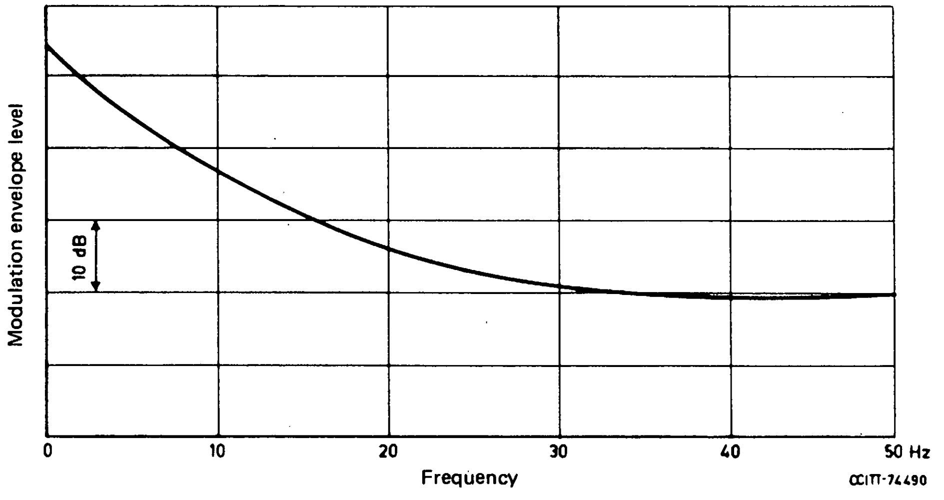


FIGURE 5/P.51

Spectrum of modulation envelope

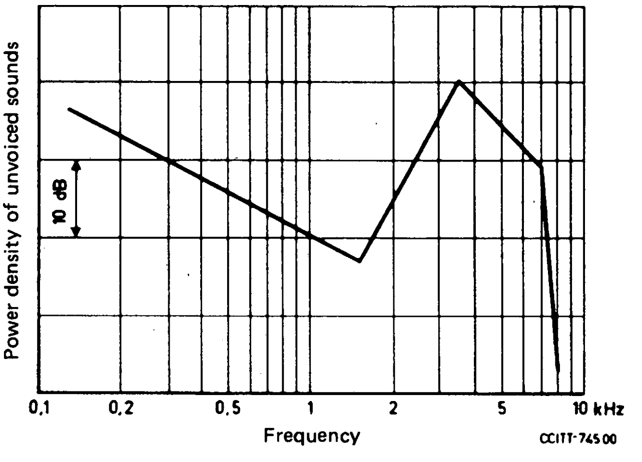


FIGURE 6/P.51

Power density spectrum of unvoiced sounds

References

[1] International Electrotechnical Commission Report *IEC provisional reference coupler for the calibration of earphones used in audiometry*, IEC publication 303, Geneva, 1970.

[2] International Electrotechnical Commission Recommendation *An artificial ear, of the wideband type, for the calibration of earphones used in audiometry*, IEC publication 318, Geneva, 1970.

[3] International Electrotechnical Commission Recommendation *Occluded ear simulator for the measurement of earphones coupled to the ear by ear insert*, IEC Publication 711, Geneva, 1981.

[4] CCITT – Contribution COM XII-No. 76, Study Period 1981-1984.

[5] CCITT – Contribution COM XII-No. 108, Study Period 1981-1984.

[6] CCITT – Contribution COM XII-No. 11, Study Period 1981-1984.

[7] CCITT – Contribution COM XII-No. 150, Study Period 1981-1984.

[8] CCITT – Contribution COM XII-No. 132, Study Period 1981-1984.

Bibliography

BRÜEL (P. V.), FREDERIKSEN, (E.) and RASMUSSEN (G.): Artificial ears for the calibration of earphones of the external type; *B. & K. Tech. Rev.* No. 4 (1961) and No. 1 (1962).

DELANY (M. E.): The acoustical impedance of human ears; *J. Sound Vib. 1* (1964), 455.

DELANY (M. E.), WHITTLE (L. S.), COOK (J. P.) and SCOTT (V.): Performance studies on a new artificial ear; *Acustica 18* (1967), 231.

ITHELL (A. H.): A determination of the acoustical input impedance characteristics of human ears; *Acustica 13* (1963), 311.

ITHELL (A. H.), JOHNSON (E. G. T.) and YATES (R. F.): The acoustical impedance of human ears and a new artificial ear; *Acustica 15* (1965), 109.

Recommendation P.52

VOLUME METERS

The CCITT considers that, in order to ensure continuity with previous practice, it is not desirable to modify the specification of the volume meter of the ARAEN employed at the CCITT Laboratory.

Table 1/P.52 gives the principal characteristics of various measuring devices used for monitoring the volume or peak values during telephone conversations or sound-programme transmissions.

The measurement of active speech level is at present under study in Question 26/XVII. A description of the principle of measurement is found in Supplement No. 8. Study Group XII intends to propose this supplement as a draft Recommendation. Administrations are invited to use and study it with a view to deciding whether or not it should become a Recommendation.

Note – Descriptions of the following devices are contained in the Supplements to *White Book*, Volume V:

- ARAEN volume meter or speech voltmeter: Supplement No. 10 [1].
- Volume meter standardized in the United States of America, termed the "VU meter": Supplement No. 11 [2].
- Peak indicator used by the British Broadcasting Corporation: Supplement No. 12 [3].
- Maximum amplitude indicator Types U 21 and U 71 used in the Federal Republic of Germany: Supplement No. 13 [4].

The volume indicator, SFERT, which formerly was used in the CCITT Laboratory is described in [5].

Comparative tests with different types of volume meters

A note which appears in [6] gives some information on the results of preliminary tests conducted at the SFERT Laboratory to compare the volume indicator with different impulse indicators.

The results of comparative tests made in 1952 by the United Kingdom Post Office appear in Supplement No. 14 [7].

TABLE 1/P.52

Principal characteristics of the various instruments used for monitoring the volume or peaks during telephone conversations or sound-programme transmissions

Type of instrument	Rectifier characteristic (see Note 3)	Time to reach 99% of final reading (milliseconds)	Integration time (milliseconds) (see Note 4)	Time to return to zero (value and definition)
(1) "Speech voltmeter" United Kingdom Post Office Type 3 (S.V.3) identical to the speech power meter of the ARAEN	2	230	100 (approx.)	equal to the integration time
(2) Vu meter (United States of America) (see Note 1)	1.0 to 1.4	300	165 (approx.)	equal to the integration time
(3) Speech power meter of the "SFERT volume indicator"	2	around 400 to 650	200	equal to the integration time
(4) Peak indicator for sound-programme transmissions used by the British Broadcasting Corporation (BBC Peak Programme Meter) (see Note 2)	1		10 (see Note 5)	3 seconds for the pointer to fall to 26 dB
(5) Maximum amplitude indicator used by the Federal German Republic (type U 21)	1	around 80	5 (approx.)	1 or 2 seconds from 100% to 10% of the reading in the steady state
(6) OIRT – Programme level meter:		for both types:		for both types:
type A sound meter		less than 300 ms for meters with pointer indication and less than 150 ms for meters with light indication	10 ± 5	1.5 to 2 seconds from the 0 dB point which is at 30% of the length of the operational section of the scale
type B sound meter			60 ± 10	

Note 1 – In France a meter similar to the one defined in line (2) of the table has been standardized.

Note 2 – In the Netherlands a meter (type NRU-ON301) similar to the one defined in line (4) of the table has been standardized.

Note 3 – The number given in the column is the index n in the formula $[V_{(output)} = V_{(input)}{}^n]$ applicable for each half-cycle.

Note 4 – The "integration time" was defined by the CCIF as the "minimum period during which a sinusoidal voltage should be applied to the instrument for the pointer to reach to within 0.2 neper or nearly 2 dB of the deflection which would be obtained if the voltage were applied indefinitely". A logarithmic ratio of 2 dB corresponds to a percentage of 79.5% and a ratio of 0.2 neper to a percentage of 82%.

Note 5 – The figure of 4 milliseconds that appeared in previous editions was actually the time taken to reach 80% of the final reading with a d.c. step applied to the rectifying/integrating circuit. In a new and somewhat different design of this programme meter using transistors, the performance on programme remains substantially the same as that of earlier versions and so does the response to an arbitrary, quasi-d.c. test signal, but the integration time, as here defined, is about 20% greater at the higher meter readings.

Note 6 – In Italy a sound-programme meter with the following characteristics is in use:
Rectifier characteristic: 1 (see Note 3).
Time to reach 99% of final reading: approx. 20 ms.
Integration time: approx. 1.5 ms.
Time to return to zero: approx. 1.5 s from 100% to 10% of the reading in the steady state.

References

[1] *ARAEN volume meter or speech voltmeter*, White Book, Vol. V, Supplement No. 10, ITU, Geneva, 1969.

[2] *Volume meter standardized in the United States of America, termed VU meter*, White Book, Vol. V, Supplement No. 11, ITU, Geneva, 1969.

[3] *Modulation meter used by the British Broadcasting Corporation*, White Book, Vol. V, Supplement No. 12, ITU, Geneva, 1969.

[4] *Maximum amplitude indicators, types U 21 and U 71 used in the Federal Republic of Germany*, White Book, Vol. V, Supplement No. 13, ITU, Geneva, 1969.

[5] *SFERT volume indicator*, Red Book, Vol V, Annex 18, Part 2, ITU, Geneva, 1962.

[6] CCIF *White Book*, Vol. IV, pp. 270-293, ITU, Bern, 1934.

[7] *Comparison of the readings given on conversational speech by different types of volume meter*, White Book, Vol. V, Supplement No. 14, ITU, Geneva, 1969.

Recommendation P.53

PSOPHOMETERS (APPARATUS FOR THE OBJECTIVE MEASUREMENT OF CIRCUIT NOISE)

Refer to Recommendation O.41, CCITT Red Book, Volume IV, Fascicle IV.4

Recommendation P.54

SOUND LEVEL METERS (APPARATUS FOR THE OBJECTIVE MEASUREMENT OF ROOM NOISE)

(amended at Mar del Plata, 1968 and Geneva, 1972)

The CCITT recommends the adoption of the sound level meter specified in [1] in conjunction, for most uses, with the octave, half, and third octave filters in accordance with [2].

References

[1] International Electrotechnical Commission Standard *Sound level meters*, IEC publication 651 (179), Geneva, 1979.

[2] International Electrotechnical Recommendation *Octave, half-octave and third-octave band filters intended for the analysis of sounds and vibrations*, IEC publication 225, Geneva, 1966.

APPARATUS FOR THE MEASUREMENT OF IMPULSIVE NOISE

(Mar del Plata, 1968)

Experiments have shown that clicks or other impulsive noises which occur in telephone calls come from a number of sources, such as faulty construction of the switching equipment, defective earthing at exchanges and electromagnetic couplings in exchanges or on the line.

There is no practical way of assessing the disturbing effect of isolated pulses on telephone calls. A rapid succession of clicks is annoying chiefly at the start of a call. It is probable that these series of clicks affect data transmission more than they do the telephone call and that connections capable of transmitting data, according to the noise standards now under study, will also be satisfactory for speech transmission.

In view of these considerations, the CCITT recommends that Administrations use the impulsive noise counter defined in Recommendation O.71 [1] for measuring the occurrence of series of pulses on circuits for both speech and data transmission.

Note – At the national level, Administrations might continue to study whether the use of this impulsive noise counter is sufficient to ensure that the conditions necessary to ensure good quality in telephone connections are met. In those studies, Administrations may use whatever measuring apparatus they consider most suitable – for example a psophometer with an increased overload factor – but the CCITT does not envisage recommending the use of such an instrument.

Reference

[1] CCITT Recommendation *Specification for an impulsive noise measuring instrument for telephone-type circuits*, Vol. IV, Rec. O.71.

SECTION 5

OBJECTIVE ELECTRO-ACOUSTICAL MEASUREMENTS

Recommendation P.61

METHODS FOR THE CALIBRATION OF CONDENSER MICROPHONES

(amended at Malaga-Torremolinos, 1984)

Primary and secondary calibrations of condenser microphones can be carried out using the methods described below.

1 Primary calibration by the reciprocity method

The recommended procedure for primary calibration of condenser microphones is the reciprocity calibration technique. A precision method for reciprocity pressure calibration is described in [1]. A simplified method, suitable for calibration over the frequency range of interest for telephonometric measurements, is given in [2]. Although the methods described are specifically for one-inch microphones, similar methods are applicable to half-inch microphones. Methods suitable for half-inch microphones are under study by IEC.

A precision method for free-field reciprocity calibration is given in [3]. Alternatively, the free-field correction curves given in [4] may be applied to the pressure calibration of one-inch condenser microphones to determine their free-field responses. The reciprocity free-field calibration method may in principle be extended to half-inch microphones. Free-field correction curves have not been standardized for half-inch microphones.

2 Secondary calibration by the comparison method

The secondary calibration of a condenser microphone may be achieved by direct comparison with a physically identical microphone having a known calibration. The procedure used is a modification of the "two microphones and auxiliary sound source" method described in [1] to [3]. The output of the calibrated microphone is first determined for a given drive level applied to the auxiliary sound source. The calibrated microphone is then replaced by the microphone to be calibrated, and its output is determined for the same drive level applied to the auxiliary sound source. The difference in level (in dB) between the outputs of the two microphones is then applied to the known calibration of the first microphone to determine the calibration of the second. The procedure is repeated at each frequency of interest.

3 Secondary calibration using pistonphones and other sound level calibrators

Secondary calibrations can also be made using pistonphones and other sound level calibrators which produce a known sound level. Such devices are often used to check the calibration of a microphone at a single frequency [1)]. Care should be taken to follow the manufacturer's instructions when using such devices; in particular, it may be necessary to apply corrections for barometric pressure, coupler volume, microphone type, etc. Standardization of these calibrators is currently under study by the IEC.

---

1) Calibrations with an accuracy of $\pm$ 0.3 dB are possible.

References

[1] International Electrotechnical Commission *Precision method for pressure calibration of one-inch standard condenser microphones by the reciprocity technique*, IEC publication 327, Geneva, 1971.

[2] International Electrotechnical Commission *Simplified method for pressure calibration of one-inch condenser microphones by the reciprocity technique*, IEC publication 402, Geneva, 1972.

[3] International Electrotechnical Commission *Precision method for free-field calibration of one-inch standard condenser microphones by the reciprocity technique*, IEC publication 486, Geneva, 1974.

[4] International Electrotechnical Commission *Values for the difference between free-field and pressure sensitivity levels for one-inch standard condenser microphones*, IEC Publication 655, Geneva, 1979.

Recommendation P.62

MEASUREMENTS ON SUBSCRIBERS' TELEPHONE EQUIPMENT

(amended at Malaga-Torremolinos, 1984)

1 Measurement of the attenuation distortion of a telephone set

The curve of the variation of the absolute sensitivity of an item of telephone equipment (sending or receiving system) as a function of frequency does not supply complete information on the manner in which this equipment reproduces the human voice or music, although such a curve may often be called the frequency characteristic.

However, the curve of variation of the absolute sensitivity of telephone equipment as a function of frequency gives useful indications from the point of view of the transmission of speech. On the other hand, for the transmission of music, in the absence of a precise criterion of the quality of transmission (corresponding to articulation, or repetition rate, in commercial telephony) such curves should be sufficient to enable the quality of the terminal equipment used (microphone or loudspeakers) to be appreciated.

For tracing sensitivity/frequency characteristics the methods described in Recommendation P.64 and its associated Annex B may be used.

2 Measurement of the nonlinear distortion of a telephone set and of microphone noise

While the nonlinear distortion of telephone receivers is in general negligible, microphones (and particularly carbon microphones of the type generally used in commercial telephone equipment) show considerable nonlinearity: the relationship between the variation of microphone resistance and the acoustic pressure on the diaphragm is not linear. This nonlinearity becomes more important as the variation of resistance in relation to the total resistance of the microphone increases, i.e. when the microphone is more sensitive. Furthermore, there may be two supplementary effects:

1) The microphone is less sensitive to acoustic pressure lower than a certain value (threshold of excitation).

2) As a consequence of the mechanical inertia of the carbon granules (delay in establishing electrical contact between the granules), the various states of agitation of the carbon under the influence of acoustic waves are not the same for all frequencies (for example, slow beats between two sounds are in general enhanced in reproduction by a carbon microphone).

Existing information on the general effect of harmonic distortion on telephone speech quality indicates that the effect of second order distortion is considerably less than that of third order distortion. Absolute detection thresholds obtained in different test are, however, difficult to compare because of differences in definition and measurement of the distortion.

Note 1 – Summaries of information available in this area are given in [1] and [2]. It is clear that measurements with sinusoidal signals can predict the speech transmission performance of nonlinear systems only to a limited extent, particularly if the peak value of the test signal is much smaller than the transmitted speech signal. A complex signal having the same spectral density at the same amplitude density function as real speech is therefore expected to be a more useful test signal.

Note 2 – The application of complex test signals or actual speech signals for the measurement of nonlinearity in telephone circuits is studied under Question 13/XII [3].

Certain types of carbon microphones may produce an audible stationary noise, often depending on the size of feeding current. The measurement of this kind of noise and its effect on transmission quality is the same as for other kinds of additive circuit noise.

3 Objective measurement of R25 equivalent, corrected reference equivalent (CRE) or loudness rating (LR)

At present no objective methods for measuring R25 equivalent or CRE have been recommended.

The apparatus described in [4] is employed by the French Administration for measuring "Planning equivalents". The "Planning equivalents" are numerically equal to the CRE, which could be computed for the same telephone sets, from subjective determination of their Reference equivalents. In addition this apparatus is capable of measuring STRE. The same measurements can also be made using the apparatus of NTT (see [5]).

Examples of apparatus that objectively measure LRs conforming to Recommendation P.79 are "CERF" of the French Administration [4], "AURAL" of NTT [5], "TIGGER" [6] of British Telecom and "Loudness Rating Meter" [7] of STL. Short descriptions of the apparatus named above can be found in Chapter 5 of the manual *Telephonometry* [8].

References

[1] CCITT – Question 13/XII, Annex 1, Contribution COM XII-No.1, Study Period 1981-1984, Geneva, 1981.

[2] CCITT – Question 13/XII, Annex, Green Book, Vol. V, ITU, Geneva, 1973.

[3] CCITT – Question 13/XII, Contribution COM XII-No. 1, Study Period 1985-1988, Geneva, 1985.

[4] CCITT – Contribution COM XII-No. 184, Equipment for the objective measurement of equivalent R25 and of the sidetone – used by the French Administration, France, Study Period 1981-1984.

[5] CCITT – Contribution COM XII-No. 79, Objective loudness rating measurement system, Nippon Telegraph and Telephone Public Corporation, Study Period 1981-1984.

[6] WARD (H. F.) and CROSS (R. C.) TIGGER: An Automatic Test System for measuring the Transmission Performance of Telephones, *British Telecommunications Engineering*, Volume 2, July 1983.

[7] CCITT – Question 15/XII, Annex 6, Contribution COM-No. 1, Study Period 1985-1988.

[8] CCITT manual *Telephonometry* (to be published in 1985).

Recommendation P.63

METHODS FOR THE EVALUATION OF TRANSMISSION QUALITY ON THE BASIS OF OBJECTIVE MEASUREMENTS

These measuring methods are being studied by the CCITT under Question 7/XII [1]. Annexes A and B to Recommendation P.11 and Supplements No. 2 and 3, at the end of this Fascicle, describe methods used respectively by British Telecom and AT&T. Attention is also drawn to methods for calculating loudness ratings given in Recommendation P.79.

Reference

[1] CCITT – Question 7/XII, Contribution COM XII-No. 1, Study Period 1985-1988, Geneva, 1985.

DETERMINATION OF SENSITIVITY/FREQUENCY CHARACTERISTICS OF LOCAL TELEPHONE SYSTEMS TO PERMIT CALCULATION OF THEIR LOUDNESS RATINGS

(Geneva, 1976; amended at Malaga-Torremolinos, 1984)

See Recommendation P.76 for general principles concerning the determination of loudness ratings.

1 Introduction

The sending, receiving or sidetone sensitivity/frequency characteristic of a local telephone system (LTS) is usually measured directly.

Note 1 – The sending, receiving or sidetone sensitivity/frequency characteristic can also be calculated provided the relevant information of the telephone line and feeding bridge is known. Some of the information required for sidetone is outside the scope of our existing Recommendations.

Note 2 – The same principles also apply to the measurement of microphones and earphones.

Since electro-acoustical measurements of the type being considered may be required for different purposes, it is important to distinguish the following:

a) supplying the designer of a transducer with information concerning the success he has achieved in aiming at a given sensitivity/frequency response;

b) checking that the manufactured product meets the specified requirements;

c) supplying sensitivity/frequency characteristics suitable for use in estimating loudness ratings, reference equivalents or other subjectivity-determined quantities.

The present Recommendation is concerned only with c) and, for this purpose, measurements under real conditions must form the basis. Artificial mouths and artificial ears must be used with due regard to obtaining good agreement between these measurements and those from real mouth and ear determinations. Measurements under real conditions are complicated, time-consuming and not reproducible with great precision, especially when carbon microphones are involved.

The present Recommendation describes measurement methods using recommended forms of artificial mouths and artificial ears (see Recommendation P.51).

2 Sending sensitivities of LTS

For the present purposes, the sending sensitivity of a local telephone system is specified in terms of the free-field sound pressure at a reference point in front of the mouth[1], and the electrical output from the local telephone system or the microphone as the case may be. The input sound pressure cannot be measured simultaneously with the electrical output and therefore the measurement must be made in an indirect manner. The sound pressure at the reference point is measured in the absence of the handset and, with the artificial mouth source unchanged, the handset is placed in the defined position in front of the mouth and the output measured. When a human mouth and voice are used, the source cannot be relied upon to maintain its output constant between the measurement of free-field sound pressure and that of the electrical output from the microphone. Artificial mouths suffer from imperfect representation of the source impedance and field distribution that applies to real mouths.

[1] The mouth reference point used in the present Recommendation is defined in Annex A.

In addition to providing the proper source conditions, it is necessary to ensure that the mouthpiece is located for every design of telephone handset at the position that would be used in the real situation. This can be achieved by locating the mouthpiece properly with respect to an ear reference point; this ensures that longer handsets are measured with a greater mouth-to-microphone distance than is the case for shorter handsets. The success of using a given handset measuring position for measurement of sensitivity/frequency characteristics can be judged only by making comparisons, for handsets of different lengths, between real conversation test results using the artificial mouth and real mouths under suitably controlled measuring conditions. For the present Recommendation, the telephone handset shall be located as defined in Annex A of Recommendation P.76.

Special problems are encountered when making measurements with real mouths and real voices, even under controlled talking conditions. Under such circumstances the sound pressure cannot be measured directly at the required mouth reference point and therefore it has to be measured at some other point and referred indirectly to the mouth reference point. Some previous determinations have made use of a measuring microphone 1 metre from the mouth but this requires anechoic surroundings and is affected by obstruction from the handset under test.

When the sound pressure input to a carbon microphone is increased, the corresponding increase in output voltage does not bear a linear relationship to the increase in sound pressure. This nonlinearity is a very complicated function of applied sound pressure, frequency, feeding current, conditioning and granule-chamber orientation. Reproducible results are obtained with an artificial mouth only if proper attention is paid to all these factors.

3 Receiving sensitivities of the LTS

The IEC-318 model artificial ear (see Recommendation P.51) provides means for precise measurements of the receiving sensitivities of the LTS. However, the sound pressures measured with it do not always agree well with those existing at the ear reference point in real ears under the test conditions used when subjective determinations of loudness ratings are being made. This can be attributed partly to the presence of appreciable acoustical leakage (L_E) between the earphone and the real ear (such leakage is not represented in available recommended forms of the artificial ear) and partly to an increase in enclosed volume between the forms of earphone and the forms of real ear. Therefore, to use the results of measurements made according to the present Recommendation, it is necessary to make a correction (see § 7 below).

Clearly, it would be very desirable if the artificial ear could be modified so as to avoid the need for the correction. Some further work has been done on this matter but it is not yet clear whether a single modification to the artificial ear would suffice for all types of telephone earphone. Further evidence is required, preferably from several laboratories so that a much wider variety of types of earphone can be examined.

4 Artificial mouth and voice

The following properties are required:

a) the distribution in sound pressure around the orifice must be a good approximation to that around a human mouth;

b) the acoustical impedance looking into the mouth must simulate that for human mouths, so that the pressure increase caused by the obstruction effect of telephone microphones will be representative;

c) it must be possible to establish definite sound pressures at the mouth reference point as a function of frequency. A convenient feature to embody in a practical artificial mouth is the linearity, over a suitable range of sound pressures, of the ratio of sound pressure at the mouth reference point to the voltage input to the artificial mouth. The ratio must be independent of frequency at least over the range 200 to 4000 Hz but preferably 100 to 8000 Hz.

For the present purposes the mouth reference point (MRP) is defined by the point on the axis of the artificial mouth located 25 mm in front of the equivalent lip position (see Annex A).

Recommendation P.51 defines the requirements for artificial mouths and for the artificial voice suitable for the present purposes.

Note – However, the send loudness ratings calculated from the sending sensitivities measured when using such a mouth do not always agree well with the loudness ratings determined subjectively using real mouths. This can in part be attributed to the fact that items a) and b) above have yet been sufficiently well defined. The subject is still under study in Question 12/XII.

5 Artificial ear

The following properties are required:

a) the acoustical impedance presented to telephone earphones must simulate that presented by real ears under practical conditions of use of telephone handsets;

b) the sensitivity of the artificial ear is defined as the pressure sensitivity of the measuring microphone. It should be constant within ± 0.5 dB over the frequency range 100-8000 Hz.

For a human ear, the ear reference point (ERP) is defined in Annex A. The corresponding point when the ear-cap is fitted to an artificial ear will usually differ from the place at which the sound pressure is measured and for this and other reasons certain corrections are necessary when the results are used for calculating loudness ratings (see § 3 above).

6 Definition of sending sensitivity of a local telephone system (LTS)

The sending sensitivity of an LTS, depends upon the location of the handset relative to the equivalent lip position of the artificial mouth. For the present purposes the speaking position defined in Annex A to Recommendation P.76 shall be used.

The sending sensitivity of a local telephone circuit is expressed as follows:

$$S_{mJ} = 20 \log_{10} \frac{V_J}{p_m} \text{ dB rel 1 V/Pa}$$

where V_J is the voltage across a 600 ohms termination and p_m is the sound pressure at the mouth reference point. Note that p_m must be measured in the absence of the "unknown" microphone of the test item.

6.1 *Measurement of complete telephone sets containing carbon*

It is intended that the Recommendation should apply for measuring systems containing carbon microphones as well as those having noncarbon microphones. When measuring local telephone systems (LTSs) that contain linear items, it does not matter at which sound pressure the measurements are made as long as it is known and does not cause overloading. However, when carbon microphones are present, different sensitivities will be obtained depending upon the sound pressure and characteristics of the acoustic signal used. For calculation of sending loudness rating, these must be reduced to single values at each frequency and the method of reduction must take account of the characteristics of human speech. At present, there is no single method that can be recommended for universal use. The problem is being studied under Question 8/XII [1]. Until a suitable method can be defined, Administrations may take note of the various methods that have been suggested and are undergoing appraisal; they are indicated in Annex B.

7 Definition of receiving sensitivity of a local telephone system (LTS)

The receiving sensitivity of an LTS, as measured directly with an artificial ear complying with Recommendation P.51, is expressed as follows:

$$S_{Je} = 20 \log_{10} \frac{p_e}{\frac{1}{2} E_J} \text{ dB rel 1 Pa/V}$$

where p_e is the sound pressure in the artificial ear and $\frac{1}{2} E_J$ is half the emf in the 600 ohm source.

Note – The receiving sensitivity suitable for use in calculation of loudness is given by:

$$S_{Je} = S_{JE} - L_E$$

where L_E is a correction explained above in § 3 and S_{JE} is the receiving sensitivity determined using a large number of real ears.

Further information on this topic is given in Recommendation P.79.

8 Methods of sidetone sensitivity of a local telephone system (LTS)

The sidetone sensitivity of an LTS is a function of the sending and receiving sensitivities of the telephone set, but also depends on a number of factors including the local subscriber's line conditions, the effective terminating impedance at the local exchange and the sidetone balance circuit within the telephone set.

The sidetone sensitivity is expressed as:

$$S_{meST} = 20 \log_{10} \frac{p_e}{p_m} \text{ dB}$$

where p_e is the sound pressure developed in the artificial ear for a free-field sound pressure, p_m is the sound pressure at the MRP of an artificial mouth.

Note 1 – For calculation of STMR in the unsealed condition the correction factor, L_E must be taken into account (see Recommendation P.79).

Note 2 – The sidetone sensitivity determined with an artificial mouth and an artificial ear will be subject to the same considerations as described in §§ 2 and 3.

9 Methods for determining S_{mJ}, S_{Je} and S_{meST}

When the sending, receiving and sidetone sensitivities of an actual local telephone system are required, the measurements according to the definitions given in §§ 6, 7 and 8 above can be made as illustrated in Figures 1/P.64, 2/P.64, 3/P.64 and 4/P.64. These methods have been used by CCITT Laboratory and elsewhere successfully.

More detail may be found in Chapter 3 of the manual *Telephonometry* [2].

Figure 1/P.64 shows the method of setting up the artificial mouth so that the sound pressure p_m at the mouth reference point is known at each test frequency or frequency band. It is recommended to provide equalization in the artificial mouth drive circuit to maintain the free-field sound pressure constant at the MRP to within ± 1 dB over the frequency range 100 to 8000 Hz. In no case should the deviation exceed ± 2 dB over the frequency range 200 to 4000 Hz and +2/−5 dB over the frequency range 100 to 8000 Hz. It is recommended that any deviations from the desired sound pressure level be taken into account when determining the sending or sidetone sensitivity of a local telephone system. This is particularly true if the deviation exceeds ± 1 dB.

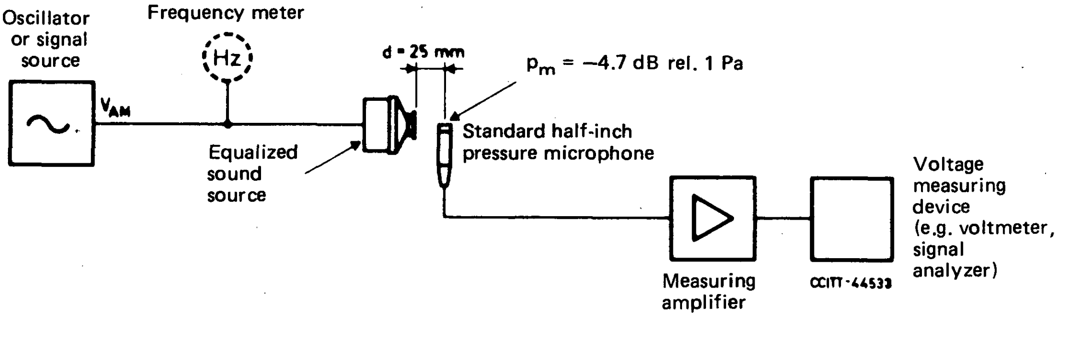


FIGURE 1/P.64

Measurement of acoustic pressure p_m at the mouth point 25 mm from the artificial lip plane of the sound source

Figure 2/P.64 shows the measurement of output V_J from the local telephone circuit when the microphone is placed at the appropriate position in front of the artificial mouth and the artificial mouth is energized in the same manner as when the sound pressure p_m was set up in the absence of the test microphone (Figure 1/P.64).

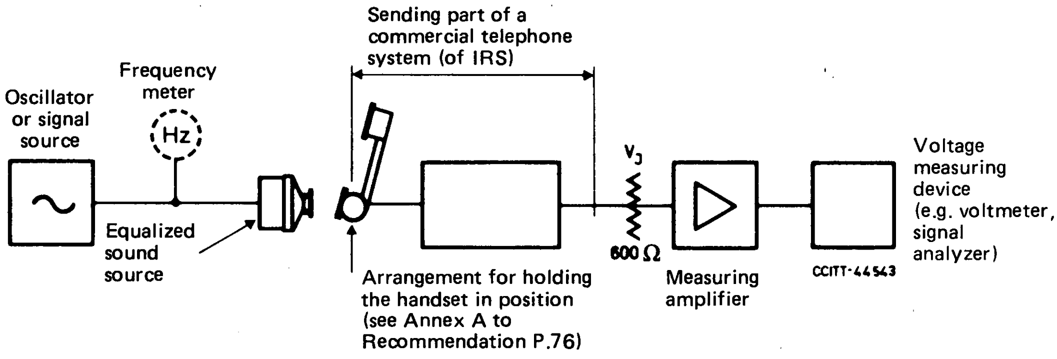


FIGURE 2/P.64

Voltage V_j, measured across the terminals of a 600 ohms pure resistance connected to the output of a commercial sending system

Figure 3/P.64 shows the measurement of the sound pressure p_e in the artificial ear when the local telephone system is connected to a 600-ohm source of internal emf E_J. Note that the definition of S_{Je} is in terms of $1/2\ E_J$ and not the potential difference across the input terminals of the local telephone system; this potential difference will, of course, differ from $1/2\ E_J$, if the input impedance of the local telephone system is not 600 ohms.

Note – Some receiving systems incorporate electronic circuits to provide special features, for example, compression to limit the level of the received sound signal. Particular care must be exercised during the measurement of such systems to ensure that the resulting sensitivity is correct and relevant. In some cases it may be necessary to determine the receiving sensitivity over a range of input levels.

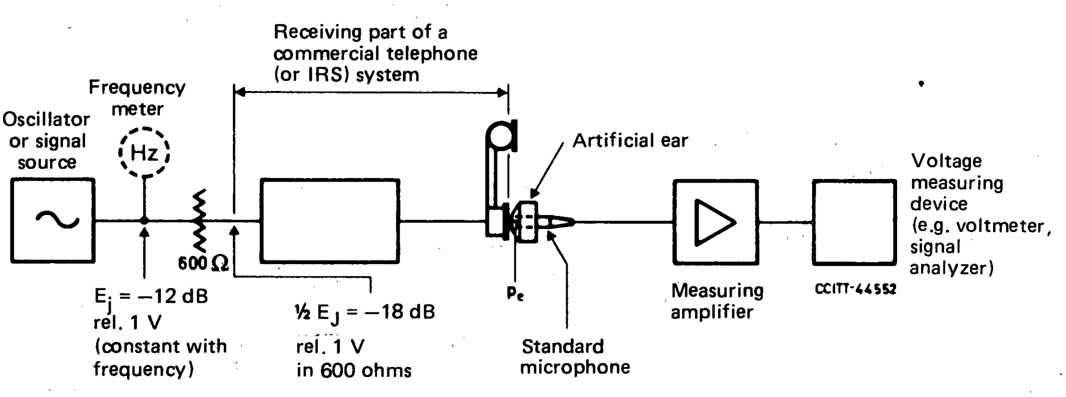


FIGURE 3/P.64

Calibration of a commercial telephone receiving system or the receiving part of the intermediate reference system

Figure 4/P.64 shows the measurement of sidetone sensitivity. The resulting value of S_{meST} is highly dependent on the impedance connected to the telephone set terminals and therefore, under short line conditions, on the exchange termination. As this impedance often deviates considerably from 600 ohms, particularly when there is a complete connection present, 600 ohms is given only as an example.

Note – The method of Figures 1/P.64 and 4/P.64 may be used to determine room noise sidetone sensitivity if the artificial mouth is replaced by a distant sound source, or by multiple sound sources to give a random sound field at the MRP. For this measurement, sinewave signals are unsuitable and it is necessary to make use of a continuous spectrum sound having, for example, a Hoth or pink noise spectrum (see Annex B, § B.3).

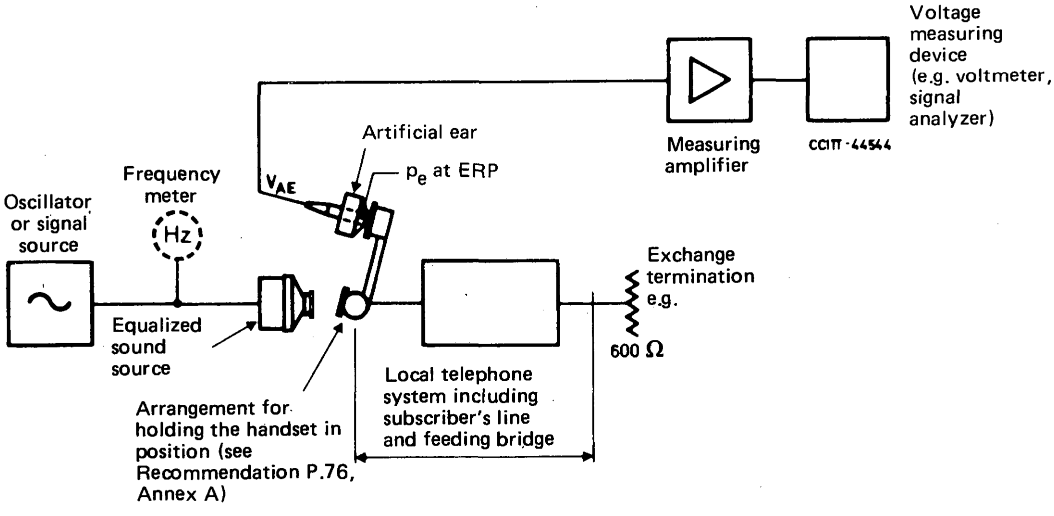


FIGURE 4/P.64

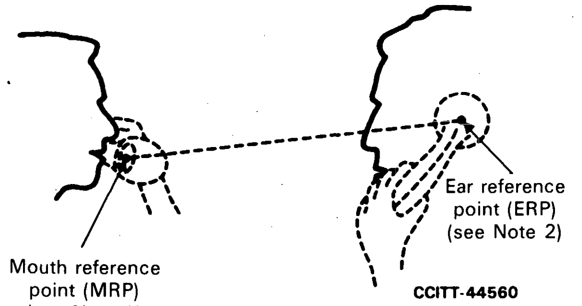
Measurement of the sidetone sensitivity of a commercial telephone system by determination of the sound pressure p_e developed in an artificial ear for a given sound signal at the MRP

ANNEX A

(to Recommendation P.64)

Definitions of mouth reference point and ear reference point

The definitions of mouth reference point (MRP) and ear reference point (ERP) are illustrated in Figure A-1/P.64.



Note 1 – The mouth reference point is located at a distance of 25 mm in front of the lips on the horizontal axis through the centre of the opening of the mouth. It is defined in the absence of any obstruction.

Note 2 — The ear reference point is located at the entrance to the ear canal of the listener's ear. It is defined as lying at the centre of the front plane of a circular concave earcap when sealed to the ear.

Note 3 — The ERP should not be confused with the earcap reference point (ECRP) which is a point in the earphone reference plane used as a handset reference parameter (see Recommendation P.10).

FIGURE A-1/P.64

Definitions of mouth and ear reference points

ANNEX B

(to Recommendation P.64)

Measurement of complete telephone sets containing carbon microphones

For the measurement of complete telephone sets containing carbon microphones, various methods have been suggested and tried. The following gives, as examples, some of these methods. These same methods can also apply to telephones using linear microphones.

B.1 The *upper envelope method* has been used in the CCITT Laboratory with success for some types of carbon microphone but has been less successful with others. The upper envelope method is as follows:

a) Determine the sensitivity as a function of frequency at the sound pressure level of −4.7 dB relative to 1 Pa. This is somewhat higher than the mean power of active speech of a talker, emitting speech at the vocal level used to determine reference equivalents in accordance with Recommendation P.72 and loudness ratings in accordance with the subjective test method described in Recommendation P.78;

b) Repeat a) but with the sound pressure level increased by 10 dB;

c) Repeat a) but with the sound pressure level decreased by 10 dB;

d) Select from a), b) and c) the highest sensitivity at each frequency.

Carbon microphones must be given appropriate conditioning treatment at suitable intervals during the measurements (see Recommendation P.75).

B.2 *Sweeping frequency method*

Currently available types of objective instrumentation for measuring loudness ratings use a sweeping frequency covering the range from 200-4000-200 Hz at a periodicity of 1 sweep per second; the instantaneous level within any narrow frequency band varies as a function of frequency approximately in accordance with the spectrum of speech emitted from the human mouth.

B.3 *Pink-noise method*

The handset containing the carbon microphone is placed in front of an artificial mouth producing at the MRP pink noise (power spectrum density diminishing by 3 dB/octave) over 1/3rd octave frequency bands centred on the preferred frequencies specified in ISO Standard 266-1975 at 1/3rd octave intervals in the range 100 to 8000 Hz with the band edges conforming to the filters described in IEC 225.

The total level of the signal, measured over the same bandwidth, should be −4.7 dB Pa with a tolerance of ± 1.0 dB.

Note – This may not be practical with the current artificial mouth, and a narrower bandwidth of 200 to 8000 Hz may have to be used.

The sensitivity/frequency characteristic is obtained by finding the ratio of the power spectrum density of the signal delivered by the carbon microphone to the power spectrum density of the signal obtained using a small linear microphone placed at the MRP under free-field conditions (after removing the carbon microphone).

B.4 *Shaped gaussian noise method*

The method uses shaped gaussian noise at the MRP, whose long-term average spectrum density is the same as shown in Table 2/P.51. The total level of the signal should be −4.7 dB Pa ± 1 dB.

The sensitivity frequency characteristic is obtained as in § B.3.

B.5 *Real-voice calibration*

This may be performed by measuring speech spectra emitted alternately or simultaneously from the carbon microphone under test and a calibrated linear microphone. A very small linear microphone can be mounted on the telephone being tested. Naturally the most appropriate results will be obtained when the talkers are conducting telephone conversations, but it is then difficult to have reliable knowledge of the sensitivity/frequency characteristic of the linear microphone. It is usually necessary to rely upon a suitable artificial mouth to provide the calibration of the linear microphone.

B.6 *Application of a wideband signal*

The wideband signal is generated by a pseudo-random binary sequence and is then equalized to have a long-term average spectrum density as defined in Recommendation P.51. The output from the carbon microphone is then processed by a digital computer using the Fourier transform. This method, like the previous method, requires calibration by a linear microphone of known sensitivity/frequency characteristic. This method has the advantage that the frequency characteristic may be obtained with a very short duration (e.g. 50 ms) sample of test signal.

B.7 *Method using the artificial voice*

The method uses the artificial voice, having spectral and time characteristics similar to those of speech.

The sensitivity/frequency characteristic is obtained as in § B.3 above, but with the artificial mouth supplying the acoustic artificial voice, defined in Recommendation P.51, § 3.3.3.

Note to Annex B – The efficiency of the artificial mouth used is not generally constant with frequency, so it is necessary, for most of the methods described above, to insert appropriate equalization networks between the electrical signal generator and the loudspeaker of the artificial mouth. It is the free field acoustical signal which shall conform to the complex signal or the artificial voice specified.

References

[1] CCITT – Question 8/XII, Contribution COM XII-No. 1, Study Period 1985-1988, Geneva, 1985.

[2] CCITT manual *Telephonometry* (to be published in 1985).

Recommendation P.65

OBJECTIVE INSTRUMENTATION FOR THE DETERMINATION OF LOUDNESS RATINGS

(Malaga-Torremolinos, 1984)

1 Introduction

This Recommendation describes the essential features of objective instrumentation suitable for the determination of loudness ratings. These features are drawn from current Recommendations relating to loudness ratings, the principles of which are defined in Recommendation P.76.

It is possible to realize objective instrumentation for loudness rating purposes in a number of ways, for example by the assembly of a number of separate instruments, each having its own defined function, and possibly under some central control, or by means of a dedicated piece of apparatus specially designed for the purpose. However, in order to ensure that loudness rating measurements made in different laboratories have an acceptable level of agreement, say ± 1 dB, it is essential that the Recommendations relating to the measurement of the electro-acoustic performance of telephone systems should be followed.

The relevant Recommendations are:

P.48 Specification for an intermediate reference system

P.51 Artificial voices, artificial mouths, artificial ears

P.64 Determination of sensitivity/frequency characteristics of local telephone systems to permit calculation of their loundness ratings

P.75 Standard conditioning method for handsets with carbon microphones

P.76 Determination of loudness ratings; fundamental principles

P.79 Calculation of loudness ratings[1].

2 Instrumentation

Below are described the four electro-acoustic sections that are required to be included in equipment intended for use in determining loudness ratings. In each case appropriate calibration is required as a function of frequency, and calibration values recorded in the fifth section where the particular sensitivity/frequency characteristic is derived and the loudness rating calculated.

It is necessary to provide certain auxiliary apparatus, such as feeding circuits, artificial subscriber cable and exchange terminations, as required by the particular Recommendation(s) being followed for any given measurement.

2.1 *Artificial ear*

See *a)* of Figure 1/P.65.

The artificial ear in the system should be in accordance with Recommendation P.51 and contain within it a measuring amplifier so that the pressure p_e occurring at the artificial ear cavity can be measured as a function of frequency, or in frequency bands within the recording and measurement system, *e)* of Figure 1/P.65. Means must also be available to calibrate the standard microphone used in the artificial ear employing, for example, an acoustic calibrator or piston-phone.

2.2 *Artificial voice*[2]

See *b)* of Figure 1/P.65.

An artificial mouth complying with Recommendation P.51 must be part of the system and be able to produce a prescribed sound field at the MRP 25 mm in front of the lip plane. A signal source will be part of the artificial voice and this source may be sine waves (swept or discrete frequencies) or a wideband signal (for example shaped gaussian noise as defined in Recommendation P.64, § B.4). Equalization and gain control should be part of the drive system to the artificial mouth such that the sound pressure at the MRP can be controlled in accordance with the requirements of Recommendation P.64, §§ B.1 and B.4.

Calibration of the sound pressure and/or spectrum at the MRP may be carried out using the standard microphone used in the artificial ear of § 2.1 above, making use of the recording and measurement system of § 2.5 below to determine p_m as a function of frequency, or in frequency bands.

Mechanical means must be provided to hold the test handset in the LRGP (loudness rating guard-ring position), in accordance with the requirements of Recommendation P.76, Annex A. If handsets having carbon microphone are being tested, conditioning in accordance with Recommendation P.75 must be provided.

2.3 *Electrical termination*

See *c)* of Figure 1/P.65.

The system should contain a 600 ohm balanced terminating impedance with means for measuring the terminating voltage, V_J (see Recommendation P.64, §§ 6 and 9, as a function of frequency, or in frequency bands, using the recording and measurement system of § 2.5 below. Calibration of this section may be carried out using a calibrated voltage source.

[1] Other algorithms are being studied by Study Group XII, under Question 15/XII.

[2] The artificial voice in this context does not refer to that recommended in Recommendation P.51.

2.4 *Electrical signal source*

See *d)* of Figure 1/P.65.

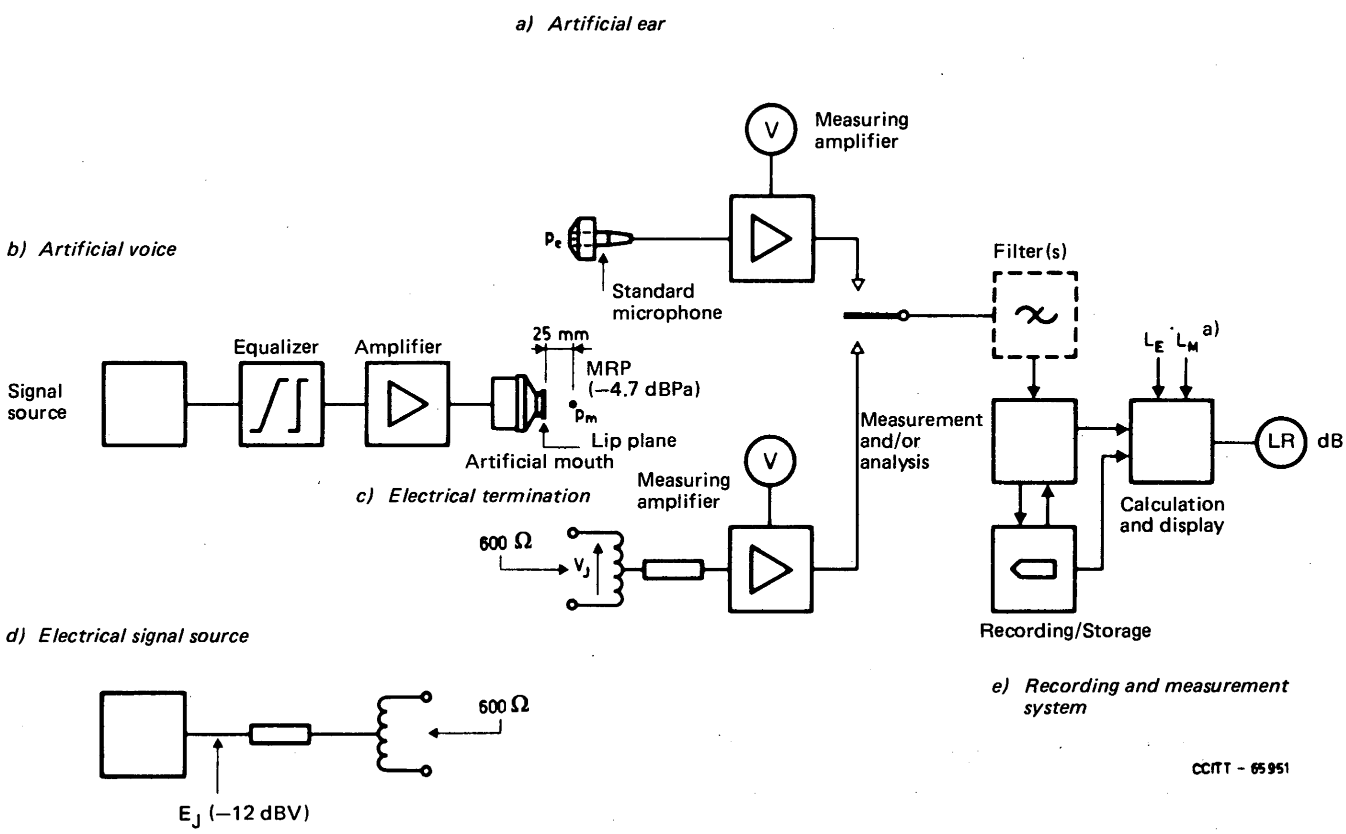
An electrical signal source must be provided having a 600 ohm balanced impedance. The electrical source need not be the same as that used for the artificial voice but should either be sine waves or a wideband signal. There should be means for calibrating and adjusting the generator voltage, E_J, to the requirements of Recommendation P.64, §§ 7 and 9 over the frequency range 100-8000 Hz. This may be carried out using calibration of the electrical termination of § 2.3 above.

2.5 *Recording and measurement system*

See *e)* of Figure 1/P.65.

In order to determine the sound pressure p_e at the artificial ear or the voltage V_J at the electrical termination it will be necessary to provide a recording and measurement system. This measurement system may, using hardware of software, contain filters in order to improve signal-to-noise ratio or for analysing the output of the telephone set in 1/3rd octave frequency bands. Where a bank of 1/3rd octave filters is used these should be centred on the preferred frequencies in accordance with ISO 266 and have the characteristics in accordance with IEC Publication 225.

Within this part of the system there should be recording or storage facilities so that calibration and measurement data may be used to derive the necessary sensitivity/frequency characteristics in accordance with Recommendation P.64. The various loudness ratings are then computed in accordance with Recommendation P.79 from the sensitivity/frequency characteristics, taking into account any recognized adjustments, for example L_E or L_M. Values for L_E and L_M may be fed into the calculation using default values (e.g. those listed for L_E in Table 4/P.79) or from other more appropriate data sources when available.



a) The artificial real mouth corrections L_M has been assumed to be zero but is currently under study in Question 8/XII.

FIGURE 1/P.65

Essential features of instrumentation for the determination of loudness ratings

3 Measurements

Facilities should be provided to enable the various sections of the instrumentation to be connected to allow the measurement of the necessary sensitivity/frequency characteristics and calculation of the loudness ratings.

A summary of these interconnections, together with the sensitivity/frequency characteristics (SFC) measured for particular loudness rating determinations, are given below.

3.1 *Send loudness rating (SLR)*

Source: *b)* of Figure 1/P.65

Load: *c)* of Figure 1/P.65

Send SFC given by:

$$S_{mJ} = 20 \log_{10} \frac{V_J}{p_m} \text{ dB}$$

3.2 *Receive loudness rating (RLR)*

Source: *d)* of Figure 1/P.65

Load: *a)* of Figure 1/P.65

Receive SFC given by:

$$S_{Je} = 20 \log_{10} \frac{p_e}{\frac{1}{2} E_J} \text{ dB}$$

3.3 *Sidetone masking rating (STMR)*

Source: *b)* of Figure 1/P.65

Load: *a)* of Figure 1/P.65

Sidetone SFC is given by:

$$S_{meST} = 20 \log_{10} \frac{p_e}{p_m} \text{ dB}$$

Note – The quantity L_{meST} used in the calculation of STMR is given by:

$$L_{meST} = -S_{meST} \text{ dB}$$

3.4 *Overall loudness rating (OLR) (Overall Send + Receive, (OSR))*

Source: *b)* of Figure 1/P.65

Load: *a)* of Figure 1/P.65

Overall SFC given by:

$$S_{me} = 20 \log_{10} \frac{p_e}{p_m} \text{ dB}$$

3.5 *JLR Junction loudness rating*

Source: *d)* of Figure 1/P.65

Load: *c)* of Figure 1/P.65

Junction loss/frequency characteristics given by:

$$X_{JJ} = 20 \log_{10} \frac{\frac{1}{2} E_J}{V_J} \text{ dB}$$

Note – Impedance terminations of 600 ohms are assumed.

SECTION 6

SUBJECTIVE VOICE-EAR MEASUREMENTS

Recommendation P.70

MODULATED NOISE REFERENCE UNIT (MNRU)

(Malaga-Torremolinos, 1984)

The CCITT,

considering

(a) that the use of digital processes (8-bit A-law or μ-law, A/D/A encoder decoder pairs, A/μ-law or μ/A-law converters, digital pads based on 8-bit PCM words) in the international telephone network has grown rapidly over the past several years, and this growth is expected to continue;

(b) that there is a need for standard tools to measure the quantization distortion performance of digital processes (for example, 32 kbit/s ADPCM) so that the tools can be used for estimating the subjective transmission performance of international connections containing digital processes;

(c) that at the present time subjective tests incorporating reference system conditions represent the only suitable method for measuring the speech transmission performance of digital processes;

(d) that expressing results in terms of a common reference system may facilitate comparison of subjective test results obtained at different laboratories,

recommends provisionally

The use of Modulated Noise Reference Unit (MNRU) of this Recommendation as the reference system in terms of which subjective performance of digital processes should be expressed.

Note 1 – The MNRU can be realized using laboratory equipment or by computer simulation. Further information on the MNRU is given in the references listed at the end of this Recommendation.

Note 2 – The listening-only method presently proposed for using the MNRU in subjective tests is described in Supplement No. 14 at the end of this volume. See Recommendation P.74, § 3, for precautions concerning the use of listening-only tests.

Note 3 – Objective measurement methods which suitably reflect subjective quantization distortion performance of various types of digital processes do not exist at present. (For example, the objective technique of Recommendation G.712, based on sine-wave and band-limited noise measurements, are designed for PCM and do not measure appropriately the distortion introduced by other systems such as ADPCM.) The work on artificial speech signal under Question 12/XII may be relevant. Even if an objective method is developed, subjective tests will be required to establish correlation of subjective results/objective results for particular digital process types.

1 Introduction

The MNRU was originally devised to produce distortion subjectively similar to that produced by logarithmically companded PCM systems [1]. This approach was based on the views:

1) that network planning would require extensive subjective tests to enable evaluation of PCM system performance over a range of compandor characteristics, at various signal levels and in combination with various other transmission impairments (e.g., loss, idle circuit noise, etc.) at various levels, and

2) that it would be as reliable and easier to define a reference distortion system, itself providing distortion perceptually similar to that of PCM systems, in terms of which the performance of PCM systems could be expressed. This requires extensive subjective evaluation of the reference system when inserted in one or more simulated telephone connections, but leads to the possibility of simplified subjective evaluation of new digital processing techniques.

The MNRU concept has been used extensively by various organizations (Administrations, recognized private operating agencies and scientific/industrial organizations) in evaluating the subjective performance of digital processes. However, the actual devices used, while based on common principles, may have differed in detail and, thus, subjective test results obtained may also have differed. (Differences in subjective testing methodology are also relevant.) The purpose of this Recommendation is to define the MNRU as completely and in as much detail as possible in order to minimize the effects of the device and objective calibration procedures on subjective test results.

2 General description

A simplified arrangement of the MNRU is shown in Figure 1/P.70. Speech signals entering on the left are split between 2 paths, a signal path and a noise path. The signal provides an undistorted (except for bandpass filtering) speech signal at the output. In the noise path, the speech signal instantaneously controls a multiplier with an applied gaussian noise "carrier" which has a uniform spectrum between 0 Hz and a frequency at least twice the cutoff frequency of the lowpass portion of the bandpass filter. The output of the multiplier, consisting of the noise modulated by the speech signal, is then added to the speech signal to produce the distorted signal.

The attenuators and switches in the signal and noise paths allow independent adjustment of the speech and noise signal levels at the output. Typically, the system is so calibrated that the setting of the attenuator (in dB) in the noise path represents the ratio of speech power (in milliwatts, while active) and noise power (in milliwatts). For this Recommendation, the decibel representation of the ratio is called Q.

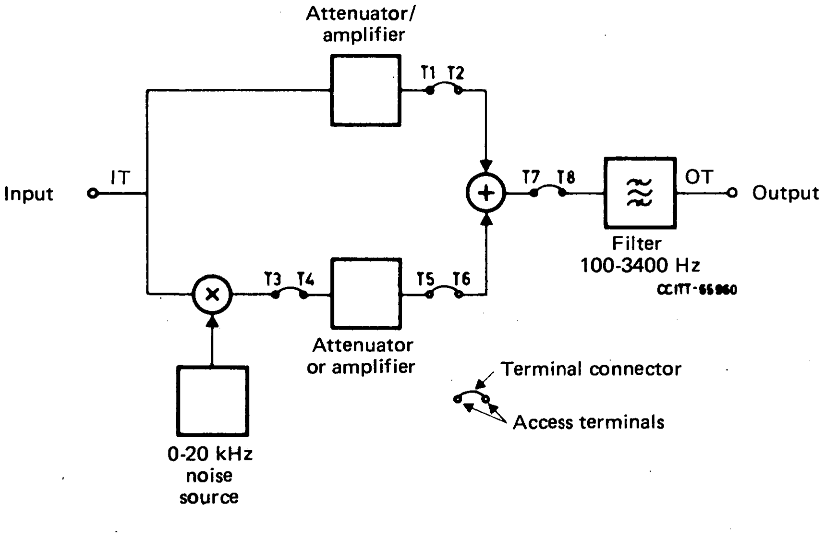


FIGURE 1/P.70

Basic arrangement of MNRU

3 Performance specifications

3.1 *General*

The specifications in this section apply both to hardware implementations and software simulations.

For practical implementations, the actual signal levels and noise levels may be increased or decreased to meet special needs. In such cases, the level requirements detailed below will have to be modified accordingly.

3.2 *Signal path*

The requirements under this heading refer to the MNRU with the noise path replaced by a resistive termination at terminal T6 in Figure 1/P.70.

The frequency range of the signal path and of the noise path is determined by the bandpass filter in their common output circuit. The frequency response of the signal path (i.e., between terminals IT and OT in Figure 1/P.70) should be within the limits of Figure 2/P.70.

The loss between terminals IT and OT for a 0 dBm, 1 kHz input sinewave should be 0 dB. Over the input level range +10 dBm to −50 dBm, the loss should be 0 dB ± 0.1 dB.

Any harmonic component should be at least 50 dB below the fundamental at the system output (terminal OT in Figure 1/P.70) for any fundamental frequency between 125 Hz and 3000 Hz.

The idle noise generated in the signal path should be less than −60 dBm measured at the system output. For this measurement, the signal source should be replaced by an equivalent resistive termination.

It is recommended that the level of speech signals applied to the terminals IT should be less than −10 dBm (mean power while active) to avoid amplifier peak clippings of the signal and be greater than −30 dBm to ensure sufficient speech signal-to-noise ratio.

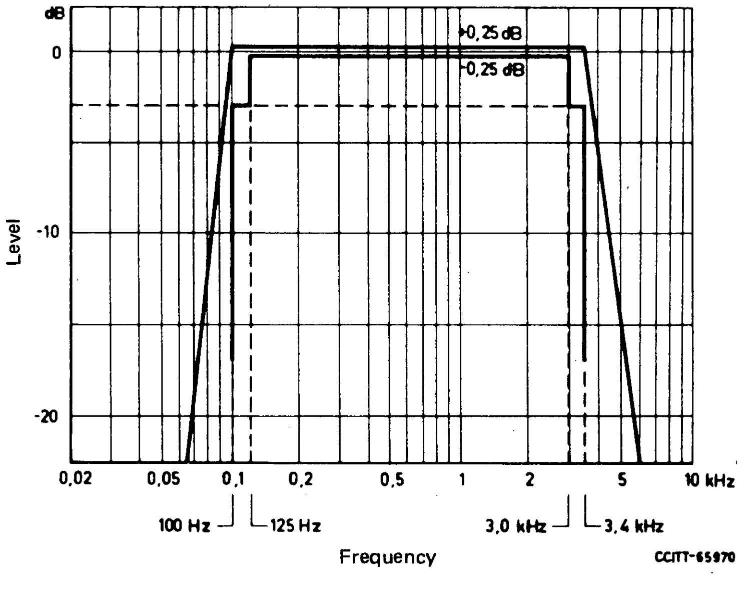


FIGURE 2/P.70

Requirements for output filter of the MNRU

3.3 *Noise path*

The requirements under this heading refer to the MNRU with the signal path replaced by an equivalent resistive termination at terminal T2 in Figure 1/P.70.

3.3.1 *Linearity as a function of input level*

With a Q-setting equal to 0 dB, the noise level at the system output (terminal OT of Figure 1/P.70) should be numerically equal to the sinewave level at the input (terminal IT). A correspondence within ± 0.5 dB should be obtained for input levels from +5 dBm to −45 dBm and input frequencies from 125 Hz to 3000 Hz.

3.3.2 *Noise spectrum*

For a Q-setting of 0 dB, input sinewaves applied to terminals IT in Figure 1/P.70 with levels from +5 to −45 dBm and frequencies from 125 Hz to 3000 Hz should result in a flat noise spectrum density at the output of the multiplication device (terminal T3 of Figure 1/P.70) within ± 1 dB over the frequency range 75 Hz to 5000 Hz. The spectrum density should be measured with a bandwidth resolution of maximum 50 Hz.

3.3.3 *Amplitude distribution*

The amplitude distribution of the noise at the system output should be approximately gaussian.

Note – A noise source consisting of a gaussian noise generator followed by a peak clipper with a flat spectrum from near zero to 20 kHz will produce a satisfactory output noise at terminal OT.

3.3.4 *Noise attenuators*

The loss of the noise attenuator(s) i.e., between terminals T4 and T5 in Figure 1/P.70, should be within ± 0.1 dB of the nominal setting. The attenuators should at least allow Q settings in the range −5 dB to 45 dB, i.e. a 50 dB range.

References

[1] LAW (H. B.), SEYMOUR (R. A.): A reference distortion system using modulated noise, *The Institute of Electrical Engineers*, November 1962, pp. 484-485.

[2] CCITT – Contribution COM XII-No. 63, *Some considerations on specifications for modulated noise reference unit*, NTT, Japan, Study Period 1981-1984.

[3] CCITT – Contribution COM XII-No. R4, pp. 71-79, Study Period 1981-1984.

[4] CCITT – Contribution COM XII-No. 119, *Description and method of use of the modulated noise reference unit (MNRU/MALT)*, France, Study Period 1981-1984.

Recommendation P.71

MEASUREMENT OF SPEECH VOLUME

(amended at Mar del Plata, 1968)

Each volume meter should be used in accordance with the relevant specifications (see Recommendation P.52). When the normal speech power for voice-ear measurements is to be used, the information provided in Recommendations P.42, § 1.2 and P.72, § 4.2, should be borne in mind.

Recommendation P.72

MEASUREMENT OF REFERENCE EQUIVALENTS AND RELATIVE EQUIVALENTS

(amended at Malaga-Torremolinos, 1984)

Preface

This Recommendation describes the methods used to obtain a subjective evaluation of reference equivalent (RE) and R25 equivalent (R25E) making use of the master reference system for the determination of reference equivalents, NOSFER. In certain cases, NOSFER may be replaced by a working standard system (e.g. SETED, see Recommendation P.42) provided it has been calibrated directly against NOSFER: resulting quantities then determined are known as relative equivalents.

1 Introduction and background

Recommendation P.76, § 1 describes how the speech path between two subscribers may be broken down into send and receive ends of the telephone connection and the junction circuit forming a transmission path between them. Figure 1/P.72, in path 1, shows an unknown conplete connection comprising a local telephone system (LTS) at the send and receive ends, and a chain of circuits interconnecting them.

The reference equivalent method is a means of assigning a loudness related rating to each of the three main parts of a telephone connection. For this purpose, a reference connection has been established, NOSFER (see Recommendation P.42), and the evaluation of a reference equivalent consists of a comparison using the human voice and ear of the NOSFER and the connection under test. The complete connection is not usually available for this purpose and therefore representative parts of the connection are evaluated separately under laboratory conditions.

Thus, to determine the send reference equivalent (SRE), path 2 is compared with path 3 and x_2 and/or x_3 adjusted until the two paths sound equally loud. Then

$$\mathrm{SRE} = x_2 - x_3 \text{ dB}.$$

Similarly, receive reference equivalent (RRE) is obtained by the comparison of paths 2 and 4, and

$$\mathrm{RRE} = x_2 - x_4 \text{ dB}.$$

A rating for representation of the chain of circuits forming the junction between JS and JR in Figure 1/P.72 may be made by comparing first path 5 with path 2, and then path 6 with path 2. The rating for the circuit is then given by $x_6 - x_5$.

The units of reference equivalent are analogous to loss and therefore the higher the value assigned to a test system, the quieter it is. Negative values of RE indicate that a system is louder than NOSFER.

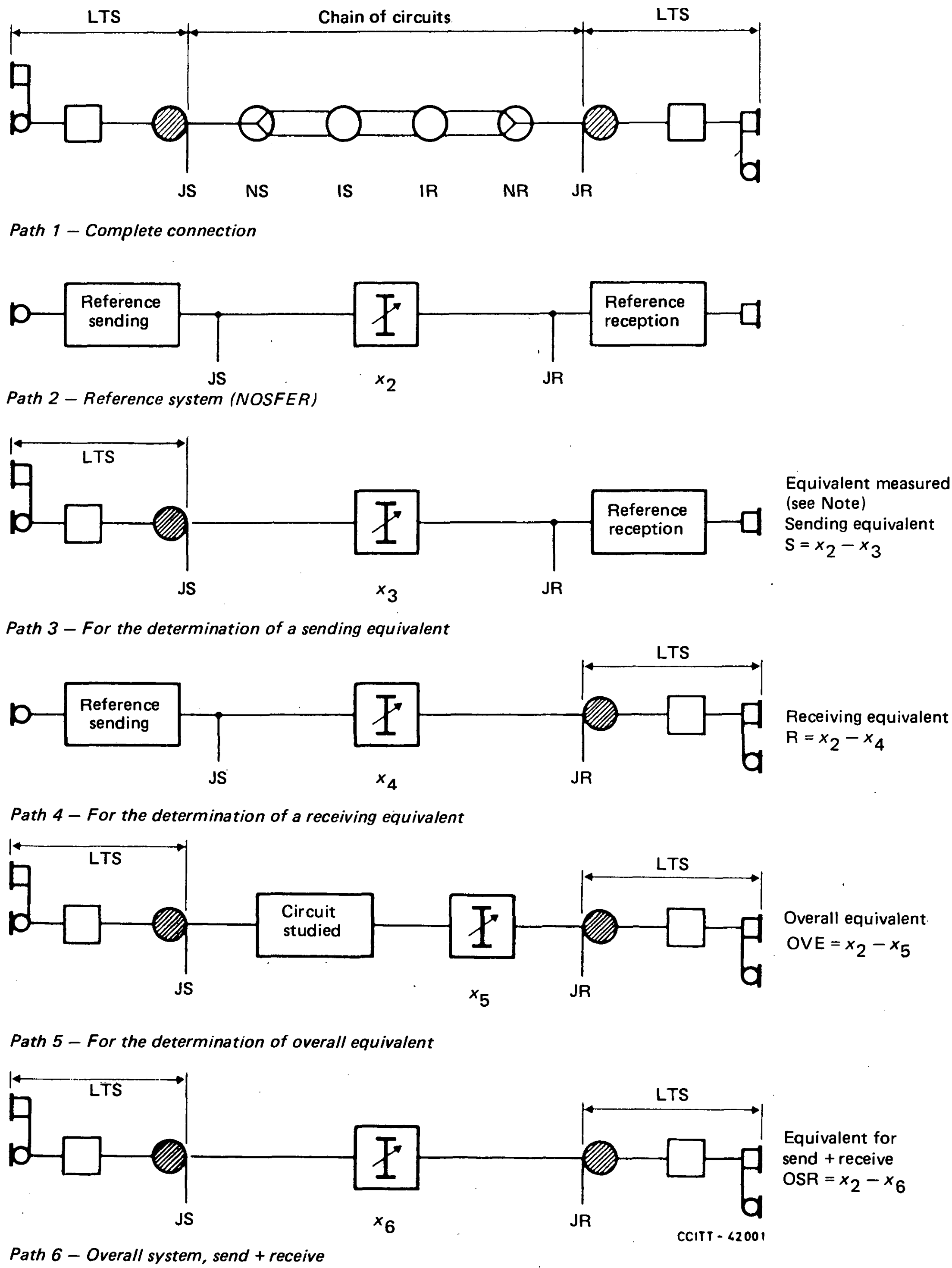
1.1 *Reference equivalent*

Reference equivalents are expressed in decibels and indicate the loss that has to be inserted into the reference system (NOSFER) in order to obtain subjectively the same ("equivalent") loudness as the system (or part thereof) under test. In practice for sending and receiving, x_3 and x_4 (Figure 1/P.72) are chosen to lie between 24 and 34 dB, and x_2 is adjusted to obtain the appropriate loudness balance.

1.2 *R25 equivalent*

From § 1.1 above it can be seen that the variation of the balance attenuator x_2 inserted in the NOSFER path causes the sensitivity of NOSFER to be altered while reference equivalents are being evaluated. The sound level at which the loudness balance is carried out is therefore variable depending on the settings of attenuators x_3 or x_4 (shown in paths 3 and 4 of Figure 1/P.72) and on the sensitivity of the unknown system under test.

It results from this that the reference equivalent of the circuit, with a loss of x throughout the transmitted band, is not equal to x which is a serious disadvantage for telephone network planning. To avoid this problem, the R25 equivalent has been developed, in which the NOSFER is maintained at a constant overall loudness, having a 25 dB attenuation inserted between send and receive ends (i.e. attenuator x_2, path 2 in Figure 1/P.72). This ensures that the loudness balance is carried out at a constant listening level whereby the test system is adjusted by attenuators to the same loudness as NOSFER.



JS : Junction send side
JR : Junction receive side
NS : National send side
IS : International send side
IR : International receive side
NR : National receive side

LTS = Local telephone system studied (telephone set + line + feeding bridge).

Note – x is obtained in each case by balancing the given path with NOSFER. The equivalents measured are:
– reference equivalents, if x_2 is varied so as to obtain the balance,
– R25 equivalents, if x_2 is set to 25 dB.

FIGURE 1/P.72

Connections and systems considered for the definition of reference equivalents and R25 equivalents

1.3 *Corrected reference equivalent*

An empirirical relationship between reference equivalent and R25 equivalent has been established based on a very large number of comparative measurements made in the CCITT Laboratory. Thus a reference equivalent may be "corrected" to estimate the R25 equivalent by the relationship:

$$\text{Corrected reference equivalent (CRE)} = 0.0082\ (\text{RE} + 70)^2 - 39.7\ \text{dB}.$$

For sending in a complete system, using a carbon microphone, this value should be reduced by 1 dB.

For a more detailed definition and the properties of CRE, see Annex A to Recommendation G.111.

1.4 *Relative equivalent*

If, instead of NOSFER, a working standard system is used as the reference standard against which commercial telephone systems are compared, the resulting rating is termed a relative equivalent. The reference equivalent of the working standard may then be taken into account in the following manner, e.g.:

Relative equivalent (or uncorrected mean result [1])		−5.0 (5.0 dB better, or louder)
Reference equivalent of the working standard system		+1.3 (1.3 dB worse, or quieter)
Reference equivalent of system under test	−5.0 + (+1.3) =	−3.7 (3.7 dB better, or louder)

2 Subjective determination of reference equivalent (RE)

In determining the RE or R25E of an LTS, a number of precautions are necessary, and these are detailed in § 4. Other facilities useful in carrying out the measurements may be found in the manual *Telephonometry* [1].

If NOSFER (or a replica meeting the requirements of Recommendation P.42) is used, results are expressed as x dB louder or "better" (M) or x dB quieter or "worse" (P) than NOSFER.

The descriptions given below apply when a NOSFER or a working standard system is used as the reference system.

2.1 *Method termed "two operator with hidden loss method"*

This method is based on the simultaneous use of three adjustable attenuators; one of these, the "balancing" attenuator (E) in the NOSFER path (see Figures 2/P.72 to 4/P.72), serves the purpose of adjusting the sound intensity at the receiving end. The second attenuator (S) which is the "hidden loss" and the third attenuator (A) which is the "common", are adjusted arbitrarily before the test (their attenuation values are unknown to the listening operator) in order to modify the apparent sensitivity of the compared systems. All three attenuators are distortionless, with input and output non-reactive impedances of 600 ohms and variable in 1 dB steps.

The settings of attenuators S and A are chosen arbitrarily between zero and 34 dB, but their total attenuation (A + S) should vary only between 24 and 34 dB.

Final reference equivalent values are obtained from the average of a number of individual telephonometric tests called "individual balances", carried out by different talker/listener pairs of operators from a team (see § 4.4).

2.1.1 *Send reference equivalent (SRE)* (see Figure 2/P.72)

Before an SRE may be determined, the handset must be fitted with a "guard-ring" to enable the talker's lips to be positioned correctly at the reference equivalent speaking position (RESP), (see § 4.1).

[1] Loudness of test system relative to the working standard.

To carry out an elementary balance, the talker speaks alternately into the two microphones repeating successively into each, one of the following conventional phrases, chosen so as to contain each of the principal vowel sounds:

Berlin, Hamburg, München, Koblenz, Leipzig, Dortmund (used in the Federal Republic of Germany)

Joe took father's shoe bench out. She was waiting at my lawn. (used in the United States of America)

Paris, Bordeaux, Le Mans, Saint-Leu, Léon, Loudun (used in France and in the CCITT Laboratory).

When speaking, the talker positions his lips to touch the plane of the guard-ring, and maintains the normal speaking level as defined for telephonometric measurements in § 4.2 below. He makes use of the speech voltmeter (SV) to monitor his talking level when speaking into the NOSFER microphone, then maintains the same speaking level into the microphone of the test system (when the SV is switched out of circuit). After each repetition of the conventional phrase the talker operates a changeover switch, thus connecting alternately the send parts of NOSFER and the system under test. (This switch may also be used to operate a lamp in the listening cabin to indicate which path is connected, thus helping to keep talker and listener in proper synchronization and to disconnect the SV when the test circuit is connected).

The listener, comparing the speech signals from the two paths in the NOSFER receiver, adjusts the attenuator E until the two systems sound equally loud. The attenuation settings of A, S, and E are then used to determine the SRE as detailed in Figure 1/P.72. Thus:

$$\text{SRE} = (x_2 - x_3) = (\text{E} + \text{A}) - (\text{S} + \text{A}) = \text{E} - \text{S dB}.$$

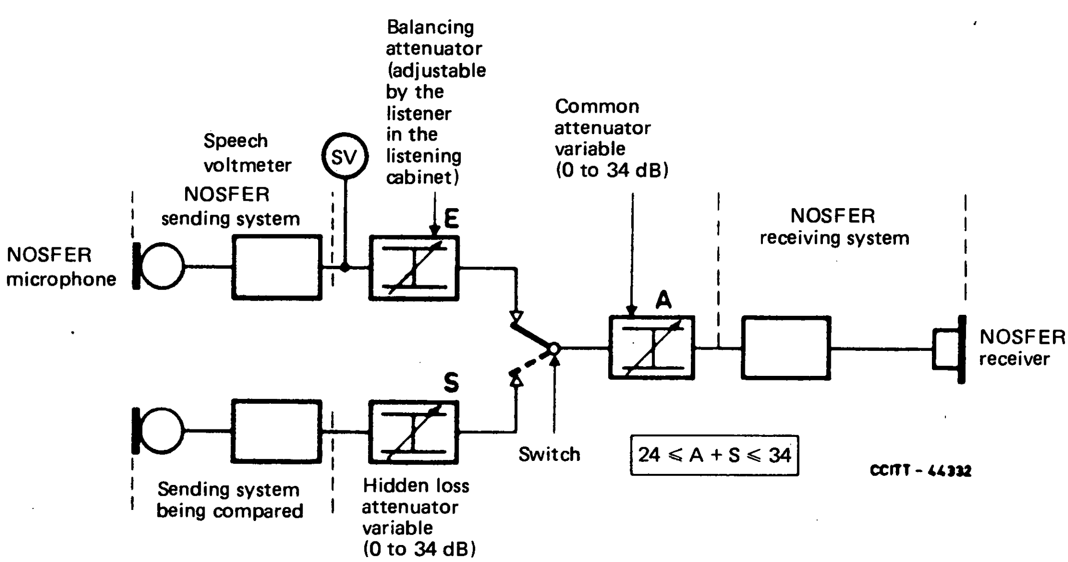


FIGURE 2/P.72

Comparison of a given sending system
with the NOSFER sending system for reference equivalents

2.1.2 *Receive reference equivalent (RRE)* (see Figure 3/P.72)

For this measurement, the talker repeats the conventional phrase into the NOSFER microphone maintaining the standard speech level for telephonometric tests (see § 4.2), while ensuring that his lips are just touching the guard-ring of the microphone. After each utterance he switched the circuit to connect alternately to the receive part of the NOSFER or to that of the LTS under test.

The listener holds in one hand the NOSFER receiver and the test receiver, listening to them alternately with the same ear while adjusting the balancing attenuator E until the speech signals in the two receivers sound equally loud. During this operation he must ensure as far as is possible that each receiver is coupled as closely as possible when applied to his ear.

The attenuation settings of A, S and E are then used to determine the RRE as detailed in Figure 1/P.72. Thus:

$$\text{RRE} = x_2 - x_4 = (\text{E} + \text{A}) - (\text{S} + \text{A}) = \text{E} - \text{S dB}.$$

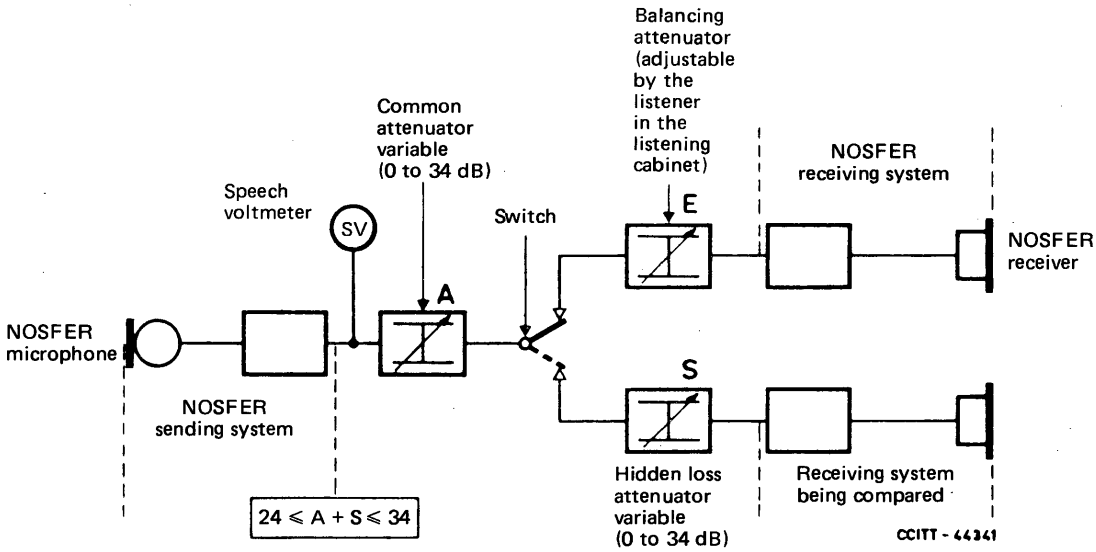


FIGURE 3/P.72

Comparison of a given receiving system
with the NOSFER receiving system for reference equivalents

2.1.3 *Reference equivalent of an overall system (send + receive, OSR)* (see Figure 4/P.72)

For this test the talker follows the same procedure as used for evaluating a sending system, speaking alternately between reference and test system microphones (see § 2.1.1). The listener follows the procedure for evaluating a receiver system, holding both reference and test system receivers (see § 2.1.2), and balancing the apparent loudness of the two paths using attenuator E.

The subjective test to find the overall reference equivalent of a complete system is rarely carried out at the CCITT Laboratory. However, in the case of a primary standard system or any other working standard system, the overall reference equivalent is often required.

The attenuation settings of attenuators A, S and E are then used to determine the OSR as detailed in Figure 1/P.72, path 6. Thus,

$$\text{OSR} = x_2 - x_6 = (\text{E} + \text{A}) - (\text{S} + \text{A}) = \text{E} - \text{S dB}.$$

2.2 *Method termed "three operator without hidden loss method"*

This method is not in current use in the CCITT Laboratory, therefore is not presented here. However the description of this method can be found in the former version of Recommendation P.72 [2].

3 Subjective determinaton of R25 equivalents (R25E)

To evaluate R25 equivalents the same procedures for talking and listening are used as for RE using NOSFER. The only difference is that in the NOSFER overall path a fixed 25 dB attenuation is inserted and the balance attenuator E is inserted in the path of the test system in series with the hidden loss attenuator S as illustrated in Figures 5/P.72 to 7/P.72. This means that as the listener adjusts attenuator E to find a balance, the test system loudness is changed rather than the NOSFER. The final R25E value is calculated as a mean of a number of individual balances found by different talker/listener pairs from the team.

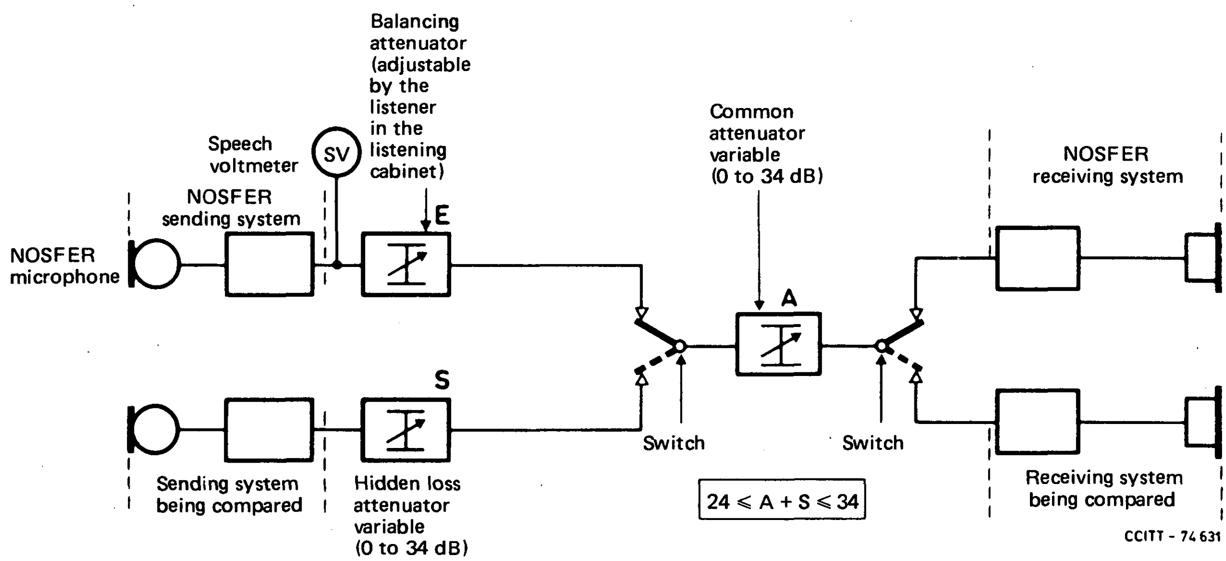


FIGURE 4/P.72

Comparison of a given overall system (send + receive) with the NOSFER system for reference equivalent

R25 equivalent may also be determined using the "three attenuators" method as used to determine RE, in which the sum of the settings of the two attenuators in the NOSFER path remains constant at 25 dB. For a full explanation of this method the reader is referred to § 2.2 of the manual *Telephonometry* [1]. If this method is used, the equations given in §§ 3.1 and 3.2 below are altered to reflect the different positions of the attenuators.

If a working standard system is used to determine R25 equivalent its overall path loudness should be adjusted to have a loudness quivalent to NOSFER with 25 dB connected between the send and receive parts.

3.1 *R25 equivalent of a sending system* (see Figure 5/P.72)

The procedure follows that given in § 2.1.1 and the sending R25E is calculated from the settings of attenuators S and E as detailed in Figure 1/P.72. Thus,

$$\text{SR25E} = x_2 - x_3 = 25 - (\text{E} + \text{S}) \text{ dB.}$$

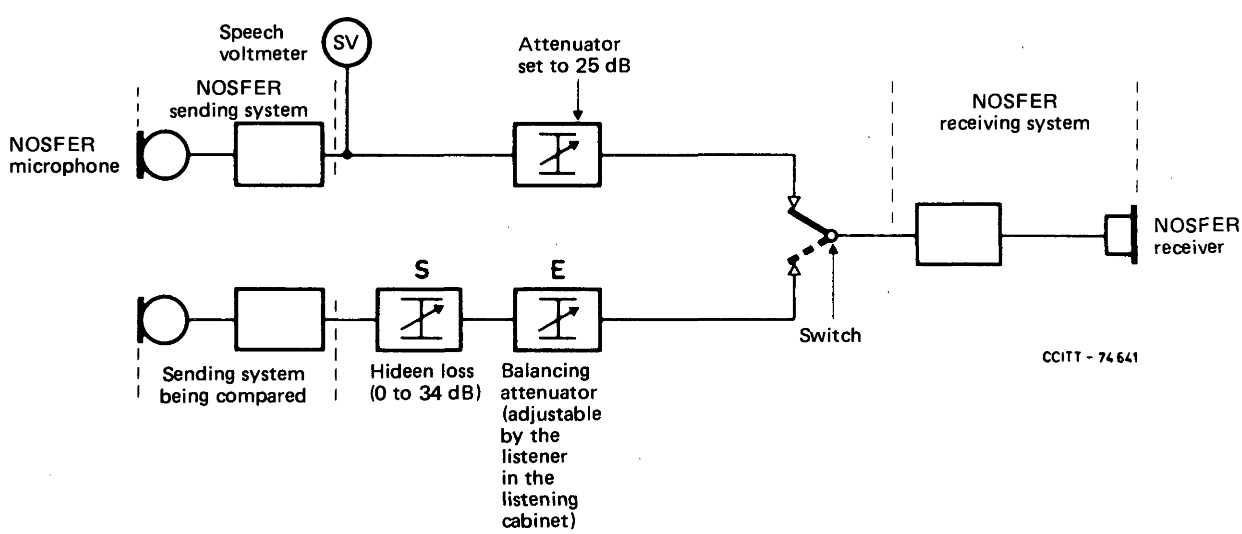


FIGURE 5/P.72

Comparison of a given sending system with the NOSFER sending system for R25 equivalents

3.2 *R25 equivalent of a receiving system* (see Figure 6/P.72)

The procedure follows that given in § 2.1.2 and the receiving R25E is calculated from the settings of attenuators S and E as detailed in Figure 1/P.72. Thus,

$$RR25E = x_2 - x_4 = 25 - (E + S) \text{ dB}.$$

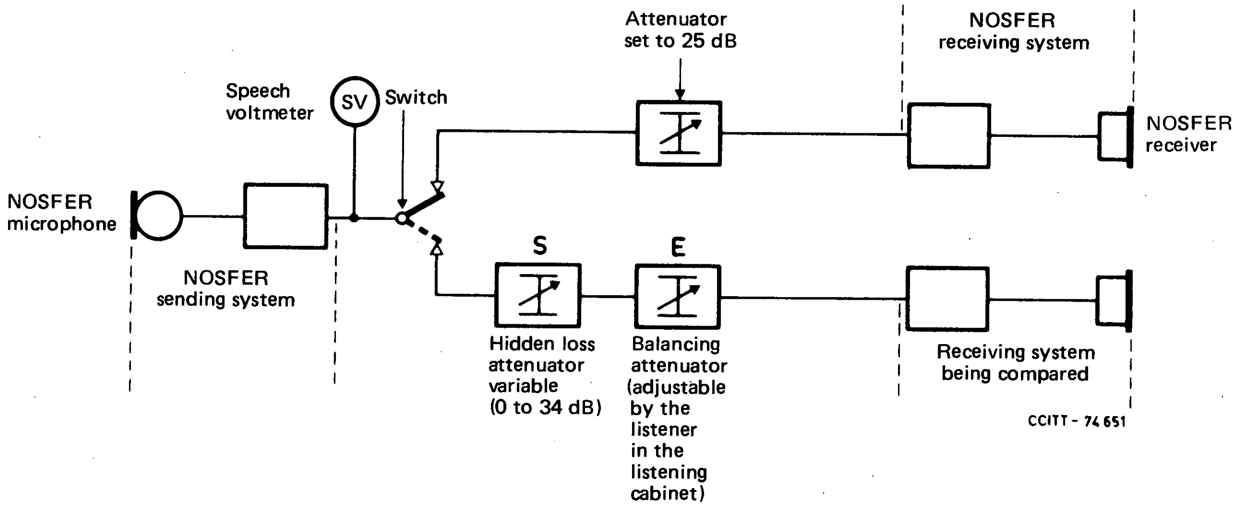


FIGURE 6/P.72

Comparison of a given receiving system with the NOSFER receiving system for R25 equivalents

3.3 *R25 equivalent of an overall system (OSR)* (see Figure 7/P.72)

The procedure follows that given in § 2.1.3 for an overall system. In this case no switching is required as there is no common path between NOSFER and the test system. Nevertheless the talker, repeating the conventional phrases alternately into the two microphones, operates the changeover switch to remove the speech voltmeter from the circuit when talking into a test system. This switch may also be used to break the overall NOSFER connection and/or operate an indicator lamp appropriately so as to keep talker and listener in proper synchronism for the balancing procedure.

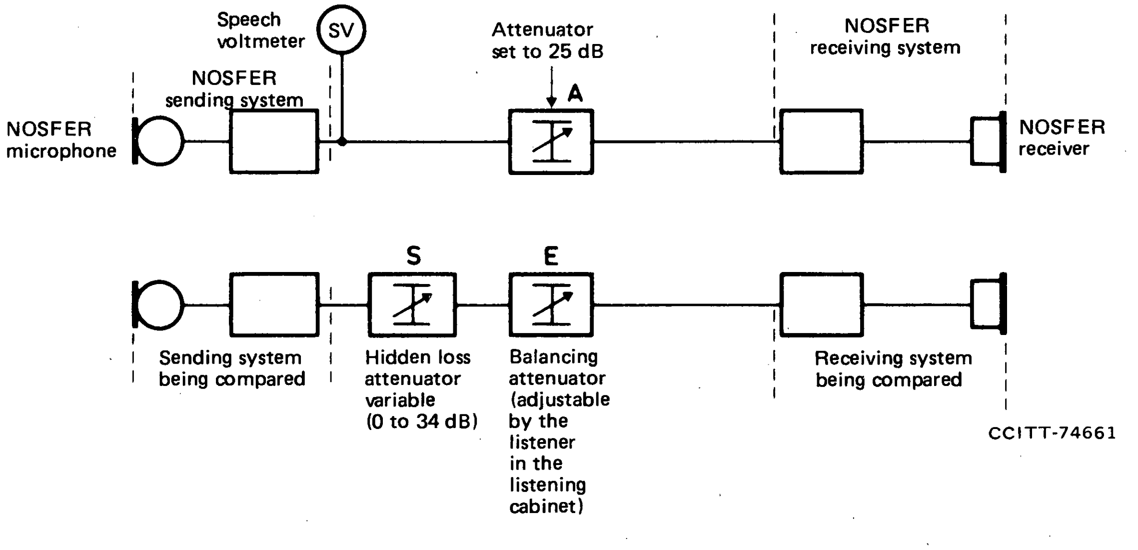


FIGURE 7/P.72

Comparison of a given overall system (send + receive) with the NOSFER system for R25 equivalents

3.4 *Overall R25 equivalent(OVE)*

The procedure is the same as for overall equivalent (send + receive) described in § 3.3, except that the electrical circuit forming the "junction" is included as shown in Path 5 of Figure 1/P.72. Balancing is performed by adjusting attenuator x_5, which in principle should be inserted in the connection in such a manner that the electrical impedance mismatches are not disturbed. Then:

$$\text{OVE} = x_2 - x_5 \text{ dB.}$$

This measurement is required in order to determine loudness insertion loss (LIL) of the junction circuit.

3.5 *Loudness insertion loss (LIL) of a junction circuit*

The LIL of an unknown junction circuit is given by the difference between the R25 equivalent of the overall system with the unknown junction included (Path 5 in Figure 1/P.72) and with the unknown removed (Path 6 in Figure 1/P.72). Then:

$$\text{LIL} = \text{OVE} - \text{OSR} = (x_2 - x_5) - (x_2 - x_6) = (x_6 - x_5) \text{ dB.}$$

4 Precautions to be taken during subjective telephonometric tests

4.1 *Lip position*

When the talker is uttering conventional phrases into the NOSFER or test microphone his lips must be maintained rigorously in the correct position using a small circular gauge termed the guard-ring. This ring which is of 25 mm internal diameter is fitted to the reference and test microphones as described below. The talker uses them by maintaining his lips just touching and approximately tangential to the plane of the guard-ring throughout the test.

4.1.2 *NOSFER guard-ring*

As described in Recommendation P.42 the NOSFER guard-ring is firmly mounted in a vertical position at a distance of 140 mm from the central axis of the reference microphone.

4.1.3 *Handset guard-ring for RE or R25E tests* (Reference equivalent speaking position, RESP)

From measurements made on the heads of a large number of individuals the characteristic head dimensions of an average subscriber have been determined together with the position in which he holds the handset to his ear during a telephone conversation. On the basis of these measurements the position of an average subscriber's lips relative to the opening of the ear canal has been determined. Figure A-1/P.76 shows this relationship.

From the midpoint of the earcap, a straight line is drawn towards the microphone in the plane of symmetry of the handset, making an angle α with the intersection of the plane of the earpiece of the receiver and the plane of symmetry of the handset and the distance δ is marked off along this line. The point thus determined is the centre of the guard-ring, which should coincide with the midpoint of the lips.

The intersection of the plane of this ring with the plane of symmetry will be a straight line, perpendicular to the direction of the speech defined above, i.e. the perpendicular to this straight line in the plane of symmetry will make an angle γ with the intersection of the plane of the earpiece. The position of the guard-ring in relation to the reference point in the centre of the earcap can be defined for reference equivalents and R25 equivalents by the following:

α = 15.5°

γ = 18°

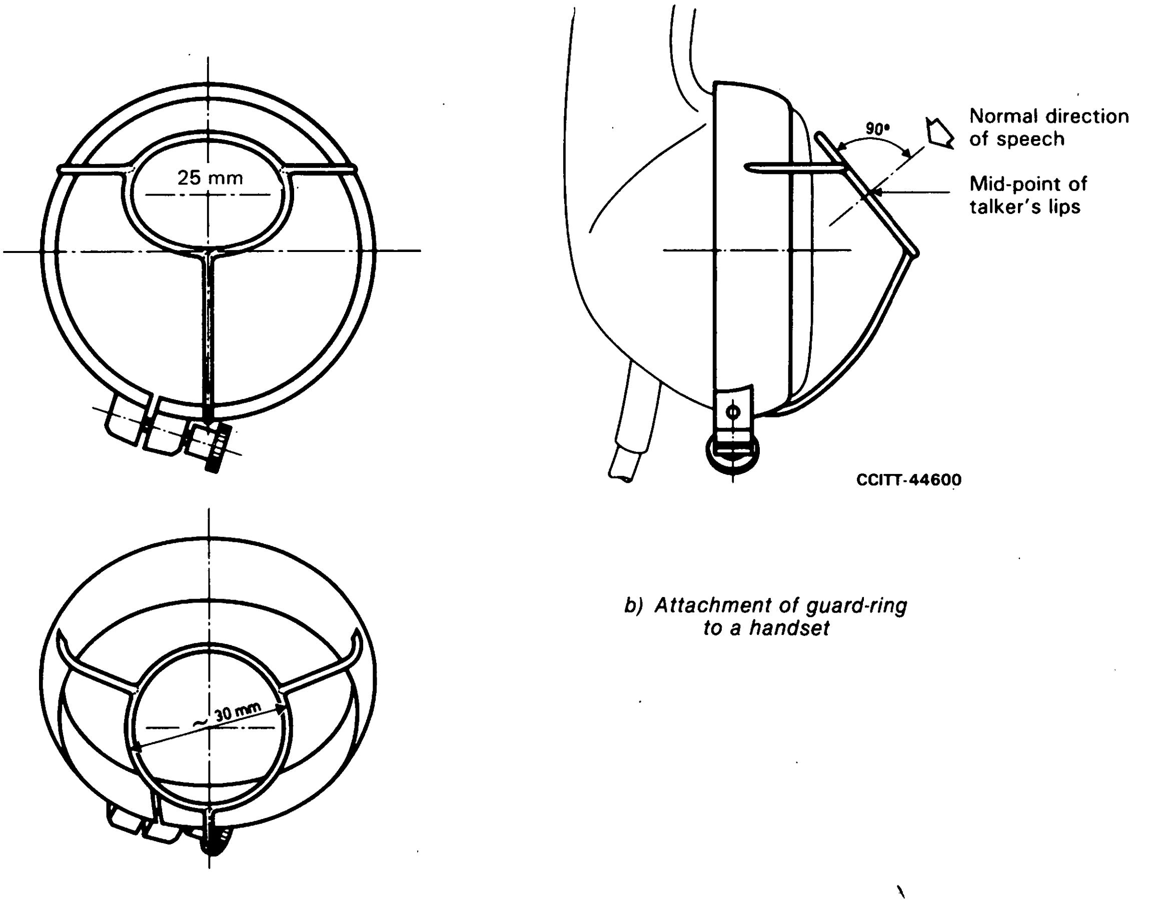
δ = 140 mm

The position of the guard-ring is thus completely determined and fixed with respect to the handset. The operator talks in such a manner that the median plane of his head is vertical, with the plane of the guard-ring perpendicular to the plane of the head. He then holds the handset with the inclination of the guard-ring at about 45° to the horizontal plane, with the head being inclined slightly forward, corresponding to a normal posture during conversation.

It should be noted that the position of the guard-ring, thus defined, has been fixed without reference to the inclination of the diaphragm of the microphone and does not necessarily correspond to the best operating conditions of the latter. When the handset is in a position described above the receiver is near the talker's ear and so he could be influenced by sidetone or by acoustic feedback. To prevent any problem in this respect the receiver may be blanked off with masking tape.

It is essential that the guard-ring and its mounting should be of a light construction in order not to cause any disturbance in the acoustic field in front of the microphone. It is equally important that the strain on the microphone case should not effect the mechanical and electrical properties of the microphone.

A device similar to that shown in Figures 8/P.72 and 9/P.72 is recommended.



a) Example of guard-ring to be fitted to a handset for RE or R25E

b) Attachment of guard-ring to a handset

FIGURE 8/P.72

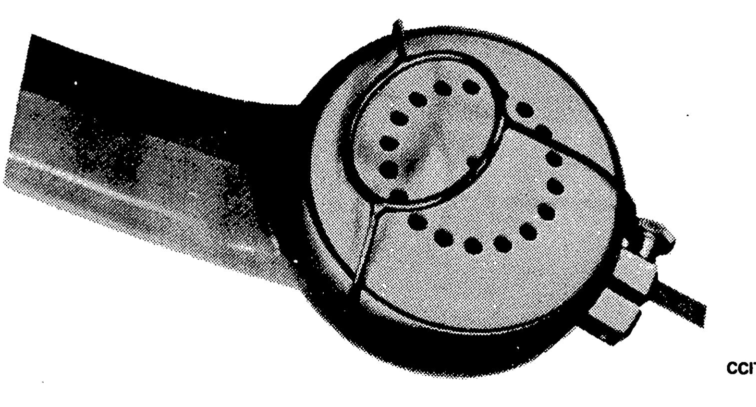


FIGURE 9/P.72

Guard-ring used by one Administration for tests of handsets

4.2 *Speech level control*

The speech volume produced by the talker during telephonometric measurements is of great importance as it influences the absolute and relative sensitivities of the equipment (especially in the case of carbon microphones). This volume must correspond to the standard level for telephonometric measurements used in the CCITT Laboratory.

4.2.1 *Standard volume*

A talker, speaking at the guard-ring of the NOSFER microphone, will be speaking at the standard volume when a speech voltage of -10 dBV, as measured using the ARAEN speech voltmeter, is generated across the output of the NOSFER send end into a 600 ohm resistive termination. The description of the ARAEN speech voltmeter can be found in Recommendation P.42, § 1.2 and its principal characteristics in Table 1/P.52.

4.3 *Packing effect*

To prevent the packing of carbon microphones under test, it is recommended that the microphone should be conditioned in accordance with Recommendation P.75 before each subjective balance adjustment.

4.4 *Choice and number of voice test team*

It has been found by experience in the CCITT and other laboratories that as a compromise between accuracy and speed of performing the tests, 6 test team members are generally adequate. Recommendation P.78, § 4 gives some guidance for their selection, and for their use of suitable speech material. It is preferable to make use of all 6 team members in one of the combinations given in § 5 below. Occasional use of only 5 members, if unavoidable, may be acceptable, but precision can be reduced compared with that using all 6 members.

4.5 *Auxiliary amplifier*

In certain circumstances when determining RE or R25E, it is necessary to include in series with the balance attenuator (E), an auxiliary amplifier (having 10 or 20 dB gain), in order to enable a loudness balance to be reached. This occurs in RE measurements, when the test system is louder than or close to the loudness of NOSFER, and for R25E when the test system is very much quieter than NOSFER. The reader is referred to § 2.2 of the manual *Telephonometry* [1], for a detailed explanation.

5 Experimental design and presentation of results

5.1 *Use of test team*

Many of the comments made in Recommendation P.78, § 3 are also relevant to the determination of RE and R25E. The team of six may be used in combination in various ways. A minimum of twelve operator pair combinations is suggested with a normal maximum of twenty. Twelve operator pair combinations can be derived from two teams of three (see Table 1a/P.78), eighteen operator pair combinations can be derived from one team of six (see Table 1b/P.78) and twenty operator pair combinations from one team of five (see Table 2a/P.78).

One team of six giving thirty operator pair combinations, known as the 6/6 operators method, produces a larger test for only slightly more precision than the previously mentioned team sizes. However, this combination of operator pairs is recommended for calibration of a primary or a working standard system and when performed on a regular basis (e.g. monthly) gives valuable data about the performance of each individual team member with time. For this latter purpose an analysis of the type presented in CCITT *Red Book*, Volume V, 1960, Annex 6 to Series-P Recommendations, is very suitable.

5.2 *Presentation of results*

Results of the given test condition repeated for *n* individual balances should be tabulated not only to ease the calculation of the results but also to help later verification. The method of recording the data and subsequent calculations depends on the type of equivalent (e.g. reference equivalent, R25 equivalent) being evaluated, irrespective of the operator method being used (5/5, 6/6, 3/6, or 3/6′).

By way of example Table 1/P.72 shows a subjective test sheet for reference equivalent evaluation using the 5/5 operator method. The individual reference equivalent result is calculated for each talker/listener pair as E – S as explained in § 2.1, and these values are entered in the column headed "Res". The mean result for the team and the standard deviation can then be derived using these individual results (see Recommendation P.78, Annex B).

For R25 equivalents it is also possible to produce individual results for each talker/listener pair from the equation,

$$\text{Res} = 25 - (\text{E} + \text{S}).$$

Entries would be made in columns S, E and Res on a form similar to that in Table 1/P.72 and the subsequent mean, and standard deviation calculated as for reference equivalent.

Table 2/P.72 shows another example of an R25 equivalent subjective test sheet which is somewhat either to use and less prone to error. In use, values for the balance attenuator E and hidden loss attenuator S are entered in the columns shown and for each talker/listener pair sum in the column headed x. All the subsequent analyses are then carried out on the x values to obtain a mean x and its standard deviation. The R25 equivalent is then worked out based on the mean x value obtained.

TABLE 1/P.72

Example of a subjective test sheet for reference equivalent evaluation using the 5/5 operators method

Listeners / Talkers	A				B				C				D				E				Total talkers
	A	S	E	Res	A	S	E	Res	A	S	E	Res	A	S	E	Res	A	S	E	Res	
A					2	30	34	+4	30	2	9	+7	5	27	31	+4	14	19	15	−4	+11
B	9	18	29	+11					17	16	20	+4	23	5	6	+1	1	24	29	+5	+21
C	23	3	11	+8	14	13	20	+7					4	26	29	+3	10	19	19	0	+18
D	17	16	22	+6	7	22	33	+11	1	25	33	8					23	5	9	+4	+29
E	1	25	33	+8	14	13	23	+10	18	9	20	+11	16	15	17	+2					+31
Total listeners				+33				+32				+30				+10				+5	+110

General total

Reference equivalent $= +\frac{110}{20} = +5.5$

Standard deviation = 3.89

Note – Values of A in the above table are not used for the calculation but introduced as a check that $24 \leqslant (A + S) \leqslant 34$.

TABLE 2/P.72

Example of a subjective test sheet for R25 equivalent evaluation using 3/6 method

$x = E + S$

Talker \ Listener	A			C			E			B			D			F			Talkers' x totals
	E	S	x	E	S	x	E	S	x	E	S	x	E	S	x	E	S	x	
A	–	–	–	17	20	37	12	15	27										64
C	24	8	32	–	–	–	17	13	30										62
E	15	20	35	19	13	32	–	–	–										67
B										–	–	–	20	13	33	13	15	28	61
D										21	10	31	–	–	–	12	19	31	62
F										7	17	24	19	15	34	–	–	–	58
Listeners' x totals			67			69			57			55			67			59	Total 374

R25 equivalent $= |25 - \bar{x}|$

$= 6.17$ dB.

Mean x ($\bar{x}$)	31.17
Standard deviation	3.44

ANNEX A

(to Recommendation P.72)

Remark on measurements of reference equivalent

It is necessary to draw a very clear distinction between, on the one hand, measurements required in the design and development of commercial telephone equipment to satisfy service conditions as well as possible and, on the other hand, the exchange between Administrations of numerical data which enable different types of equipment to be compared, from the standpoint of reference equivalent as one of the factors which affect transmission quality.

In the first case it is necessary to measure the sending and receiving sensitivities of the equipment over a wide range of variation of either the position of the subscriber's mouth with respect to the microphone or of the volume used or even of the feeding current value.

In the second case it is sufficient to give for each item a value of sending and receiving reference equivalent corresponding to a conventional position of the mouth with respect to the microphone and at a "standard" speech volume with a specified speech voltmeter.

The CCITT considers only the second case and for this reason it is not absolutely essential that the conventional position adapted for the mouth should correspond exactly to the mean position of the subscriber's mouth nor that the "standard" speech volume for telephonometric tests should coincide exactly with the mean value of speech levels found in service.

On the other hand, it is a great advantage if this conventional mouth position and this "standard" volume for telephonometric tests are used universally when it is simply a matter of communicating from one country to another general information on reference equivalents.

It follows from this that the values of sending and receiving reference equivalents corresponding to this conventional mouth position and "standard" volume for telephonometric tests are not necesssarily the same as those that would be obtained for the same items when in actual service.

References

[1] CCITT manual *Telephonometry* (to be published in 1985).

[2] CCITT Recommendation P.72 *Measurement of reference equivalents and relative equivalents*, Yellow Book, Vol. V, ITU, Geneva, 1981.

Recommendation P.73

MEASUREMENT OF THE SIDETONE REFERENCE EQUIVALENT

(amended at Malaga-Torremolinos, 1984)

For talker sidetone, a voice and ear measurement is made of the sidetone reference equivalent while speaking in a silence cabinet into the microphone of the set concerned, with the mouth at the normal speaking distance (see Recommendation P.72 § 4.1 and Figure A-1/P.76) from the diaphragm of the microphone; the receiver of the set situated some distance away in another silence cabinet where the sound level heard in this receiver is compared with that in the receiver of the NOSFER (or with that in the receiver of a working standard whose reference equivalent is known). (See Figure 1/P.73.)

The speech power to be used for this test is that used for sending and receiving reference equivalent determinations (see Recommendation P.72, § 4.2).

Equality of sounds heard is obtained by adjusting the balancing attenuator, E. A hidden-loss attenuator, S, situated close to the talking position enables the apparent sensitivity value of the complete NOSFER to be varied at will before the measurement and by an amount unknown to the listener. The value of the telephone sidetone reference equivalent is equal to the sum S + E of the values of the hidden-loss and balancing attenuators.

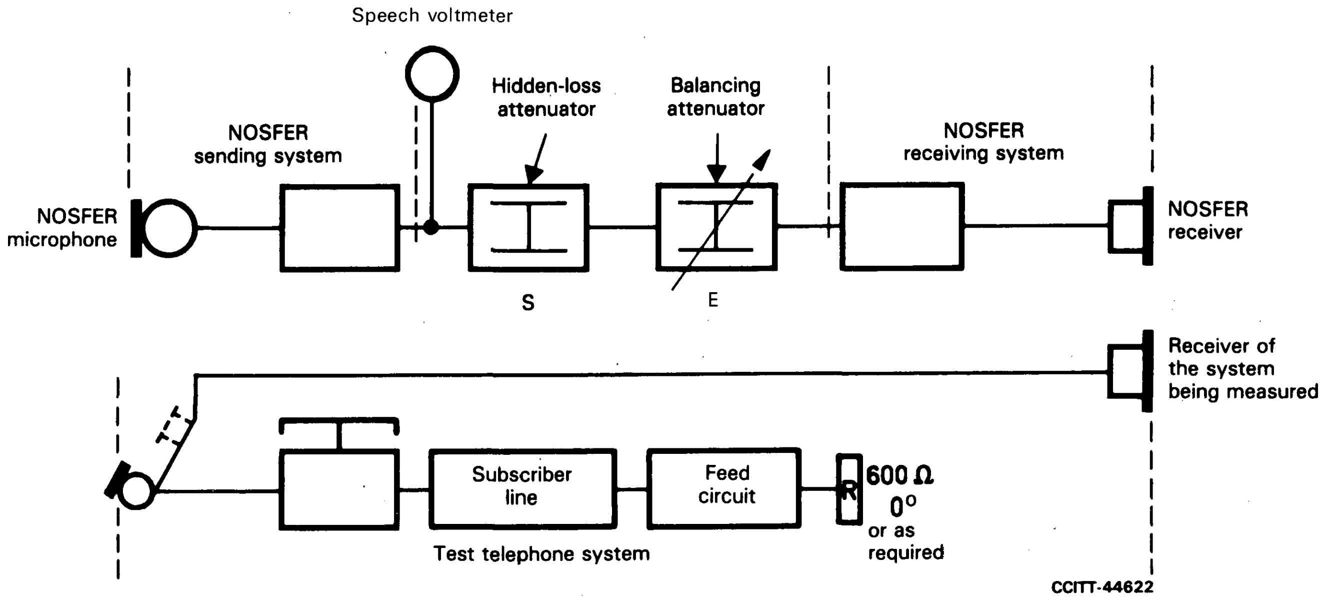


FIGURE 1/P.73

Measurement of the sidetone reference equivalent of a commercial telephone system

Whenever a result of a sidetone reference equivalent measurement is quoted for a telephone set it is necessary also to state the length and characteristics of the subscriber's line and the exchange terminating impedance to which it was connected during the measurement. The value of the feeding current and the sending and receiving reference equivalents of the telephone set may be provided as additional information.

In the past the CCITT has measured room noise sidetone by aural comparison between the NOSFER (or a calibrated working standard) and the sidetone path from microphone to receiver of the telephone set considered.

For this purpose the talking room has been subjected to room noise having appropriate level and spectrum from loudspeakers situated at specified distances from the microphones. The measurement technique used in the CCITT Laboratory is given in Figure 1/P.73 where the real voice is replaced by the room noise source.

The value of the reference equivalent of the sidetone path for room noise is equal to S + E – 17 dB. The correction of 17 dB takes account of the fact that under these conditions the NOSFER microphone is more sensitive than when used normally, e.g. as for a talker sidetone determination described above.

Note 1 – Studies have shown that for talker sidetone a rating method which correlates better with subjective effects than reference equivalent is one which takes into account the human sidetone path as a masking threshold, STMR (sidetone masking rating). STMR is documented in Recommendations P.76, § 3 and P.79, § 5.

Note 2 – The room noise aspects of sidetone are at present being studied in Study Group XII under Question 9/XII [1]. Recommendation P.11, § 2.4 describes some of the effects of sidetone in a telephone connection.

Reference

[1] CCITT – Question 9/XII, Contribution COM XII-No. 1, Study Period 1985-1988, Geneva, 1985.

METHODS FOR SUBJECTIVE DETERMINATION OF TRANSMISSION QUALITY

1 Introduction

This Recommendation contains advice to Administrations on conducting subjective tests in their own laboratories. The tests carried out in the CCITT Laboratory by using reference systems are described in Section 3 of this Volume.

In the course of developing items of telephone equipment, it is necessary to conduct various kinds of specialized tests to diagnose faults and shortcomings; such tests dedicated to the study of specific aspect of transmission quality are not discussed here. The present purpose is to indicate methods that have been found suitable for determining how satisfactory given telephone connections may be expected to be if offered as such for use by the public.

The methods indicated here are intended to be generally applicable whatever the form of any degrading factors present. Examples of degrading factors include transmission loss (often frequency dependent), circuit and room noise, sidetone, talker echo, nonlinear distortion of various kinds, propagation time, deleterious affects of voice-operated devices and changes in characteristics of telephone sets, including loudspeaking sets. Combinations of two or more of such factors have to be catered for.

2 Recommended methods

To be applicable for such a wide range of types of degrading factor given in § 1, the assessment method must reproduce as far as possible all the relevant features present when customers converse over telephone connections. Suitable methods are referred to as "Conversation Tests" and detailed prescriptions on the conduct of such tests as carried out by British Telecom are given in Supplement No. 2 at the end of this fascicle.

If the rather large amount of effort needed is available and the importance of the study warrants, transmission quality can be determined by service observations and recommended ways of performing these, including the questions to be asked when interviewing customers, are given in Recommendation P.77.

A disadvantage of the service observation method for many purposes is that little control is possible over the detailed characteristics of the telephone connections being tested. A method that largely overcomes this disadvantage but retains many of the advantages of service observations is that used by the AT&T Co. and termed SIBYL (refer to Supplement No. 5 at the end of this fascicle). According to this method, members of the staff of Bell Laboratories volunteer to allow a small proportion of their ordinary internal calls to be passed through special arrangements which modify the normal quality of transmission according to a test programme. If a particular call has been so treated the volunteer is asked to vote by dialling one of a set of digits to indicate his opinion. In this way all results are recorded by the controlling computer and complete privacy is retained.

3 Supplementary methods

Under certain conditions, it is permissible to dispense with the full conversation method and to use one-way listening-only tests. Suitable conditions apply for using a listening test when the degrading factor(s) under study affect the subjects only in their listening role. Attenuation/frequency distortion and nonlinear distortion caused by quantizing have been studied successfully by listening tests but it would be unwise to study the effects of sidetone, for example, by this method. Listening-only tests may also be misleading when assessing the effects of a factor, like circuit noise, when the magnitude of the degradation caused is substantial. In any case, sufficient comparison with the results from full conversation tests should be made before the results from listening-only tests are accepted as reliable.

Recommendation P.75

STANDARD CONDITIONING METHOD FOR HANDSETS WITH CARBON MICROPHONES

(Geneva, 1972; amended at Malaga-Torremolinos, 1984)

1 Since the characteristics of carbon microphones are strongly dependent on conditioning techniques, it is necessary to follow a consistent procedure prior to measuring sensitivity/frequency characteristics in order to obtain reproducible results. The CCITT recommends that for best reproducibility, automatic mechanical conditioning be used. The following steps are specified for the *standard conditioning method*:

a) Place the handset in a holding fixture with the handset clamped in a position corresponding to that in which the microphone is going to be measured [e.g. loudness rating guard-ring position (LRGP) according to Annex A of Recommendation P.76, or reference equivalent guard-ring position, according to Recommendation P.72]

b) Connect the microphone or telephone set terminals as required to the d.c. feed circuit and appropriate terminating loading.

c) Turn the feed current on. After 5 seconds, condition the microphone by rotating it smoothly. Rotation is made such that the plane of the granule bed moves through an arc of at least 180°. The procedure is repeated twice with the handset coming to rest finally in the test position. The time of each rotation cycle should lie within the range of 2 to 12 seconds.

2 When carrying out subjective tests with a carbon microphone telephone set, the conditioning of the handset should be done by the talker. This conditioning should conform to the conditioning for objective measuring as described under § 1 above insofar as it is practicable.

Recommendation P.76

DETERMINATION OF LOUDNESS RATINGS; FUNDAMENTAL PRINCIPLES

(Geneva, 1976; amended at Geneva, 1980 and Malaga-Torremolinos, 1984)

Preface

This Recommendation is one of a set of closely related Recommendations concerned with determination of loudness ratings. The present one deals with the fundamental principles and the others, as follows, deal with certain additional matters [1]).

Recommendation P.48	Specification for an intermediate reference system
Recommendation P.78	Subjective testing method for determination of loudness ratings in accordance with Recommendation P.76
Recommendation P.64	Determination of sensitivity/frequency characteristics of local telephone circuits to permit calculation of their loudness ratings
Recommendation P.79	Calculation of loudness ratings
Recommendation P.65	Objective instrumentation for measuring loudness ratings

1) The present Recommendation together with Recommendations P.48, P.78 and P.79 provide complete definitions of overall, sending, receiving and junction loudness ratings and Administrations are invited to use them to further their studies of Question 19/XII [1].

1 Introduction

A speech path is, broadly, a transmission path that exists between a talker's mouth and the ear of a listener or, in the case of sidetone, between the mouth and ear of a talker. In typical face-to-face conversation, the speech is transmitted by means of the air path connecting the mouth and ear. Depending on environmental conditions, transmission may be:

a) more or less direct, as in the case of two persons conversing in an open, unobstructed location, such as a golf course;

b) largely indirect, as in the case of two persons conversing in a small, hard surfaced room where a large proportion of the energy reaching the ear may be due to reflections from the walls, ceilings and floor; or

c) something between the two extremes of *a)* and *b)*.

In the case of telephony, the air path is replaced by a system comprising:

a) an air path from the mouth to the telephone microphone;

b) an air path between the telephone earphone and the ear; and

c) a telephone connection consisting of the microphone, earphone and interconnecting circuitry together with a similar system for the reverse direction of transmission. The two situations – face-to-face and using the telephone – differ appreciably in detail but, for speech transmission purposes, they are alike as far as their function is to provide a means of both-way speech communication.

Telephone engineering is concerned with providing telephone connections which, while not identical to the face-to-face situation, are comparable in effectiveness for providing a means of exchanging information by speech; such telephone connections should also optimize customer satisfaction within technical and economic constraints.

Various tools are used by transmission engineers in planning, design and assessment of the performance of telephone networks. Reference equivalent, based on the criterion of loudness of speech emitted by the talker and perceived by the listener, has been one of the most important of these tools; it provides a measure of the transmission loss, from mouth to ear, of a speech path.

The *reference equivalent method* is defined in Recommendations P.42 and P.72 and the fundamental principles are briefly explained in [2]. The method for determining *loudness ratings* of local telephone circuits is in accordance with rather similar fundamental principles to those of the reference equivalent method but embodies modifications which render the method much more flexible and should greatly simplify transmission planning.

A desire to depart from use of reference equivalents as defined by Recommendation P.72 arises for the following reasons:

1) Reference equivalents cannot be added algebraically; discrepancies of at least ± 3 dB are found.

2) Replication accuracy of reference equivalents is not good; changes in crew can cause changes of as much as 5 dB.

3) Increments of real (distortionless) transmission loss are not reflected by equal increments of reference equivalent; 10 dB increase in loss results in an increase in reference equivalent of only about 8 dB.

Use of loudness ratings defined in accordance with the principles given below should largely obviate these troubles.

In addition to these advantages, the same values of loudness ratings should be obtained whether the determination is by subjective tests, by calculation based on sensitivity/frequency characteristics or by objective instrumentation. The fundamental principles of the method are described below and these differ from those applicable to reference equivalents by the least possible extent to achieve the desirable flexibility.

The loudness rating (which has the dimensions and sign of "loss") is, in principle, like the reference equivalent, defined by the amount of loss inserted in a reference system to secure equality of perceived loudness to that obtained over the speech path being measured. Practical telephone connections are composed of several parts connected together. To enable the transmission engineer to deal with these parts in different combinations, loudness ratings must be defined in a suitable manner so that "overall", "sending", "receiving" and "junction" ratings can be used.

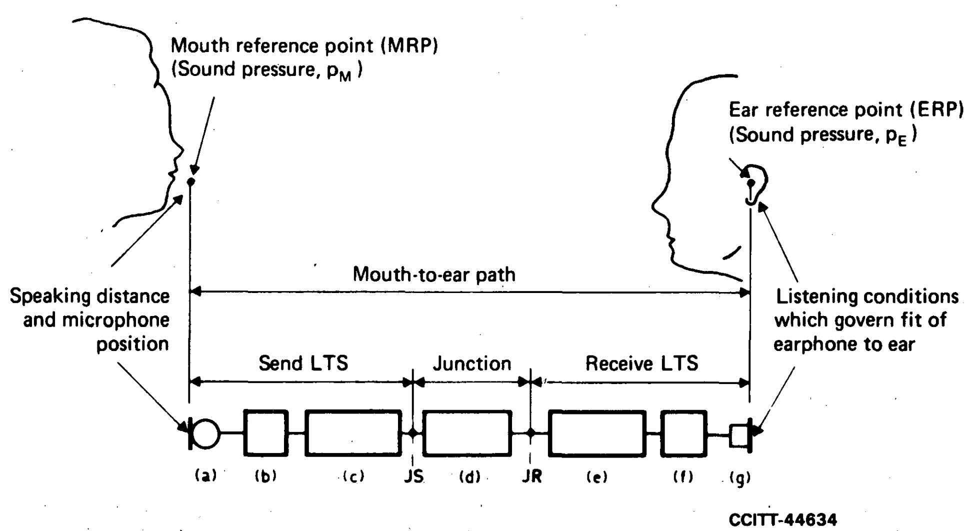
"Sidetone" loudness ratings can also be determined in an analogous manner. Sidetone reference equivalent is defined in Recommendation P.73 and sidetone loudness ratings are defined in § 3 below.

2 Definitions of loudness ratings for principal speech paths

2.1 *General*

§ 2 deals with principal speech paths, namely from a talker at one end of a connection to a listener at the other. Sidetone paths are treated in § 3 below.

In general, loudness ratings are not expressed directly in terms of actual perceived loudness but are expressed in terms of the amounts of transmission loss, independent of frequency, that must be introduced into an *intermediate* reference speech path and the *unknown* speech path to secure the same loudness of received speech as that defined by a fixed setting of NOSFER. This implies that some interface exists or could, by some arrangement, be found in the unknown speech path into which the transmission loss can be introduced. In practice the unknown speech path is composed of a sending local telephone circuit coupled to a receiving local telephone circuit through a chain of circuits interconnecting the two local systems[2]. Figure 1/P.76 shows this subdivision of one principal speech path of a telephone connection. The interfaces JS and JR separate the three parts of the connection to which loudness ratings are assigned, namely: *sending loudness rating*, from the mouth reference point to JS; *receiving loudness rating* from JR to the ear reference point; and *junction loudness rating* from JS to JR. The *overall loudness rating* is assigned to the whole speech path from mouth reference point to ear reference point.



Note – (a) represents the microphone of the sending local telephone system;

(b) represents the electrical circuit of the telephone set of the sending local telephone system;

(c) represents the subscriber's line and feeding/transmission bridge of the sending local telephone system;

(d) represents the chain of circuits interconnecting the two local systems;

(e) represents the subscriber's line and feeding/transmission bridge of the receiving local telephone system;

(f) represents the electrical circuit of the telephone set of the receiving local telephone system;

(g) represents the earphone of the receiving local telephone system.

FIGURE 1/P.76

Subdivision of a telephone connection

---

2) See Annex B for explanation of certain terms.

Note that in practical telephone connections:

a) the transmission loss of the junction may be frequency dependent;

b) the image impedances of the "junction" may not be constant with frequency and may not be resistive;

c) the impedances of the local telephone systems presented to the junction at JS and JR may not be constant with frequency and may not be resistive;

d) impedance mismatches may be present at JS or JR or both.

Overall loudness ratings (OLRs), sending loudness ratings (SLRs), receiving loudness ratings (RLRs) and junction loudness ratings (JLRs) are defined so that the following equality is achieved with sufficient accuracy for practical telephone connections.

$$\text{OLR} = \text{SLR} + \text{RLR} + \text{JLR}$$

2.2 *Definitions of overall, sending, receiving and junction loudness ratings*

Figure 2/P.76 shows the principles used to define the overall, sending, receiving and junction loudness ratings.

2.2.1 *Overall loudness rating*

Path 1 in Figure 2/P.76 shows the complete unknown speech path subdivided into local telephone systems and junction. In this example the junction comprises a chain of circuits represented by trunk junctions (JS-NS and NR-JR) and trunk circuits (NS-IS, IS-IR and IR-NR). A suitable arrangement for inserting transmission loss independent of frequency must be provided at some point such as in IS-IR.

Path 2 shows the complete IRS with its adjustable, non-reactive, 600 ohms junction between JS and JR.

The level of received speech sounds to which the additional loss x_1 in Path 1 and the junction attenuator setting x_2 of Path 2 are both adjusted is defined by using the fundamental reference system NOSFER with its attenuator set at 25 dB. When these adjustments have been made, the overall loudness rating (OLR) of the complete unknown connection is given by $(x_2 - x_1)$ dB.

2.2.2 *Sending loudness rating*

Path 3 in Figure 2/P.76 shows the IRS with its sending part replaced by the local telephone system of the unknown. The junction is adjusted to produce, via Path 3, the same loudness of received speech sounds as the NOSFER with its attenuator set at 25 dB. If x_3 is the required setting in Path 3, the sending loudness rating (SLR) is given by $(x_2 - x_3)$ dB.

2.2.3 *Receiving loudness rating*

Path 4 in Figure 2/P.76 shows the IRS with its receiving part replaced by the local telephone system of the unknown.

The junction is adjusted to produce via Path 4 the same loudness of received speech sounds as the NOSFER with its attenuator set at 25 dB. If x_4 is the required setting in Path 4, the receiving loudness rating (RLR) is given by $(x_2 - x_4)$ dB.

2.2.4 *Junction loudness rating*

Path 5 in Figure 2/P.76 shows the IRS with its junction replaced by the unknown chain of circuits as located in Path 1 of Figure 2/P.76 between JS and JR. The arrangement for introducing transmission loss, independent of frequency, must be provided as was required in Path 1. The additional loss is adjusted to produce, via Path 5, the same loudness of received speech as the NOSFER with its attenuator set at 25 dB. If x_5 is the required additional loss in Path 5, the junction loudness rating is given by $(x_2 - x_5)$ dB.

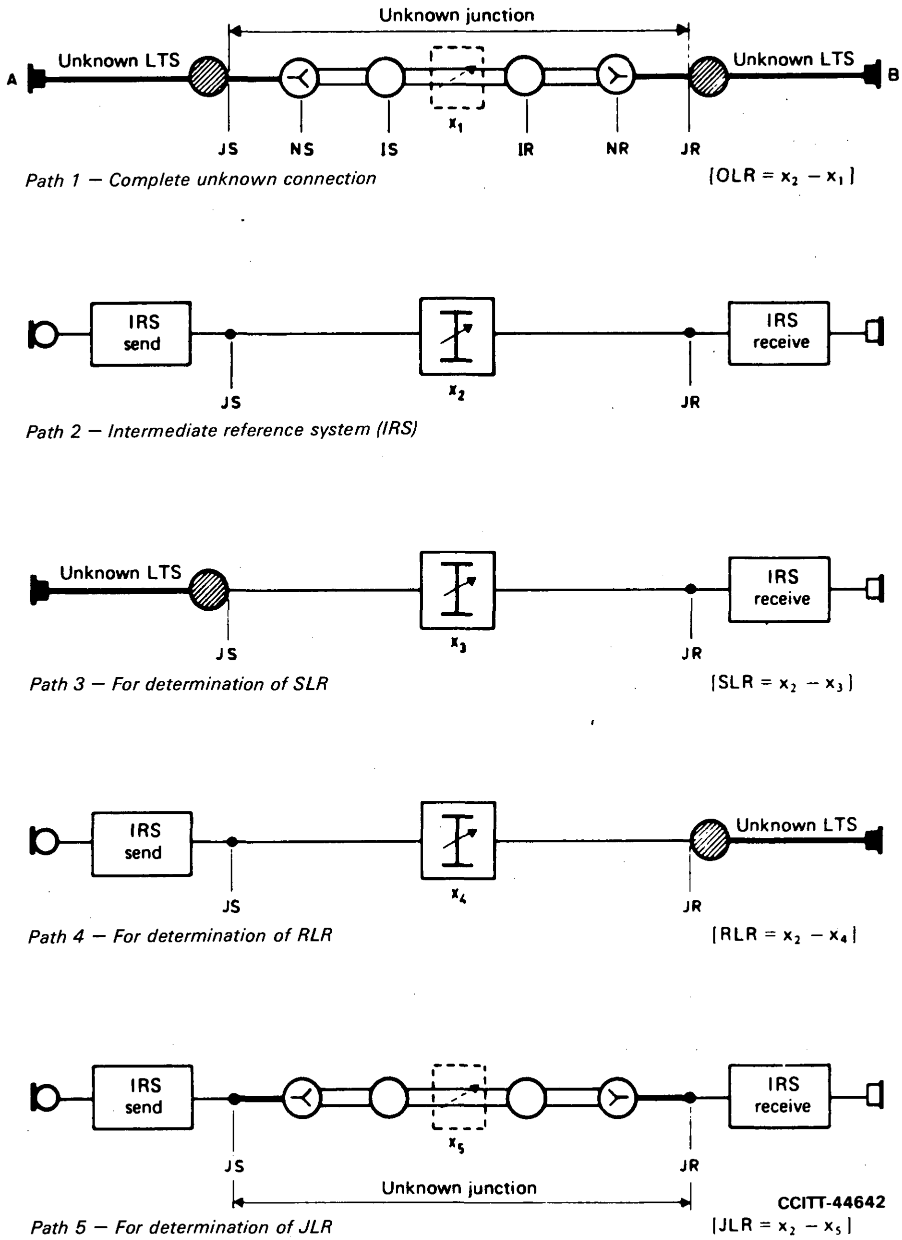


FIGURE 2/P.76

Principles used for defining OLR, SLR, RLR and JLR

2.3 *Conditions under which loudness ratings are determined*

2.3.1 *General*

The loudness of received speech sounds depends upon certain factors that are not well defined under practical conditions of use, but must be defined as precisely as possible to obtain accurately reproducible loudness ratings. Clearly, as shown in Figure 1/P.76, the loudness rating is largely governed by the characteristics of the mouth-to-ear path. This path can be made precise by defining a *mouth reference point* at which the sound pressure p_M of speech emitted by the talker is measured or referred, and an ear reference point at which to measure or to which to refer the sound pressure p_E of speech reproduced by the earphone. These points can be chosen in a fairly arbitrary manner and this becomes important when loudness ratings are to be determined objectively; suitable definitions for such purposes are given in Recommendation P.64 which deals with measurement of sending and receiving sensitivity/frequency characteristics.

It is essential, however, to define vocal level, speaking distance, microphone position and listening conditions which govern the fit of the earphone to the ear. These are indicated in Figure 1/P.76. The essential features that define the conditions under which loudness ratings are determined are indicated in Table 1/P.76.

Some remarks on the items listed in Table 1/P.76 are given below.

TABLE 1/P.76

Conditions under which loudness ratings are determined

No.	Item specified	Specification
1	Intermediate reference system	Recommendation P.48
2	Vocal level of speaker	As Recommendation P.72
3	Level of received speech sounds at which loudness is judged constant	NOSFER set at 25 dB
4	Handset position relative to talker's mouth	See Annex A
5	Direction of speech	Head erect
6	Handset arrangement for listening	See § 2.3.7
7	Conditioning of carbon microphones	Recommendation P.75

2.3.2 *Intermediate reference system*

The intermediate reference system is defined in Recommendation P.48. It has been chosen with the following in mind:

a) It shall correspond approximately, as far as the shapes of sending and receiving frequency characteristics are concerned, with those of national sending and receiving systems in use at present and likely to be used in the near future. For this reason the frequency bandwidths for sending and receiving parts are confined to the nominal range 300-3400 Hz[3].

b) The absolute sensitivity has been chosen to reduce as much as possible changes in values from reference equivalents to loudness ratings.

c) In external form its handsets are similar to conventional handsets used in actual telephone connections.

2.3.3 *Vocal level of speaker*

The vocal level at which speech is emitted from the speaker's mouth conforms to that in use for determining reference equivalents and is defined in Recommendation P.72. This approximates the level actually used by customers under good transmission conditions. It is defined in terms of the speech level at the output of the NOSFER sending system.

2.3.4 *Listening level*

The level of received speech sounds at which loudness is judged constant is defined by the vocal level (see § 2.3.3 above) and the setting (25 dB) of NOSFER against which all the speech paths shown in Figure 2/P.76 are adjusted. This corresponds to a fairly comfortable listening level of the same order as that commonly experienced by telephone users.

[3] The IRS is specified for the range 100-5000 Hz (see Recommendation P.48). The nominal range 300-3400 Hz specified is intended to be consistent with the nominal 4 kHz spacing of FDM systems, and should not be interpreted as restricting improvements in transmission quality which might be obtained by extending the transmitted frequency bandwidth.

2.3.5 *Handset position*

The position of the telephone handset relative to the talker's mouth is defined in Annex A to this Recommendation. It is intended to approximate fairly well the position used by customers under real telephone connections. The definition covers not only the distance between lips and mouthpiece but also the attitude of the microphone relative to the horizontal axis through the centre of the mouth opening. It is defined in such a way that the lips-to-mouthpiece distance becomes greater as the length of a handset is increased.

2.3.6 *Direction of speech*

The speaker shall hold his head erect and it will be assumed that speech is emitted horizontally from his mouth.

2.3.7 *Handset arrangement for listening*

The listener shall hold the handset in his hand with the earphone placed comfortably against his ear.

2.3.8 *Conditioning of carbon microphones*

Telephone handsets with carbon microphones usually require to be conditioned. This shall be done in accordance with Recommendation P.75.

3 Sidetone loudness rating

Although it is possible, in principle, to define a sidetone loudness rating exactly as overall loudness rating, it is more satisfactory to use a definition in which human sidetone is treated as a masking threshold.

3.1 *Definition of sidetone masking rating (STMR)*

When a telephone subscriber speaks, his own voice reaches his ear by several paths. These are (see Figure 3/P.76):

a) through the telephone set circuit from microphone to earphone due to mismatch of the hybrid balance impedance within the set and the line impedance;
b) through the mechanical path within the human head;
c) through the acoustic path to the ear and involving leakage at the earcap and human ear interface;
d) through the mechanical path along a handset handle [although this may be measured in fact as a contribution to a) above].

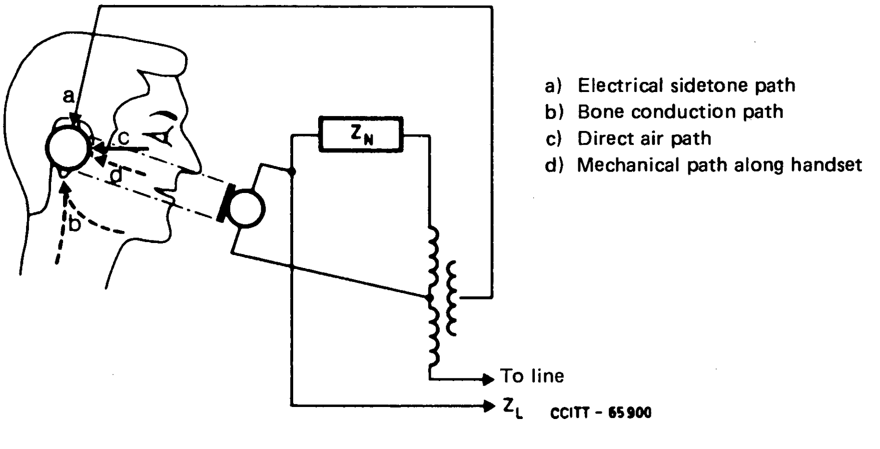


FIGURE 3/P.76

Sidetone paths through which a telephone subscriber may hear his own voice

Determination of these sidetone paths will usually resolve into two main measurements, a) + d) and b) + c). Each is referred to the speech signal at the mouth reference point (MRP) and the measurement made at the ear reference point (ERP).

Thus L_{MEST} is the loss from the mouth to ear (MRP to ERP) of the telephone sidetone path, and L_{MEHS} is the loss from mouth to ear (MRP to ERP) of the human sidetone path.

Note – Recommendation P.64, § 8 describes a method for the measurement of S_{meST}, the sidetone sensitivity/frequency characteristic of a telephone set using the artificial mouth and ear, from which an estimate of S_{MEST} using the human mouth and ear may be obtained by adding correction L_M and L_E as explained in the text. Thus:

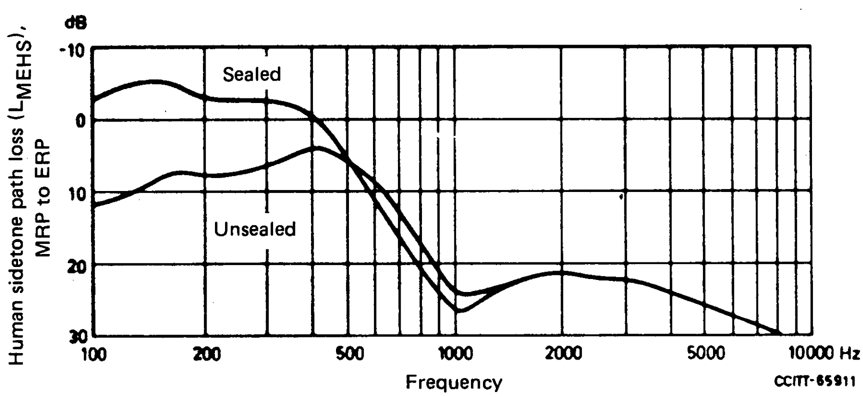
$$L_{MEST} = S_{MEST} \text{ in dB}$$

L_{MEST} and L_{MEHS} are each usually measured at a number of frequencies in the ISO range of ⅓rd octave frequencies, typically at least 200 to 4000 Hz. Where complex signals are used (for example, during the measurement of L_{MEHS} the subjects' speech signals were used), spectrum density measurements must be made.

Studies completed so far have indicated that for talker sidetone at least, the rating method which correlates best with subjective effects of sidetone is one which takes into account the human sidetone signal as a masking threshold, i.e. sidetone masking rating (STMR).

3.2 *Determination of sidetone masking rating (STMR)*

It has been determined that when a subscriber speaks with his ear occluded by a telephone receiver the signal at the ERP via the path L_{MEHS} is very significant, particularly at low frequencies (less than 1000 Hz) and warrants inclusion in any method used to rate the path L_{MEST} for talker sidetone. (Figure 4/P.76 shows the average L_{MEHS} for 6 females and 6 males for 2 earphone coupling conditions, sealed, for which stringent precautions were taken, and unsealed using a leakage typical of conversational conditions).



Note – Measurement conditions: handset, British Telecom No. 3 receiver, rocking armature, 4.T. The earphone coupling leakage path used for the unsealed case was typical of that experience under conversational conditions.

FIGURE 4/P.76

Human sidetone characteristic (for an occluded ear) for sealed and unsealed conditions (plotted from values given in Table 5/P.79)

A satisfactory method for taking L_{MEHS} into account when rating a telephone sidetone path L_{MEST}, is one which in effect treats L_{MEHS} as a masking threshold against which the loudness of L_{MEST} is determined. STMR is defined as the attenuation that must be inserted into the overall IRS (Recommendation P.48) to give an equivalent loudness to L_{MEST} when similarly making use of L_{MEHS} as a masking threshold.

It has been established that telephone sets having typical sidetone characteristics and the same value of STMR have similar subjective effects on the subscriber in his role as a talker.

Note – When the subscriber is listening, any room noise present may reach ERP through the paths a) and c) of § 3.1, and mask the received speech.

It is the high frequencies of local room noise which are most likely to mask the low-level consonants of a received signal. The STMR method strictly speaking cannot be used to rate listener sidetone. Nevertheless it has the effect of placing more control on L_{MEST} at frequencies greater than 1000 Hz. Control of these frequencies should ensure that high frequency room noise sidetone heard via the sidetone path L_{MEST} are also controlled. The whole subject of room noise sidetone, including the definition of a suitable rating method, is still under study under Question 9/XII.

For room noise, L_{MEST} may not have the same value as for talker sidetone, as the sensitivity of the microphone in a handset may not be the same for random incidence signals as for a point source close to the diaphragm (less than 5 cm). In general, room noise would arrive at the microphone at lower levels than the speech and this may result in different values of sensitivity where carbon microphones are present.

ANNEX A

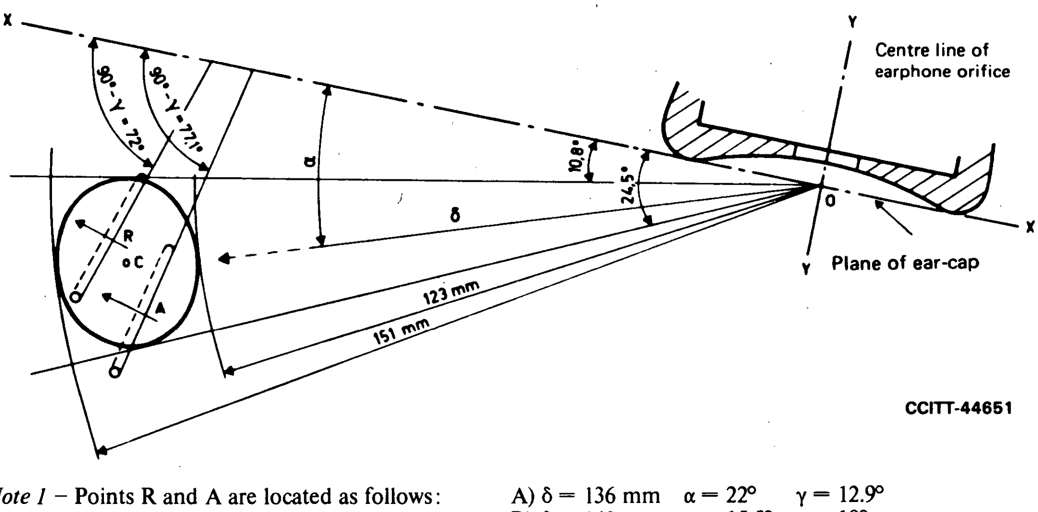
(to Recommendation P.76)

Definition of the speaking position for measuring loudness ratings of handset telephones

This annex describes the speaking position which should be used to measure the sensitivities of commercial telephone sets (by the method described in Recommendation P.64) for the determination of loudness ratings.

A.1 The definition of a speaking position falls into two parts: description of the relative positions of mouth opening and ear-canal opening on an *average* human head; and description of the angles that define the attitude in space of telephone handsets held to such a head. For any given telephone handset, these descriptions together describe the relative special disposition of the microphone opening and the talker's lips, and hence the direction in which speech sound waves arrive at the mouthpiece and the distance they have travelled from a *virtual point source.*

The relative positions of the centre of the opening of the mouth and that of the ear canal can be described in terms of a distance δ and an angle α as shown in Figure A-1/P.76. Point R in that figure represents the centre of a lip-ring located at the reference equivalent speaking position in accordance with Recommendation P.72. Position A is that used to determine ratings by the articulation method defined in Recommendation P.45. The approximately elliptical area encloses about 80% of the lip positions found in a sample of 3889 heads in the United States of America before 1930; averages of more recent results, including 4012 heads in the People's Republic of China cluster round the point A.



Note 1 – Points R and A are located as follows: A) δ = 136 mm α = 22° γ = 12.9°
R) δ = 140 mm α = 15.5° γ = 18°

Note 2 – Area shown includes about 80% of the sample of 3889 lip positions.

Note 3 – A straight line tangential to the boundary of the area shown excludes 5% of lip positions.

Note 4 – Solid line shows plane of lips; broken line shows lip-ring 1.6 mm thick.

FIGURE A-1/P.76

Location of lip position relative to opening of ear canal

A second angle is required to define the direction in which speech is emitted from the mouth into the mouthpiece of the microphone. In Recommendations P.45 and P.72 reference is made to an angle β, but this does not lie in the plane of symmetry of the handset, so it is more convenient to use an angle γ, which describes the vertical projection of the direction of speech on this plane.

A.2 The position of the centre of the lips as defined by A in Figure A-1/P.76 is used also to define the new speaking position, but two additional angles must also be defined, namely: the earphone rotational angle Φ and the handset rotational angle Θ. Earphone rotation is considered about an axis through the centre of the ear-cap (YY in Figure A-1/P.76); handset rotation is taken about a longitudinal axis of the handset (XX in Figure A-1/P.76); both angles are zero when the plane of symmetry of the handset is horizontal. Naturally, the earphone rotational angle is positive when the handle is pointed downwards away from the earphone and the handset rotational angle is positive in the sense that the upper part of the earphone is moved farther from the medial plane of the head.

The new speaking position is described by the following values for the distance and angles defined above:

$$\alpha = 22°, \quad \gamma = 12.9°, \quad \delta = 136 \text{ mm}, \quad \Phi = 39° \quad \text{and} \quad \Theta = 13°$$

The angle γ cannot be determined very precisely and is not convenient for use when setting up a handset for test in front of an artificial mouth. The semi-interaural distance ε may be used in its place, and for the new speaking position $\varepsilon = 77.8$ mm.

A.3 The foregoing description of the speaking position has shown the complexities of expressing the relative location of the ear reference point and the lip-ring centre, and the relative orientation of the earphone axis and the lip-ring axis. It is often more convenient, particularly in terms of constructing and setting up handset jigs, to express the position of the ear reference point[4)] and the direction of the earphone axis with respect to the lip-ring. This is easier since the axis of the lip-ring is horizontal as would be the axis of an associated artificial mouth.

A.4 Use has been made of a vector analysis method to determine the orthogonal coordinates of the handset ear-cap relative to the lip position when the handset is mounted in the LR guard ring position. It is necessary to define a set of cartesian axes with origin at the centre of the lips (or equivalent lip position of an artificial voice) as follows:

x-axis: horizontal axis of the mouth, with positive direction into the mouth;

y-axis: horizontal, perpendicular to the x-axis, with positive direction towards the side of the mouth on which the handset is held;

z-axis: vertical, with positive direction upwards.

The ear reference point is defined by the vector:

$$(86.5, 77.8, 70.5) \text{ mm}.$$

The handset is mounted so that the ear reference point lies at the intersection of the axis of the ear-cap with a plane in space on which the ear-cap can be considered to be resting. With some shapes of handset, this definition is not adequate; in such cases the position of the ear reference point relative to the handset should be clearly stated.

The orientation of the handset is defined by vectors normal to the plane of the ear-cap and the plane of symmetry of the handset:

Unit vector normal to plane of the ear-cap:

$$\pm (0.1441, -0.974, 0.1748)$$

Unit vector normal to plane of symmetry of the handset:

$$\pm (0.6519, -0.0394, -0.7572).$$

When using an artificial voice, the equivalent lip position must be used as the datum; this is not normally the same as the plane of the orifice of the artificial mouth.

---

4) See Recommendation P.64 for definition of ear reference point.

Alternatively, it can be convenient to define the speaking position in terms of axes with the origin at the ear reference point. These are defined as follows:

x-axis: axis of ear-cap with positive direction away from earphone;

y-axis: line of intersection of the plane of symmetry of the handset with the ear-cap plane, with positive direction towards the microphone;

z-axis: normal to the plane of symmetry of the handset with positive direction obliquely upwards.

The lip-ring centre is defined by the vector:

$$(50.95,\ 126.10,\ 0)\ \text{mm}.$$

The orientation of the lip-ring is defined by a unit vector along its axis:

$$\pm\ (0.1441,\ -0.7444,\ -0.6250)$$

and the orientation of the handset is defined by specifying the vertical by the unit vector:

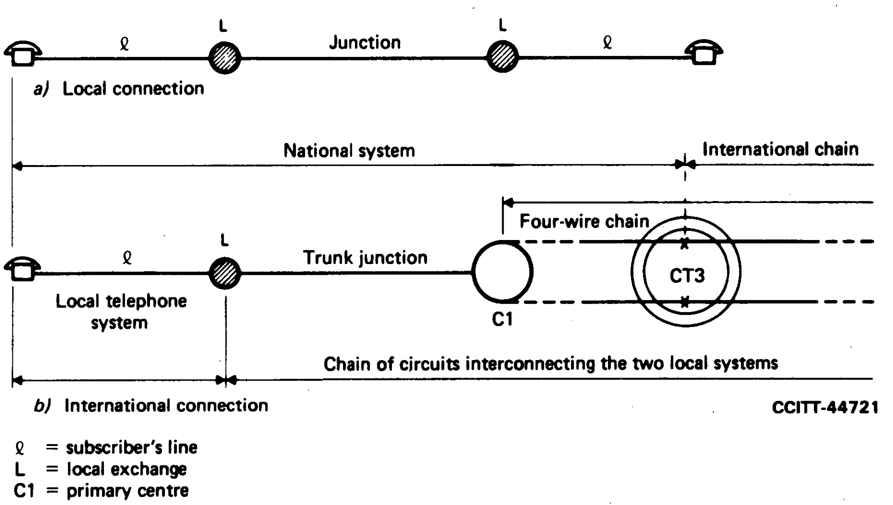
$$\pm\ (0.1748,\ -0.6293,\ -0.7572).$$

Note – The speaking position defined above differs from the special guard-ring position in the values of Φ (= 37°) and Θ (= 19°). It has been found that altering the handset position from the special guard-ring position to the loudness rating guard-ring position described above affects sensitivity measurements to a negligible extent.

ANNEX B

(to Recommendation P.76)

Explanations of certain terminology



Terminology applying to parts of a telephone connection according to Recommendations G.101 [3], G.111 [4], G.121 [5] and CCITT manuals.

Note – In the present Recommendation the word "junction" is used in a special sense to denote "chain of circuits interconnecting the two local systems" and the "junction attenuator" used in laboratory tests for determination of loudness ratings.

References

[1] CCITT – Question 19/XII, Contribution COM XII-No. 1, Study Period 1985-1988, Geneva, 1985.

[2] CCITT Manual *Transmission planning of switched telephone networks*, Chapter I, Annex 1, ITU, Geneva, 1976.

[3] CCITT Recommendation *The transmission plan*, Vol. III, Rec. G.101.

[4] CCITT Recommendation *Corrected reference equivalents (CREs) and loudness ratings (LRs) in an international connection*, Vol. III, Rec. G.111.

[5] CCITT Recommendation *Corrected reference equivalents (CREs) and loudness ratings (LRs) of national systems*, Vol. III, Rec. G.121.

METHOD FOR EVALUATION OF SERVICE FROM THE STANDPOINT OF SPEECH TRANSMISSION QUALITY

(Geneva, 1976; amended at Malaga-Torremolinos, 1984)

1 General

The CCITT recommends that Administrations make use of telephone users' surveys in the manner of Recommendation E.125 [1] as a means of measuring speech transmission quality on international calls.

Such surveys being call-related (in this instance to the last international call made) can be conducted either by the full use of the Recommendation E.125 [1] questionnaires (when other valuable information is obtained on users' difficulties, e.g. knowing how to make the call, difficulties in dialling or understanding tones, etc.) or by making use of those questions solely related to transmission quality which appear in Annex A (see Note 1).

2 Conduct of surveys

In order to make valid comparisons between data collected in different countries, Recommendation E.125 [1] should be strictly adhered to. Specifically the preamble to the Recommendation, the notes of intended use of the questionnaires and the precise order and wording of the questions should be rigidly followed. In some cases, however, an exception will be made and Question 10.0 will be replaced by the wording indicated in Annex B (detailed information is given in [3]; also see Note 3 below).

3 Treatment of results

To provide quantitative information suitable for comparisons, the subjective assessments (e.g. those obtained from Question 9.0 of Annex A) of excellent, good, fair or poor (see Note 2) should be accorded scores of 4, 3, 2 and 1, respectively and a mean opinion score (MOS) calculated for all associated responses. Similarly for all those experiencing difficulty (under Question 10.0 of Annex A or, alternatively, Question 10.0 of Annex B) a percentage of the total responses should be calculated. These two criteria of MOS and percentage difficulty are now internationally recognized and have been measured under many different laboratory simulated connections and practical situations.

The results can be classified in a number of ways, e.g. in terms of the call-destination countries or by nature/composition of the connection i.e. cable/satellite circuits, presence or otherwise of echo suppressors etc. Typical methods of presentation of the results are shown in [2], in this case for several countries. It should be noted that in all presentations it is essential to show the number of responses.

Note 1 – The evaluation of the transmission performance may be altered by difficulties in setting-up call. Hence the response to incomplete questionnaires should be considered with some reservation.

Note 2 – Among the reasons which lead to the limitation of users' opinions of transmission quality to four classes, i.e. excellent, good, fair and poor, is the following. The experience gained in human factor investigations has shown that when a question which requires a selection from several different classifications is posed in aural form, e.g. by face-to-face interview or by telephone as with Recommendation E.125 [1], the respondent is frequently unable to carry a clear mental separation of more than four categories. As a consequence, he is unable to draw on his short-term memory and judgement ability in a sufficiently precise manner to avoid confusion and gives an unreliable response. This restriction does not apply to other situations where a written presentation of the choices is used, in which case frequently five or more classes may be appropriate and shown to yield reliable responses.

Note 3 – An alternative version of Question 10.0 is described in Annex B. This alternative version has the advantage of simplifying the classification of responses to open end probes by experts, as well as increasing the sensitivity to some types of impairments such as delay. These advantages should be weighed against the additional interview time which may be required.

ANNEX A

(to Recommendation P.77)

Extract from the questionnaire annexed to Recommendation E.125

Reproduced below are the questions relating to transmission quality which appear in the questionnaire annexed to Recommendation E.125 [1].

The CCITT recommends that this Annex should be used when customers' general impressions of transmission performance are required.

9.0 . . . .

Which of these four words comes closest to describing the quality of the connection during conversation?

9.1	– *excellent*	1	☐	48
9.2	– *good*	2	☐	
9.3	– *fair*	3	☐	
9.4	– *poor*	4	☐	

10.0 *Did you or the person your were talking to have difficulty in talking or hearing over that connection?*

YES	NO	
1 ☐	2 ☐	49

(If answer is "yes") probe for nature of difficulty, but without suggesting possible types of difficulty, and copy down answers verbatim: e.g. *"Could you describe the difficulty a little more?"*

. . . .

. . . .

At end of interview, categorize the answers in terms of the items below:

10.1	– low volume	1	☐	50
10.2	– noise or hum	1	☐	51
10.3	– distortion	1	☐	52
10.4	– variations in level, cutting on and off	1	☐	53
10.5	– crosstalk	1	☐	54
10.6	– echo	1	☐	55
10.7	– complete cut off	1	☐	56
10.8	– other (specify) . . . . ←	1	☐	57

Note – Responses to Questions 10.1 to 10.8 are only obtained from customers who have expressed difficulty in Question 10.0.

ANNEX B

(to Recommendation P.77)

Alternative version for Question 10.0 of questionnaire annexed to Recommendation E.125

Studies at AT&T have shown that the verbatim responses describing impairments (requested after Question 10.0 of Annex A) are often too imprecisely worded to permit accurate classification by interviewers who are not experienced in transmission studies. A typical solution to this problem has been to convene a panel of experts to classify the responses, a method which may become impractical as the size and number of user reaction tests increases. This annex presents an alternative approach developed in 1976 and used widely since then by AT&T to measure customer's perceptions of transmission quality on domestic and international telephone connections. The approach involves a more complicated technique of probing for impairments which simplifies the ultimate task of classifying the responses. The alternative of Question 10.0 is reproduced below.

The CCITT recommends that this annex should be used for diagnostic purposes only.

10.0 *Did you have any difficulty talking or hearing over that connection?*

	YES	NO	
	1	2	49

Do not probe: If the person volunteers an explanation, write it down. . . . .

. . . .

On question 10.1-10.8, attempt to read entire text before respondent replies.

10.1 *Now I'd like to ask some specific questions about the connection.*

If the person has already described difficulty, add:

(In view of what you've already said, some of these may seem repetitious, but please bear with me). First, during your conversation on that call, did you hear your own voice echoing back, or did your own voice sound hollow to you?

10.1.1	– echo hollow (own voice)	1	50
10.1.2	– neither	2	
10.1.3	– don't remember/not sure	3	
10.1.4	– other (specify) . . . . ←	4	

10.2 *Did you hear another telephone conversation on the telephone network at the same time as your own?*

10.2.1	– other conversation	1	51
10.2.2	– no	2	
10.2.3	– don't remember/not sure	3	
10.2.4	– other (specify) . . . . ←	4	

10.3 *Now I'd like you to think about the voice of the person you were talking to. Was the volume of the voice low as if the person were faint and far away; did the voice fade in and out; or was the voice interrupted or chopped up at times?*

10.3.1	– low volume	1	52
10.3.2	– fading	2	
10.3.3	– chopping	3	
10.3.4	– none	4	
10.3.5	– don't remember/not sure	5	
10.3.6	– other (specify) . . . . ←	6	

10.4 *How did the voice of the person your were talking to sound to you: did it echo or sound hollow and tinny; or did it sound fuzzy or unnatural?*

10.4.1	– echo, hollow	1	53
10.4.2	– fuzzy, unnatural	2	
10.4.3	– none	3	
10.4.4	– don't remember/not sure	4	
10.4.5	– other (specify) . . . . ←	5	

10.5	*Now let me describe three kinds of noise. Tell me if you noticed any of these noises during your conversaiton: a rushing or hissing sound; a frying and/or sizzling, crackling sound; or a humming or buzzing sound?*		
10.5.1	– rushing, hissing	1 ☐	54
10.5.2	– frying and/or sizzling, cackling	2 ☐	
10.5.3	– humming, buzzing	3 ☐	
10.5.4	– none	4 ☐	
10.5.5	– don't remember/not sure	5 ☐	
10.5.6	– other (specify) . . . . ←	6 ☐	
10.6	*Now let me describe three more kind of noise. Tell me if you noticed any of these during your conversation: a clicking sound; a series of musical tones or beeps; or a continuous high-pitched tone?*		
10.6.1	– clicking	1 ☐	55
10.6.2	– tones or beeps	2 ☐	
10.6.3	– high-pitched tone	3 ☐	
10.6.4	– none	4 ☐	
10.6.5	– don't remember/not sure	5 ☐	
10.6.6	– other (specify) . . . . ←	6 ☐	
10.7	*Did the other person seem slow to respond, as if there were delay or time lag in the conversation?*		
10.7.1	– yes	1 ☐	56
10.7.2	– no	2 ☐	
10.7.3	– don't know	3 ☐	
10.7.4	– other (specify) . . . . ←	4 ☐	
10.8	*Would you please try to remember the background noise in the area around your telephone (e.g. noise from air-conditioning plant unit, road traffic, office equipment or other people talking) when you made the call. Which of the following categories best describes it?*		
10.8.1	– very noise	1 ☐	57
10.8.2	– noisy	2 ☐	
10.8.3	– quiet	3 ☐	
10.8.4	– very quiet	4 ☐	
10.8.5	– other (specify) . . . . ←	5 ☐	
10.9	*Which of the categories listed below best describes the extent to which you heard your own voice through your telephone when you were talking?*		
10.9.1	– could not hear it	1 ☐	58
10.9.2	– could hear it now that you have drawn my attention to it	2 ☐	
10.9.3	– did notice it – not loud	3 ☐	
10.9.4	– did notice it – loud	4 ☐	
10.9.5	– other (specify) . . . . ←	5 ☐	

10.10 *Was there anything else about the connection you'd like to mention?*

Yes – What? (Write in)

. . . .

. . . .

. . . .

Coding instructions:

	YES	NO	
– is there a written comment?	1 ☐	2 ☐	59
– does the comment apply to this call?	1 ☐	2 ☐	60
– does it mention an impairment?	1 ☐	2 ☐	61
– has it been mentioned already?	1 ☐	2 ☐	62
– other (specify)	1 ☐	2 ☐	63

Note – The responses to the specific questions are only obtained from customers who have expressed difficulty in Question 10.0. This may prevent the diagnosis of certain impairments (the bias produced is more serious than that mentionned at the end of Annex A).

References

[1] CCITT Recommendation *Inquiries among users of the international telephone service*, Vol. II, Rec. E.125.

[2] CCITT – Question 2/XII, Annex 2, Contribution COM XII-No. 1, Study Period 1977-1980, Geneva, 1977.

[3] CCITT – Question 2/XII, Annex, Contribution COM XII-No. 171, Study Period 1977-1980, Geneva, August 1979.

Recommendation P.78

SUBJECTIVE TESTING METHOD FOR DETERMINATION OF LOUDNESS RATINGS IN ACCORDANCE WITH RECOMMENDATION P.76

(amended at Malaga-Torremolinos, 1984)

Preface

This Recommendation describes a subjective testing method which has been found suitable for its purpose by use in the CCITT Laboratory. It can also be used in other laboratories. Provided that the Intermediate Reference System (IRS) used complies with the requirements of Recommendation P.48 and that other requirements given in Recommendation P.76 are adhered to, the loudness ratings obtained by using the method given in the present Recommendation can be used for forwarding the study of Question 19/XII [1] (Recommended values of loudness rating). The present Recommendation, together with Recommendations P.76 and P.48, provides a definition of loudness ratings which can be used for planning.

Summary

This Recommendation contains the essential particulars for defining the method for determining loudness ratings in accordance with Recommendation P.76 when use is made of subjects performing equal loudness balances. Details are included concerning the balancing method, choice of subjects, speech material, design of experiment, method of analysis and presentation of results.

1 Introduction

To compare the calculation of loudness ratings method (Recommendation P.79) a defined method of subjectively determining loudness ratings is required. This Recommendation deals with all aspects of a test from selection of operators to the method of analysis and finally presentation of results.

2 General

In the subjective comparisons, the Fundamental Reference System (FRS) is used (although other reference systems are permissible) as the datum for comparing the following speech paths:

a) *Path 0* – The fundamental reference system always provides the speech path against which each of the others is balanced. NOSFER set at 25 dB is used.

b) *Path 1* – The send end of the test ("unknown") local telephone circuit connected through the test ("unknown") junction and an adjustable attenuator to the receive end of the test ("unknown") local telephone circuit. The adjustable attenuator must be inserted in such a manner that the impedance relationships between the three parts of the connection (send end, junction and receive end) are not disturbed.

c) *Path 2* – The send end of the intermediate reference system connected through an adjustable attenuator to the receive end of the intermediate reference system.

d) *Path 3* – The send end of the test ("unknown") local telephone circuit connected through an adjustable attenuator to the receive end of the IRS.

e) *Path 4* – The send end of the IRS connected through an adjustable attenuator to the receive end of the test ("unknown") local telephone circuit.

f) *Path 5* – The send end of the IRS connected through the test ("unknown") junction and an adjustable attenuator to the receive end of the IRS. The adjustable attenuator must be inserted in such a manner that the impedance relationships between the three parts of the connection (send end, junction and receive end) are not disturbed.

In these subjective comparisons, the junction of the fundamental reference system is fixed, i.e. the level of speech sounds received via the fundamental reference system is kept constant, the loudness balance being obtained by the so-called "margin" method, and the balance attenuator being that inserted in the telephone (or IRS) path being tested.

The speaking position used with both the IRS and the test telephone sets should be as defined in Annex A to Recommendation P.76.

Figure 1/P.78 shows the composition of the telephone paths to be compared. The balances should be conducted using the vocal level defined in Recommendation P.72.

The loudness ratings relative to the IRS as defined in Recommendation P.76 are:

$$\mathrm{OLR} = x_2 - x_1$$

$$\mathrm{SLR} = x_2 - x_3$$

$$\mathrm{RLR} = x_2 - x_4$$

$$\mathrm{JLR} = x_2 - x_5$$

It is not necessary to include all the paths indicated above in every experiment. Paths 0 and 2 are essential but addition of only 3 and 4 is sufficient to determine sending and receiving loudness ratings of a local telephone circuit. Paths 0, 2 and 5 are required to determine a junction loudness rating. Path 1 is usually required only when it is derived to verify additivity of loudness ratings, namely that:

$$\mathrm{OLR} = \mathrm{SLR} + \mathrm{JLR} + \mathrm{RLR}$$

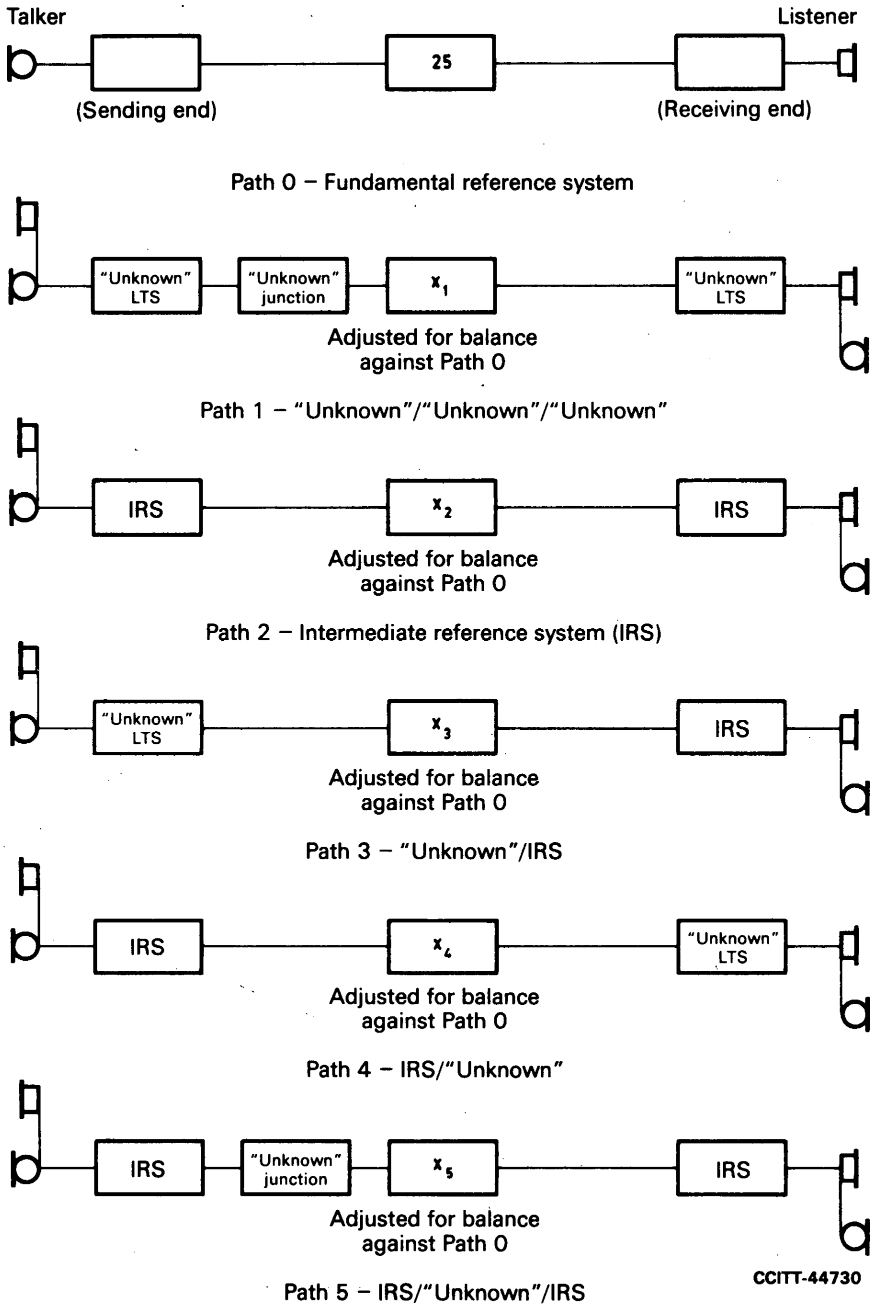


FIGURE 1/P.78

Arrangement of paths for subjective method of determination of loudness ratings

3 Experiment design

To have confidence in results requires the correct testing procedures to be followed, coupled with the correct experiment design. The procedure should be prepared such that no ambiguity can exist.

The following points must be considered in the design:

a) The experiment should be designed in such a way that all uncontrolled influences operate at random, e.g. slight day-to-day drift of subjects and/or measuring equipment;

b) If more balances are required than can be comfortably completed in one day, then the experiment must be designed such that equal numbers of each type of system are completed each day;

c) The operators who start a test should always be the same throughout the test [2];

d) A minimum of 12 operator-pair combinations is suggested with a maximum of 20. Twelve operator-pair combinations can be arrived at from two crews of 3 (see Table 1a/P.78) or one crew of 4 and 18 operator-pair combinations can be arrived at from one crew of 6 (see Table 1b/P.78) and 20 operator-pair combinations from one crew of 5 (see Table 2a/P.78).

Note – One crew of 6 giving 30 operator-pair combinations (see Table 2b/P.78) produces a larger test for only slightly more precision than the previously mentioned crew sizes;

TABLE 1a/P.78

Twelve operator-pair combinations from two crews of three, known as 3/6 operator method

Operator (talker) \ Operator (listener)	A	B	C	D	E	F
A		X	X			
B	X		X			
C	X	X				
D					X	X
E				X		X
F				X	X	

TABLE 1b/P.78

Eighteen operator-pair combinations from one crew of six, known as 3/6′ operator method

Operator (talker) \ Operator (listener)	A	B	C	D	E	F
A				X	X	X
B				X	X	X
C				X	X	X
D	X	X	X			
E	X	X	X			
F	X	X	X			

TABLE 2a/P.78

Twenty operator-pair combinations from one crew of five, known as 5/5 operator method

Operator (talker) \ Operator (listener)	A	B	C	D	E
A		X	X	X	X
B	X		X	X	X
C	X	X		X	X
D	X	X	X		X
E	X	X	X	X	

TABLE 2b/P.78

Thirty operator-pair combinations from one crew of six, known as 6/6 operator method

Operator (talker) \ Operator (listener)	A	B	C	D	E	F
A		X	X	X	X	X
B	X		X	X	X	X
C	X	X		X	X	X
D	X	X	X		X	X
E	X	X	X	X		X
F	X	X	X	X	X	

e) When using two crews of 3, one can use both crews interleaved but it is generally more practical to separate the crews and use test crew 1 before crew 2. Members should not be used in both crews as it causes a bias and complicates the analysis;

f) All operator-pair combinations should be tested in rotation, where practical, such that each operator takes a turn as talker, then listener and then has a break;

g) The design of the experiment should eliminate any effect that could be attributed to the order of presentation. That is to say that all systems should be in a randomized order. To illustrate this point two examples are as follows:

Example 1

If one type of loudness rating is required, with a given combination of telephone set and circuit condition, then the experiment design must allow for any effect associated with order of presentation for each operator-pair combination. An example is shown in Table 3/P.78.

TABLE 3/P.78

Example to illustrate the elimination of order of presentation effect for one type of loudness rating

Operator-pairs	Talker Listener	A B	B C	C A
Circuits	α α' β β'	3 2 1 4	1 3 4 2	2 4 3 1

Where α = path 0 presented before path 2

α' = path 2 presented before path 0

β = path 0 presented before path 3

β' = path 3 presented before path 0

Note – When it is proven that there is no difference for a given test crew and set of test conditions, the distinction between the order of path presentation can be eliminated.

Example 2

Now, if more than one type of loudness rating is made or more than one telephone set is used, then there need only be one balance of path 2 against path 0 and vice-versa per operator-pair combination for any experiment, but this must be randomized within the experiment. An example is shown in Table 4/P.78.

TABLE 4/P.78

Example to illustrate the elimination of order of presentation effect for two types of loudness ratings

Operator-pairs	Talker Listener	A B	B C	C A
Circuits	α α' β_1 β_1' β_2 β_2'	3 5 1 6 2 4	1 4 2 5 6 3	2 6 5 3 4 1

β_1, β_1' = have, for example, 0 km of subscriber's cable

β_2, β_2' = have, for example, 6 km of subscriber's cable

Some experiment designs can be found in Annex A.

4 Selection of crew members and speech material

Requirements for the selection of crew members including audiometric testing of subjects, as well as the speech material used by the crew for subjective tests, can be found in Annex B.

5 Calibration of the IRS

It is most important that the calibration of the IRS is made before every test so that any small change in SLR and RLR can either be compensated for in the results or the sensitivity can be changed before the test. It is good experimental practice to check the sensitivity of the IRS after each experiment. The specification of the IRS is found in Recommendation P.48 and the description of the calibration procedure is found in Recommendation P.64. The results of the calibration are used to determine the corrections to the subjective balance results (see § 9).

6 Circuit arrangements

Figure 2a)/P.78 shows a typical circuit layout for the measurement of SLR and RLR. Figures 2b)/P.78 and 2c)/P.78 show layouts for the measurement of JLR and OLR respectively. There is no reason if the experimenter so wished, why all four types of loudness rating cannot be tested in the same experiment. This, however, would require extremely intricate switching arrangements.

In Figures 2a)/P.78, 2b)/P.78 and 2c)/P.78 the 600 ohm on the second position of switch S1 allows the correct speech level to be set when Path 0 is presented after Path 1/2/3/4/5 (see Figure 1/P.78). This switch should be of the nonlocking type and should be returned to the normal position as soon as the talker has attained the correct speech level.

In order to reduce the effect of sidetone on the talker's vocal level during sending and overall determinations, the acoustic sidetone path of handset telephones should be disabled. This can be accomplished by placing the earphone in another identical handset and the electrical connections made to the correct terminals on the telephone transmission circuit. The earphone can then be sealed to an IEC/CCITT artificial ear to give the correct acoustic loading. A simpler method, used by the Australian Post Office, is to seal the earphone by means of heavy tape. Although this might not have the correct acoustic loading, in practice it has been found to have a negligible effect.

If the microphone is of the carbon-granule type, then before each balance the conditioning procedure according to Recommendation P.75 should be used.

In Figures 1/P.78 and 2/P.78 the fundamental reference system, NOSFER, has been shown but other types such as SETED and METRE-AIR-PATH could be used.

7 Recording of information

It is essential that as much information of any test should be recorded, in such a way that at any time in the future, the information can be retrieved.

7.1 *Details of the test*

Each test should always include the following information:

a) test No. – this should be unique so that one test cannot be confused with another;

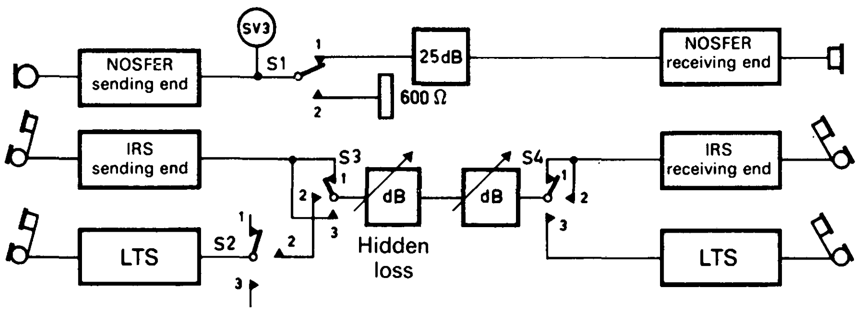
b) date;

c) title – a brief description of the test;

d) circuit conditions – describe each individual path;

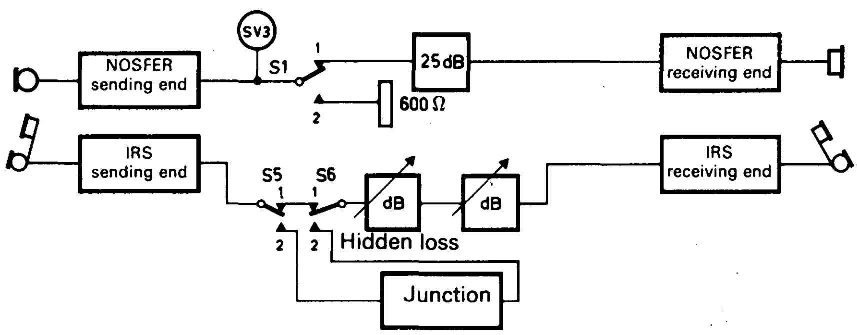
e) diagram to show switching arrangement;

f) crew members – name each operator and assign a code, as for example in Table 5/P.78. Then each operator-pair combination can be denoted by a code e.g. A-B.



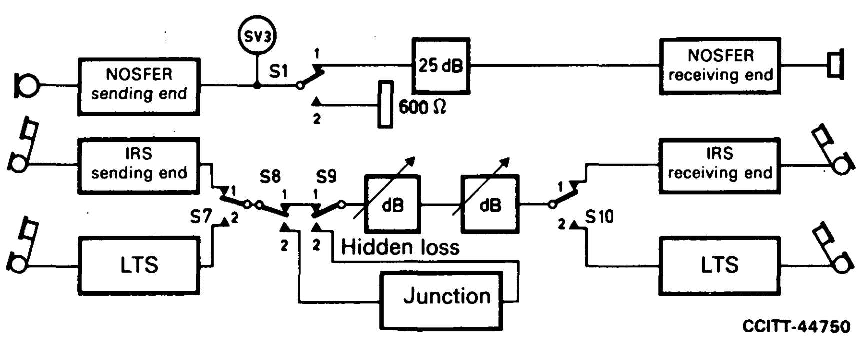
Note – S1 is a nonlocking switch. S2, S3 and S4 are all ganged.

a) Switching diagram for the measurement of SLR and RLR



Note – S1 is a nonlocking switch. S5 and S6 are ganged.

b) Switching diagram for the measurement of JLR



Note – S1 is a nonlocking switch. S7, S8, S9 and S10 are all ganged.

c) Switching diagram for the measurement of OLR

FIGURE 2/P.78

TABLE 5/P.78

Crew members	
Code	Operator
A B C D E F	

7.2 *Individual balances*

These should always include the "hidden loss" attenuation, the "balance" attenuation and finally the result of the comparison, e.g.

$$R = H + B$$

where

R = result

H = hidden loss

B = balance

8 Analysis

For any experiment most information can be obtained from an analysis of variance. However, sufficient useful information can be derived using the mean, standard deviation. The method of calculation of these parameters can be found in Annex C.

9 Presentation of results

The results of the test should be presented such that the important information can be displayed on one form. An example of such a form is shown in Table 6/P.78.

Note – In Tables 6/P.78 to 8/P.78 corrected mean = mean + correction.

Worked examples of the use of the form shown in Table 6/P.78 are shown in Tables 7/P.78 and 8/P.78. The form has been modified to allow SLR and RLR determinations to be made on a local telephone system including two line lengths. Table 7/P.78 shows the SLR results and Table 8/P.78 the RLR results.

TABLE 6/P.78

Presentation of results

Frequency Hz	IRS sending sensitivity (dBV/Pa)	IRS receiving sensitivity [a] (dBPa/V)	Operator-pair	x_0 (dB)	x_2 (dB)	x'_2 (dB)	x_3 (dB)	x'_3 (dB)	x_2 (dB)	x'_2 (dB)	x_4 (dB)	x'_4 (dB)	SLR (dB)	SLR′ (dB)	RLR (dB)	RLR′ (dB)	$\frac{\text{SLR} + \text{SLR}'}{2}$ (dB)	$\frac{\text{RLR} + \text{RLR}'}{2}$ (dB)
100																		
125																		
160																		
200																		
250																		
315																		
400																		
500																		
630																		
800																		
1000																		
1250																		
1600																		
2000																		
2500																		
3150																		
4000																		
5000																		
6300																		
8000																		
Calculated LR of IRS			Mean: dB															
			Std. dev.: dB															
			95% confidence limits: dB															
				Corrected mean: dB														

[a] Artificial ear conforming to Recommendation P.51.

TABLE 7/P.78

Example to illustrate the use of the form shown in Table 6/P.78 for the determination of SLR

Frequency Hz	IRS sending sensitivity (dBV/Pa)	IRS receiving sensitivity[a] (dBPa/V)	Operator-pair	x_0 (dB)	x_2 (dB)	x_2' (dB)	x_3 (0) (dB)	x_3' (0) (dB)	x_2 (dB)	x_2' (dB)	x_3 (L) (dB)	x_3' (L) (dB)	SLR (0) (dB)	SLR' (0) (dB)	SLR (L) (dB)	SLR' (L) (dB)	$\frac{SLR + SLR'}{2}$ (0) (dB)	$\frac{SLR + SLR'}{2}$ (L) (dB)
100			A-C	25	14	15	13	14			12	10	1	1	2	5	1.0	3.5
125			D-A	25	13	13	8	10			10	11	5	3	3	2	4.0	2.5
160			C-D	25	10	11	7	11			10	11	3	0	0	0	1.5	0.0
200	−19.7		D-C	25	12	14	11	10			10	11	1	4	2	3	2.5	2.5
250	−15.3		C-A	25	17	17	17	13			12	14	0	4	5	3	2.0	4.0
315	−12.2		A-D	25	10	12	8	10			10	8	2	2	0	4	2.0	2.0
400	− 9.6		F-E	25	11	11	7	7			5	4	4	4	6	7	4.0	6.5
500	− 8.0		B-F	25	10	11	6	8			5	7	4	3	5	4	3.5	4.5
630	− 6.7		E-B	25	13	12	8	13			8	9	5	−1	5	3	2.0	4.0
800	− 5.9		E-F	25	13	13	12	11			12	8	1	2	1	5	1.5	3.0
1000	− 5.6		F-B	25	12	13	9	5			5	6	3	8	7	7	5.5	7.0
1250	− 4.2		B-E	25	12	13	9	9			9	10	3	4	3	3	3.5	3.0
1600	− 1.2																	
2000	0																	
2500	+ 1.0																	
3150	+ 0.3																	
4000	−36.5																	
5000																		
6300																		
8000																		
Calculated LR of IRS	1.09		Mean: dB	25	12.25	12.92	9.58	10.08			9.00	9.08	2.67	2.83	3.25	3.83	2.75	3.54
			Std. dev.: dB	0	1.92	1.71	3.01	2.50			2.58	2.56	1.60	2.23	2.24	1.91	1.28	1.82
			95% confidence limits: dB	0	1.22	1.08	1.91	1.59			1.64	1.63	1.02	1.42	1.42	1.21	0.81	1.16
				Corrected mean: dB									3.76	3.92	4.34	4.92	3.84	4.63

[a] Artificial ear conforming to Recommendation P.51.

TABLE 8/P.78

Example to illustrate the use of the form shown in Table 6/P.78 for the determination of RLR

Frequency (Hz)	IRS sending sensitivity (dBV/Pa)	IRS receiving sensitivity [a] (dBPa/V)	Operator-pair	x_0 (dB)	x_2 (dB)	x'_2 (dB)	x_4 (0) (dB)	x'_4 (0) (dB)	x_2 (dB)	x'_2 (dB)	x_4 (L) (dB)	x'_4 (L) (dB)	RLR (0) (dB)	RLR′ (0) (dB)	RLR (L) (dB)	RLR′ (L) (dB)	$\frac{\text{RLR}+\text{RLR}'}{2}$ (0) (dB)	$\frac{\text{RLR}+\text{RLR}'}{2}$ (L) (dB)
100			C-B	25	10	11	20	20			15	13	−10	− 9	−5	−2	− 9.5	−3.5
125			B-E	25	15	9	19	21			13	13	− 4	−12	2	−4	− 8.0	−1.0
160			B-C	25	14	17	23	23			17	14	− 9	− 6	−3	3	− 7.5	0.0
200		− 3.8	E-B	25	11	10	19	19			13	15	− 8	− 9	−2	−5	− 8.5	−3.5
250		2.0	C-E	25	8	11	16	18			14	15	− 8	− 7	−6	−4	− 7.5	−5.0
315		6.6	E-C	25	13	13	18	18			13	16	− 5	− 5	0	−3	− 5.0	−1.5
400		9.8	D-F	25	8	9	13	13			12	9	− 5	− 4	−4	0	− 4.5	−2.0
500		11.2	F-A	25	14	14	22	21			17	16	− 8	− 7	−3	−2	− 7.5	−2.5
630		12.1	D-A	25	12	10	18	18			13	13	− 6	− 8	−1	−3	− 7.0	−2.0
800		12.8	A-D	25	12	8	21	19			12	11	− 9	−11	0	−3	−10.0	−1.5
1000		13.4	A-F	25	10	9	15	18			9	9	− 5	− 9	1	0	− 7.0	0.5
1250		13.8	F-D	25	11	9	19	16			10	10	− 8	− 7	1	−1	− 7.5	0.0
1600		14.0																
2000		13.2																
2500		11.0																
3150		10.4																
4000		−15.8																
5000																		
6300																		
8000																		
Calculated LR of IRS		− 0.16	Mean: dB	25	11.50	10.83	18.58	18.67			13.17	12.83	− 7.08	− 7.83	−1.67	−2.00	− 7.46	−1.83
			Std. dev.: dB	0	2.18	2.51	2.75	2.46			2.30	2.44	1.89	2.23	2.46	2.12	1.51	1.56
			95% confidence limits: dB	0	1.38	1.59	1.75	1.56			1.46	1.55	1.20	1.42	1.56	1.35	0.96	0.99
				Corrected mean: dB									− 7.24	− 7.99	−1.83	−2.16	− 7.62	−1.99

[a] Artificial ear conforming to Recommendation P.51.

ANNEX A

(to Recommendation P.78)

Examples of experiment designs

Tables A-2/P.78, A-3/P.78 and A-4/P.78, give typical designs for different crew sizes.

As an example, using Table A-2/P.78, the order of balances is as given in Table A-1/P.78.

TABLE A-1/P.78

Balance No.	Operator-pair	Circuit
1	BA	β_1
2	CB	α
3	DC	β_2
ǀ		ǀ
ǀ		ǀ
ǀ		ǀ
13	BA	β'_1
14	CB	β_1
15	DC	β'_2
ǀ		ǀ
ǀ		ǀ
ǀ		ǀ
25	BA	β_2
26	CB	β'_2
27	DC	α
ǀ		ǀ
ǀ		ǀ
ǀ		ǀ
71	AC	β_1
72	DA	α'

The operator-pairs in rotation do all balances in numerical order starting with "1" and finishing with "6".

Similar tables can be drawn up for a test requiring only one type of loudness rating where only 4 circuits are required e.g. α, α', β and β' for a SLR test, where numbers 1, 2, 3 and 4 would be assigned respectively in the experiment design.

For a test involving more circuits the same principles can be followed assigning as many numbers as there are circuits.

It may be necessary to improve the validity of results and a replication of the same experiment design using the same operator-pairs can be made.

TABLE A-2/P.78

Design for one crew of 4 or two crews of 3

One crew of 4 Operator-pairs	Talker	B	C	D	A	C	B	A	B	C	D	A	D
	Listener	A	B	C	D	A	D	B	C	D	B	C	A
Two crews of 3	Talker	B	C	A	C	B	A	E	F	D	F	E	D
	Listener	A	B	C	A	C	B	D	E	F	D	F	E
Circuits	α	4	1	3	2	6	5	3	6	1	5	4	2
	α'	6	5	4	3	2	1	2	4	5	3	1	6
	β_1	1	2	5	6	3	4	5	3	2	1	6	4
	β'_1	2	4	6	5	1	3	4	2	3	6	5	1
	β_2	3	6	1	4	5	2	6	1	4	2	3	5
	β'_2	5	3	2	1	4	6	1	5	6	4	2	3

TABLE A-3/P.78

Design for one crew of 6

Operator-pairs	Talker Listener	D A	E B	F C	E A	F B	D C	F A	D B	E C	A D	B E	C F	A E	B F	C D	A F	B D	C E
Circuits	α	4	1	3	2	6	5	3	6	1	5	4	2	1	2	6	3	5	4
	α'	6	5	4	3	2	1	2	4	5	3	1	6	5	4	1	6	2	3
	β_1	1	2	5	6	3	4	5	3	2	1	6	4	4	6	2	1	3	5
	β'_1	2	4	6	5	1	3	4	2	3	6	5	1	3	1	4	5	6	2
	β_2	3	6	1	4	5	2	6	1	4	2	3	5	6	5	3	2	4	1
	β'_2	5	3	2	1	4	6	1	5	6	4	2	3	2	3	5	4	1	6

TABLE A-4/P.78

Design for one crew of 5

Operator-pairs	Talker Listener	B A	C B	D C	E D	A E	C A	E C	B E	D B	A D	D A	B D	E B	C E	A C	E A	D E	C D	B C	A B
Circuits	α	4	1	3	2	6	5	3	6	1	5	4	2	1	2	6	3	5	4	1	6
	α'	6	5	4	3	2	1	2	4	5	3	1	6	5	4	1	6	2	3	2	5
	β_1	1	2	5	6	3	4	5	3	2	1	6	4	4	6	2	1	3	5	3	4
	β'_1	2	4	6	5	1	3	4	2	3	6	5	1	3	1	4	5	6	2	4	3
	β_2	3	6	1	4	5	2	6	1	4	2	3	5	6	5	3	2	4	1	5	2
	β'_2	5	3	2	1	4	6	1	5	6	4	2	3	2	3	5	4	1	6	6	1

ANNEX B

(to Recommendation P.78)

Selection of crew members, audiometric testing of subjects and speech material

B.1 *Crew members*

The crew should, wherever possible contain an equal number of both men and women.

The following points are a guide for selection:

a) Good hearing – no operator should exceed a hearing loss of a 15 dB at all frequencies up to and including 4 kHz and no more than 25 dB at 8 kHz. This is shown in Figure B-1/P.78. If it is intended that contra-lateral balances are required and this necessitates the use of both ears, then the maximum difference between ears should be ± 10 dB at all frequencies. An example of an audiometric testing procedure of subjects is presented below in § B.2;

b) Clear speech – each operator should be free from obvious speech impediments;

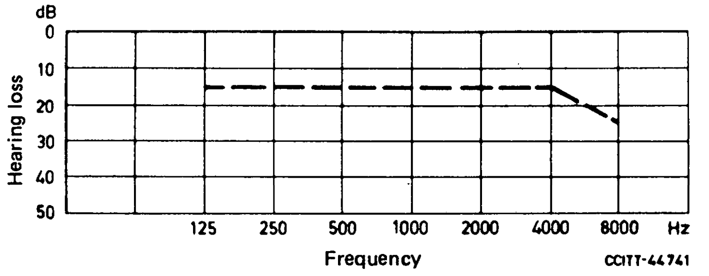
c) The operator should be able to work harmoniously with other people;

d) The operator should be able to make simple arithmetical calculations;

e) The operator should be able to talk at a constant level, with the aid of a meter, after sufficient training;

f) The operator must not suffer from claustrophobia as each operator must, during the test, spend a certain amount of short-term solitary confinement;

g) Regular checks should be made to determine the performance of each operator as both a talker and as a listener to disclose any unusual changes. A full description can be found in Reference [3].



Note – Nominal hearing is at 0 dB.

FIGURE B-1/P.78

Mask of maximum loss of hearing of subjects

B.2 *Audiometric testing of subjects – simple screening procedure [4]*

B.2.1 Visual examination of ears for wax, ask if subject has a cold, sinusitus or any other abnormality.

B.2.2 *Frequencies of test*

125, 250, 500, 1000, 2000, 3000, 4000, 6000, 8000 Hz.

B.2.3 *Example of presentation*

1000, 2000, 3000, 4000, 6000, 8000, 125, 250, 500, 1000 Hz.

Note – It is common for the second reading at 1000 Hz to be lower than the first.

Follow the above sequence for one ear, then repeat for the other ear.

B.2.4 *Example of finding threshold:*

Start above estimated threshold (say 20 dB hearing loss), approach in 10 dB steps until inaudible (no response). Return to last audible level and descend in 5 dB steps. Then approach this threshold from below in 5 dB steps. Signal duration 1 to 2 seconds.

Threshold is that value at which two equal responses are obtained from four successive stimuli.

B.2.5 *Room noise* [5]

Using supra-aural type headsets the maximum permissible levels in the test room are given in Table B-1/P.78.

If circum-aural type headsets are used then it is normally permissible to allow higher levels of noise.

TABLE B-1/P.78

Octave band	Sound pressure level (dB)
125	22.0
250	16.0
500	18.0
1000	26.0
2000	36.0
3000	39.5
4000	38.5
6000	40.0
8000	34.5

B.3 *Speech material*

The test phrase or phrases can be either a "nonsense" or "meaningful" phrase. Examples are:

a) Joe took father's shoe bench out,

b) Paris – Bordeaux – Le Mans – Saint-Leu – Léon – Loudun.

Due consideration should be given to the following points:

i) The ability of each operator to pronounce the chosen test phrase or phrases fluently and at a steady speech level. The sound structure of the native languages of the operators has therefore a bearing on the choice of test phrase or phrases;

ii) The phrase or phrases should be chosen so that the agreed measurement method to control the speech level (i.e. deflection of meter) can give a consistent and readily appreciated indication of vocal level.

ANNEX C

(to Recommendation P.78)

Simplified statistical analysis

C.1 *Mean*

The mean is obtained by using the following formula:

$$\bar{x} = \frac{\Sigma x}{n}$$

C.2 *Standard deviation*

It cannot be assumed that the operators are a sample drawn at random from a population and that the operator-pair combinations are independent of each other. Under these circumstances the standard deviation must be of the sample and not an estimate of a population.

The formula for the standard deviation is:

$$\sigma = \sqrt{\frac{\Sigma (x - \bar{x})^2}{n}}$$

C.3 A more detailed statistical analysis is possible to calculate confidence intervals as explained in § 1.3.4 of the manual *Telephonometry* [6]. The confidence interval is governed by the dispersion between the crew members, the number of crew members and the arrangement of the experimental design. Typical values in a well-conducted test are ± 5 dB for the arrangements shown in Table 1a/P.78, ± 4 dB for Table 1b/P.78, ± 3 dB for Table 2a/P.78 and ± 2 dB for Table 2b/P.78.

References

[1] CCITT – Question 19/XII, Contribution COM XII-No. 1, Study Period 1985-1988, Geneva, 1985.

[2] *The design and analysis of loudness efficacy measurements*, Red Book, Vol. V, Annex 7, ITU, Geneva, 1962.

[3] *Extract from a study of the differences between results for individual crew members in loudness balance tests*, Red Book, Vol. V, Annex 6, p. 214, ITU, Geneva, 1962.

[4] BURNS (W.): Noise and man, *Murray*, pp. 70-80, 1968

[5] *Ibid.*, pp. 298-300.

[6] CCITT Manual *Telephonometry* (to be published in 1985).

CALCULATION OF LOUDNESS RATINGS

(Geneva, 1980; amended at Malaga-Torremolinos, 1984)

Preface

The method given in this Recommendation is provisional for the reason stated in detail below, that its applicability to local telephone systems containing carbon microphones has not been confirmed beyond doubt. Nevertheless, Administrations who are studying Question 19/XII [1] (recommended values of loudness ratings) may use the method given here for studies relating to new types of telephone set which do not contain carbon microphones [1)].

Administrations are also encouraged to use the method in studying Question 7/XII [2] for expressing loudness loss on a common scale in quality evaluation experiments.

The Recommendation describes a calculation method which gives results in good agreement with those from subjective tests by the CCITT Laboratory [2)] (see Recommendation P.78) using local telephone systems having noncarbon microphones. For such local telephone systems, the methods given in Recommendation P.64 should be used to determine the values of sending and receiving sensitivities.

When local telephone systems containing carbon microphones are to be considered, the results obtained so far from tests in the CCITT Laboratory suggest that the method given in the present Recommendation can still be used provided a suitable method is used to obtain the sending sensitivities. Various measuring methods are being considered for this purpose and are listed in Annex B to Recommendation P.64. The results of extensive tests by the CCITT Laboratory using the "upper-envelope" method show that this method gives good results for some types of carbon microphone. The matter is being studied under Question 8/XII [3] (Measuring the efficiency of a microphone or a receiver).

1 Introduction

Loudness ratings according to the principles described in Recommendation P.76 can be determined without recourse to subjective tests provided that all the following conditions are fulfilled:

a) a theoretical model is available having suitable structure;

b) the appropriate values of the essential parameters of the model are known;

c) the sending and receiving sensitivities of the intermediate reference systems are known;

d) the sending and receiving sensitivities of the "unknown" local telephone systems and the insertion loss of the intervening chain of circuits are known.

The methods of determining sending and receiving sensitivities using an artificial mouth and artificial ear are defined in Recommendation P.64. The characteristics of the intermediate reference system determined according to the same methods are given in Recommendation P.48. The receiving sensitivities obtained using the artificial ear now mentioned in Recommendation P.64 are not directly suitable for use in calculating loudness ratings but must be corrected to allow for differences between sound pressures in real ears under conditions of telephone conversations and those measured by the artificial ear. Information concerning this correction (L_E) is given in § 6.

---

1) The method may also be used for determining receiving loudness ratings irrespective of whether the telephone set contains a carbon microphone or not.

2) The calculation method described in the Recommendation is based on weighting factors which have been determined for the 20 ISO-preferred frequencies. General applicability of the method would be improved if smoothed analytic expressions were also available for use with other sets of frequencies.

2 Definitions and symbols concerning sound pressures, sensitivities and transmission losses

Definitions and symbols used in the subsequent description of theoretical principles are listed below. Figure 1/P.79 illustrates these.

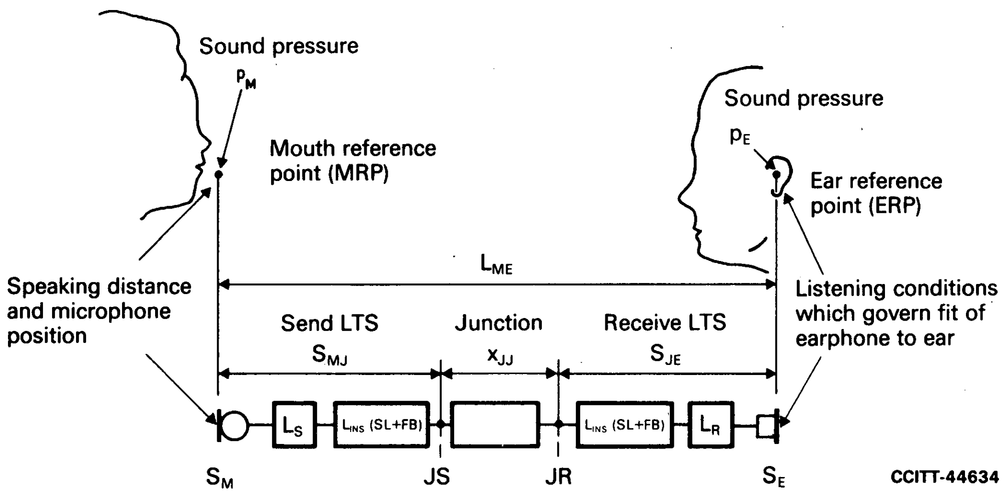


FIGURE 1/P.79

Factors effecting loudness of received speech

2.1 *Concerning talking*

These definitions and symbols characterize the situation where a subject is talking and they include his physical relationship to the telephone or reference connection.

MRP	Point defining the mouth reference point; MRP is at a defined location relative to the talker's lips. (See Recommendation P.64.)
p_M	Sound pressure at MRP[3] in absence of any obstruction.
B'_S	Spectrum density (long-term mean pressure)[4] of speech referred to a MRP in dB relative to 20 μPa in a bandwidth of 1 Hz.
VL	Vocal level, i.e. speech sound pressure (long-term rms while talker is active) level of talker at the MRP; usually referred to a reference vocal level as datum.
SP	Speaking position, i.e. the relative location of the microphone of the telephone or reference system and the lips of the talker.

2.2 *Concerning listening*

These definitions and symbols characterize the situation where a subject is listening and they include his physical relationship to the telephone or reference connection:

ERP	Point defining the ear reference point (see Recommendation P.64).
p_E	Sound pressure at ERP.
β_0	Hearing threshold for pure tones referred to an ERP in dB relative to 20 μPa.
K	A number, related to Fletcher's critical frequency bands, required to convert hearing threshold for pure tones to that for continuous-spectrum sounds like speech.

3) The reference level or datum must be specified, e.g. 1 Pa, 20 μPa, etc.

4) In practice, measurements are made in terms of sound pressure, and that convention is retained for convenience of explanation. It is worth noting that sound pressure relative to 20 μPa in a bandwidth of 1 Hz is approximately equal to sound intensity relative to 1 pW/m^2 per Hz.

$\beta_0 - K$	Hearing threshold for continuous-spectrum sounds referred to an ERP in dB relative to 20 μPa in a bandwidth of 1 Hz.
HL	Hearing loss, usually referred to "normal" hearing threshold.
LC	Listening conditions; the manner in which the earphone and its coupling to the ear is related to the ERP.

2.3 *Concerning telephone or reference connections*

These definitions and symbols serve to characterize the telephone or reference connections in objective terms:

L_{ME}	Air-to-air transmission loss, in dB, from a MRP to an ERP.
JS, JR	Electrical interfaces at the output of a sending local telephone circuit and the input to a receiving local telephone circuit.
LTC	Local telephone circuit.
S_{MJ}	Sending sensitivity of a local telephone circuit from the MRP to the electrical output (JS).
	Note – S_{MJ} relates to a median real mouth; for practical purposes, sensitivities measured according to Recommendation P.64 using the recommended artificial mouth may be used for handset telephones.
S_{JE}	Receiving sensitivity of a local telephone circuit from the electrical input (JR) to the ERP.
	Note – S_{JE} relates to a median real ear; sensitivities measured with the artificial ear referred to in Recommendation P.64 and according to the method described therein are denoted by the symbol S_{Je}. Such values must be corrected to give appropriate values for S_{JE} (see § 6).
x_{JJ}	Transmission loss between local telephone circuits, i.e. between JS and JR in Figure 1/P.79. The circuits concerned in real telephone connections will consist of trunk junctions, trunk circuits, switching centres, etc. For assessment purposes this chain of lines is replaced by nonreactive attenuators and filters, etc. and referred to collectively by the word "junction".
S_{RMJ}, S_{RJE}, L_{RME}, etc.	are values of S_{MJ}, S_{JE}, L_{ME}, etc., applicable to a reference speech path, e.g. NOSFER or the IRS defined in Recommendation P.48.
S_{UMJ}, S_{UJE}, L_{UME}, etc.	Are values of S_{MJ}, S_{JE}, L_{ME}, etc., applicable to an unknown speech path, e.g. a telephone connection.
x_{UR}, x_{RU}	values of x applicable to combinations of "unknown" sending to reference receiving and reference sending to "unknown" receiving speech paths.
S_M	Sensitivity of a telephone microphone referred to a MRP.
S_E	Sensitivity of a telephone receiver referred to an ERP.
L_S	Electrical transmission loss from the terminals of a microphone to the line terminals of a telephone set.
L_R	Electrical transmission loss from the line terminals of a telephone set to the terminals of a receiver.
$L_{INS}\,(SL + FB)$	Transmission loss of the combination of subscriber's line and feeding bridge.

3 Structure of the theoretical model

3.1 *Definitions concerning loudness, its relationship to sensation level and loudness ratings*

These definitions and symbols relate to factors concerning loudness and loudness ratings of telephone speech paths:

Z	Sensation level, in dB, of the received speech signal at a given frequency; describes the portion of the received speech signal which is above threshold and is, therefore, effective in producing the sensation of loudness.
Z_{RO}	Value of Z when $L_{ME} = 0$ dB.

$Q(Z)$	Function of Z related to loudness; transforms sensation level expressed in terms of Z, to loudness numerics.
m	A parameter which can be used to define $Q(Z)$; represents the slope of $10 \log_{10} Q(Z)$ as function of Z.
S	A monotonic function of frequency such that equal increments of S are of equal importance to loudness, provided the associated values of Z are the same.
S'	The derivative of S with respect to frequency; $S' = \mathrm{d}S/\mathrm{d}f$. S' can be considered as a frequency weighting factor.
$\mathrm{d}S$	From the foregoing, $\mathrm{d}S = S'\,\mathrm{d}f$.
$\overline{Q(Z)}$	Weighted average of $Q(Z)$ which is related to the total loudness in a received speech signal.
λ	Loudness of the sound being considered.
OLR, SLR, RLR, JLR	Overall, sending and receiving and junction loudness ratings.

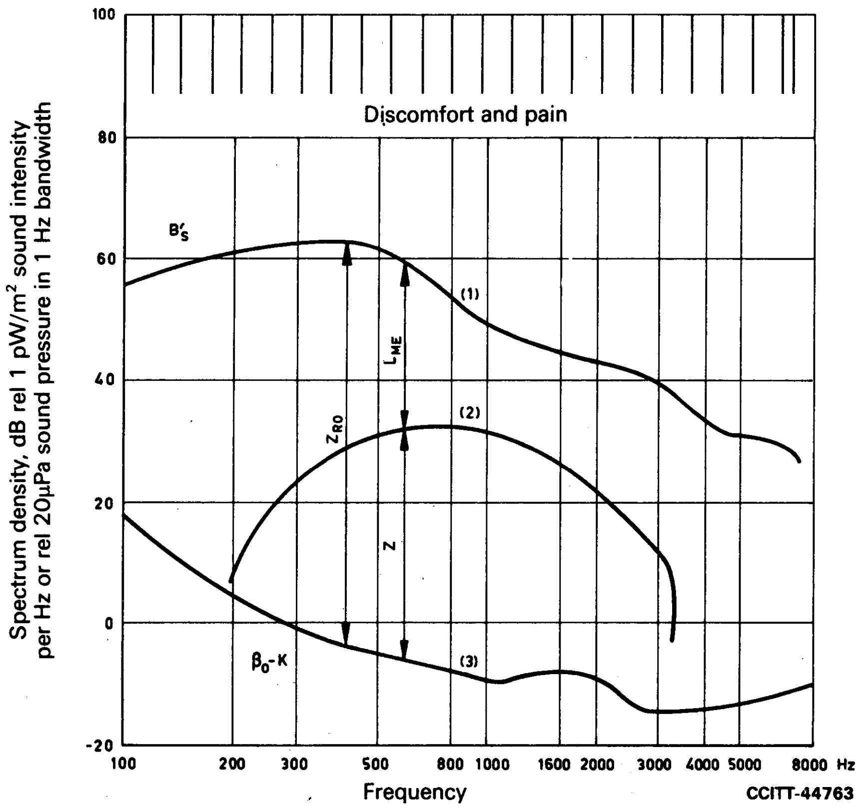
3.2 *Loudness model*

In considering speech transmission paths, it is necessary to define acoustical terminals of the paths. This can be done in terms of MRP and ERP. There are no unique definitions of such reference points, but those used here are defined in Recommendation P.64.

Curve 1 in Figure 2/P.79 shows the spectrum density B'_S of speech emitted at a certain vocal level and measured at the MRP in the absence of any obstruction in front of the mouth[5]. The measurement may be thought of as made with the aid of a very small measuring microphone. When the speech reaches the ear of the other participant in a telephone conversation, it will have been subjected to transmission loss and distortion in the telephone speech path and the spectrum density may then be as shown in Curve 2; the ERP to which Curve 2 is referred can, for explanation, be thought of as located at the opening of the ear canal, but might equally well be the tympanum, i.e. eardrum of the listener's ear. The studies at present in hand make use of an ear reference point located at the opening of the air canal (as referred to in Annex A to Recommendation P.64). The interval L_{ME} between curves 1 and 2 represents the "mouth-to-ear" transmission loss and is, in general, frequency-dependent.

The received spectrum represented by Curve 2 does not contribute uniformly to loudness, i.e. those portions of the spectrum lower in level than the listener's threshold of hearing contributes very little compared with those well above the threshold. Account is taken of this by defining a quantity termed "sensation level" (symbol Z) which is the interval between the received spectrum, Curve 2, and the threshold of audibility for continuous spectrum sounds ($\beta_0 - K$) shown in Curve 3. Loudness of the received speech sound thus depends upon Z, which is, in general, frequency-dependent.

[5] See Annex A to Recommendation P.64 for the definition of MRP.



Curve (1) Spectrum density of speech at mouth reference point.
Curve (2) Spectrum density of speech at ear reference point received over an approximately limiting telephone speech path.
Curve (3) Hearing threshold for continuous spectrum sounds.

FIGURE 2/P.79

Determination of sensation level *Z*, the portion of the received speech signal effective in producing the sensation of loudness

Studies have shown[6] that the loudness, λ, can be expressed approximately as a function of *Z* in the following manner:

$$\lambda = \mathrm{C} \int_{f_1}^{f_2} Q(Z)\, S'\, \mathrm{d}f \tag{3-1}$$

where C is a constant, $Q(Z)$ is a "loudness growth function" which transforms Z so that equal increments of the transformed values represent equal increments in loudness, S' is a "frequency weighting function" which weights the transformed values of Z according to their positions along the frequency scale and f_1 and f_2 correspond to the lower and upper frequency limits for the band of interest.

[6] This model does not claim to represent accurately all the features that relate to perception of the loudness of speech; for example, the effects of interfrequency masking are ignored and it does not predict the increasing importance of the lower frequencies as the intensity of the sound is increased from the threshold. It is possible to construct models that represent more of the features fairly well, but no completely comprehensive model is known. Such models are unnecessarily complicated for calculating loudness ratings. The most important restrictions in use of the model described here are: a) it should only be used for comparing telephone channels similar in frequency band to the intermediate reference system or commercial telephone connections; b) it should be used to make comparisons at the constant listening level indicated in Recommendation P.76.

If desired the frequency scale can be transformed to a scale of S, equal increments of which have the same "importance" so far as loudness is concerned.

Thus:

$$S' = \frac{\mathrm{d}S}{\mathrm{d}f} \tag{3-2}$$

which gives

$$\lambda = \mathrm{C} \int_{S_1}^{S_2} \overline{Q(Z)}\, \mathrm{d}S \tag{3-3}$$

where S_1 and S_2 are points on the scale of S that correspond respectively to f_1 and f_2.

The basic elements of the loudness rating process are shown in the flow diagram of Figure 3/P.79. The flow diagram depicts a "reference" spectrum decreased by the loss of a telephone connection resulting in a received spectrum which together with the threshold of hearing produces Z, the values of which (as a function of frequency) are effective in producing the sensation of loudness. Thus:

$$Z = B'_S - L_{ME} - (\beta_0 - K) \tag{3-4}$$

and Z as a function of frequency is converted to loudness, λ, according to the equations explained above in which Z is transformed to loudness numerics which are then weighted by the frequency weighting function to produce $\overline{Q(Z)}$; a constant applied to $\overline{Q(Z)}$produces λ, the loudness of the received speech expressed on some suitable scale.

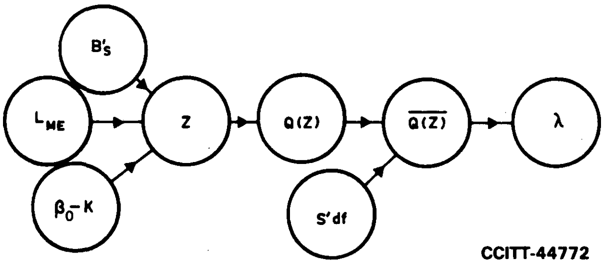


FIGURE 3/P.79

Simplified flow diagram showing how loudness, λ, is related to sensation level, Z

The flow diagram of Figure 3/P.79 represents only basic elements in the loudness rating process. These elements require further specification in order to render them unique. For example, B'_S depends on the particular speaker and his vocal level, the test phrase used, and the location of the talker's lips with respect to the telephone microphone defined by his individual method of usage and by the somewhat arbitrarily defined MRP. Similarly, the received spectrum level depends on the particular listener and his characteristics, e.g. fit between his ear and the telephone earphone when the handset is held in a prescribed manner, whether or not he has a hearing loss, and on the ERP.

Furthermore, transmission planning studies require subdivision of the connection loss, L_{ME}, into component parts, e.g. a sending component, a receiving component and an interconnecting component.

The function $Q(Z)$ can, in part, be specified in terms of a parameter m which is the slope of the logarithm of $Q(Z)$ when plotted against Z. m does, however, depend upon the listening level (or Z) in the general case but may be considered constant over a wide and useful range of Z.

Those additional factors considered at present to be of importance are included in the more detailed flow diagram of Figure 4/P.79 which is an expansion of Figure 3/P.79. The influence of these factors can be appreciated from the previous discussion and from review of the definitions given in § 3.1. Figure 3/P.79 supplements these definitions.

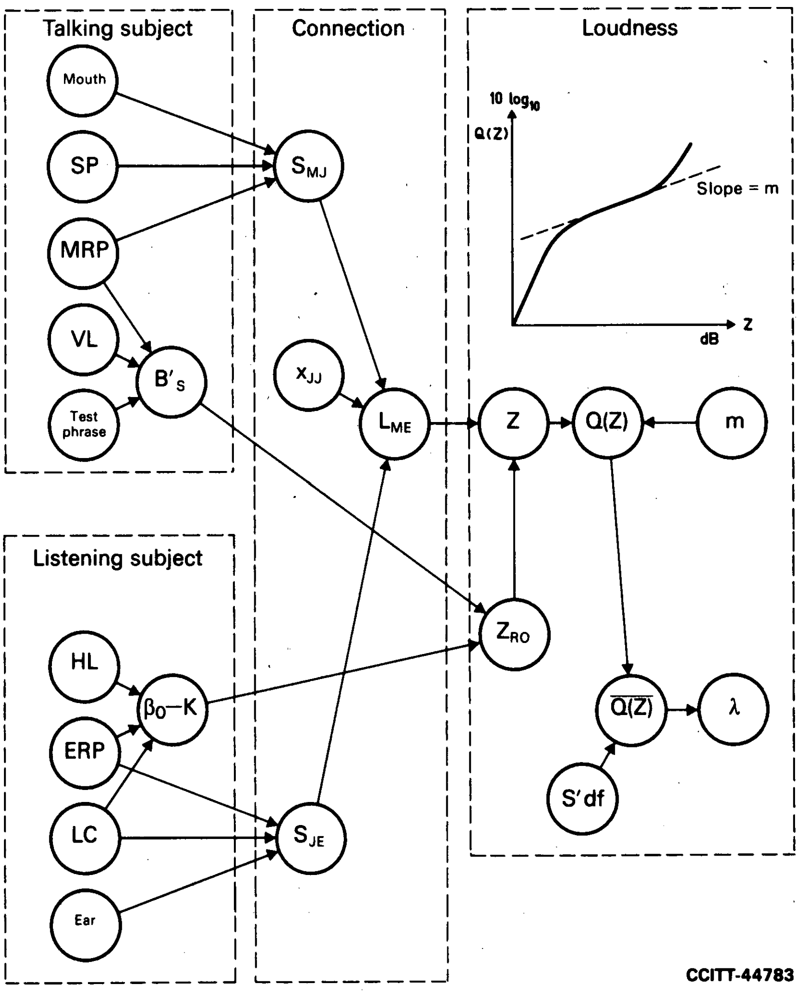


FIGURE 4/P.79

Flow diagram

4 Values of the parameters

4.1 *General*

To implement the model in the form described in § 3, it is, in principle, necessary to assign values to the following parameters:

B'_S as a function of frequency

$10 \log_{10} S'$ as a function of frequency

m which (partly) defines the loudness growth function $Q(Z)$

$\beta_0 - K$ as a function of frequency.

In fact, for the present purposes, it is convenient to group all these parameters together into a single frequency-dependent parameter which can be used with m for the purposes of calculating sending, receiving and junction loudness ratings and the loudness insertion loss of electrical elements such as channel filters in commercial telephone connections.

The theoretical derivation of this frequency-dependent parameter G, is explained below.

G, together with m, can be estimated directly from the results of subjective loudness balance tests conducted using sets of lowpass and highpass filters in a suitable reference system.

4.2 *Theoretical derivation of G*

Equation 3-1 can be written:

$$\lambda_U = C \int Q(Z_U)\ S'\ \mathrm{d}f \tag{4-1a}$$

and

$$\lambda_R = C \int Q(Z_R)\ S'\ \mathrm{d}f \tag{4-1b}$$

where λ_U and λ_R represent the loudness of speech received through the "unknown" and reference speech paths respectively and Z_U and Z_R are the corresponding values of sensation level (which are functions of frequency).

The calculation method to be described depends upon the assumption (largely verified for restricted ranges of listening level) that the function $Q(Z)$ can be put in the form:

$$Q(Z) = \text{constante} \cdot 10^{m(1/10)Z} \tag{4-2}$$

(The base 10 and the multiplier 1/10 are used merely to preserve the analogy to the decibel, in which unit Z is expressed.)

Let

$$Z_{RO} = B'_S - (\beta_0 - K) \tag{4-3}$$

and substitute in Equation 3-4 to obtain:

$$Z_U = Z_{RO} - L_{UME} \tag{4-4a}$$
$$Z_R = Z_{RO} - L_{RME} \tag{4-4b}$$

By substituting Equations (4-4a) and (4-4b) in Equations (4-1a) and (4-1b) and rearranging:

$$\lambda_U = C \int 10^{-m(1/10)L_{UME}} [10^{m(1/10)Z_{RO}}\ S']\ \mathrm{d}f \tag{4-5a}$$

$$\lambda_R = C \int 10^{-m(1/10)L_{RME}} [10^{m(1/10)Z_{RO}}\ S']\ \mathrm{d}f \tag{4-5b}$$

The loudness rating can be considered to be the Δx (independent of frequency) removed from the "unknown" speech path to render $\lambda_U = \lambda_R$.

Using the substitution:

$$G = [10^{m(1/10)Z_{RO}}\ S'] \tag{4-6}$$

and inserting $L_{UME} - \Delta x$ in Equation (4-5a) in place of L_{UME}, we obtain equality of the λ's.

Therefore

$$\int 10^{-m(1/10)(L_{UME} - \Delta x)}\ G\mathrm{d}f = \int 10^{-m(1/10)L_{RME}}\ G\ \mathrm{d}f \tag{4-7}$$

$$10^{-m(1/10)\Delta x} = \frac{\int 10^{-m(1/10)L_{UME}}\ G\ \mathrm{d}f}{\int 10^{-m(1/10)L_{RME}}\ G\ \mathrm{d}f} \tag{4-8}$$

and

$$\Delta x = -m^{-1}\, 10 \log_{10} \int 10^{-m(1/10) LUME}\, G\, \mathrm{d}f - \left\{ -m^{-1}\, 10 \log_{10} \int 10^{-m(1/10) LRME}\, G\, \mathrm{d}f \right\} \qquad (4\text{-}9)$$

Without affecting the equality, G can be scaled by multiplying with a suitable constant to render $\int G\, \mathrm{d}f = 1$; G can then be treated as a weighting factor[7] and each term on the right-hand side takes the form:

$$\Phi^{-1} \left[\int \Phi(L)\, G\, \mathrm{d}f \right] = \overline{L}$$

Then for the loudness rating we have

$$\text{loudness rating} = \Delta x = \overline{L_{UME}} - \overline{L_{RME}} \qquad (4\text{-}10)$$

The terms $\overline{L_{UME}}$ and $\overline{L_{RME}}$ can be considered as the "weighted average mouth to ear loss" of the "unknown" and reference speech paths respectively. In each of the foregoing equations, integration (and therefore averaging) is over the range between lower and upper frequency limits of interest.

For computation, the audible range of frequency is divided into a number (N) of continuous band; use is made here of the 20 ISO-preferred bands centred at frequencies spaced at approximately 1/3 octaves from 100 to 8000 Hz. Averaging the values of $\overline{L_{UME}}$ is then performed by summations of the form:

$$\overline{L_{UME}} = -m^{-1}\, 10 \log_{10} \sum_{i}^{N} 10^{-m(1/10) LUME}\, G\, \Delta f \qquad (4\text{-}11)$$

The acoustical transmission loss of a speech path is, in general, a function of frequency and can be defined as:

$$\overline{L_{UME}} = 20 \log_{10} \frac{p_M}{p_E} \qquad (4\text{-}12)$$

where p_M and p_E are as defined in §§ 2.1 and 2.2.

It is necessary to know the values of L_{UME} at each frequency together with $G\Delta f$; naturally, L_{UME} depends on the telephone speech path under consideration but $G\Delta f$ and other information common to all speech paths is described below.

4.3 *Determination of values for G*

Values have been assigned to G by analysis of results of loudness balance tests by the CCITT Laboratory using a special speech path consisting of NOSFER, but with its sending frequency response made more level by equalization. Each of a set of special low- and high-pass filters was inserted in turn in the "junction" of this speech path.

Balances were made with each filter and with the "through" path; each was treated as the "unknown" while balancing for determining relative equivalents against NOSFER with its junction set at 25 dB. Balancing was done by the "margin" method, i.e. by changing the transmission loss in the "unknown". Values of Δx were calculated for each filter and corrected for the transmission loss in the pass-band. The cut-off frequencies were taken as those frequencies at which the transmission loss was 10 dB greater than the pass-band transmission loss.

---

[7] From Equations 4-3 and 4-6 it can be seen that G as a function of frequency depends upon the value of m and the frequency-dependent functions B'_S, β_0, K and S'.

By smoothing the results and interpolating at the appropriate edges of the 20 ISO-preferred frequency bands centred at the frequencies from 100-8000 Hz, it was possible, first, to estimate m; $m = 3/\Delta x$, if we take the value of Δx at the frequency where Δx was the same for low- and for high-pass filtering. Then, by use of Equation 4-8 and some interaction, it was possible to obtain a set of values for G which satisfied the experimental data. Note that L_{RME} in Equations 4-7 to 4-10 represents the mouth-to-ear transmission loss of the "through" path and L_{UME} represents that of the same path with the filter inserted.

The results are given in Table 1/P.79, the value determined for m being 0.175.

TABLE 1/P.79

Values of 10 $\log_{10} G$ and 10 $\log_{10} G \Delta f$ determined by the CCITT Laboratory

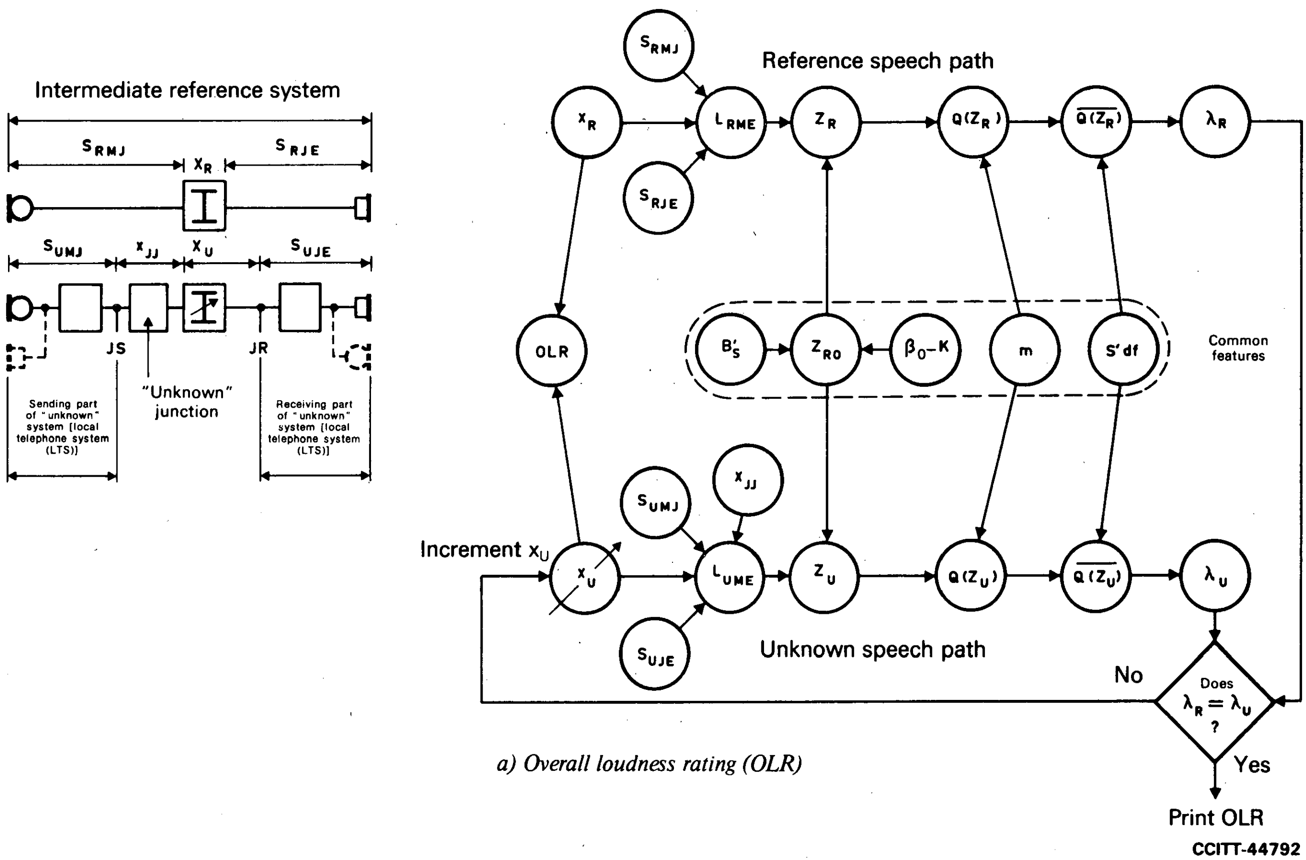
Midfrequency (Hz)	Δf (Hz)	10 $\log_{10} G$ (dB)	10 $\log_{10} G \Delta f$ (dB)
100	22.4	−32.63	−19.12
125	29.6	−29.12	−14.41
160	37.5	−27.64	−11.90
200	44.7	−28.46	−11.96
250	57.0	−28.58	−11.02
315	74.3	−31.10	−12.39
400	92.2	−29.78	−10.14
500	114.0	−32.68	−12.12
630	149.0	−33.21	−11.48
800	184.0	−34.14	−11.49
1000	224.0	−35.33	−11.83
1250	296.0	−37.90	−13.19
1600	375.0	−38.41	−12.67
2000	447.0	−41.25	−14.75
2500	570.0	−41.71	−14.15
3150	743.0	−45.80	−17.09
4000	922.0	−43.50	−13.86
5000	1140.0	−47.13	−16.56
6300	1490.0	−48.27	−16.54
8000	1840.0	−46.47	−13.82

5 Calculation of loudness ratings

5.1 *Deviation of formulas and W weights*

The method described in Recommendation P.78 can be described in terms of the flow diagrams illustrated in Figure 5/P.79 which also embody the structure of the model used here (Figure 4/P.79). The diagrams placed on the left in parts a), b), c) and d) of Figure 5/P.79 are redrawn versions of the various paths given in Figure 1/P.78.

Figure 5/P.79 illustrates the procedure when values are known for all the parameters referred to in §§ 1, 2 and 3. In a) of Figure 5/P.79, the parameters shown grouped together are those used to form the composite parameter G described in § 4. Further grouping is possible as shown in b), c) and d) of Figure 5/P.79. It will also be seen that the whole of the path from x_R to λ_R is also common to all four flow diagrams. Use can be made of this feature to reduce the calculation procedure to a formula which is very easy to compute.



a) Overall loudness rating (OLR)

Note concerning a) of Figure 5/P.79

The "unknown" path consists of four sections as follows:

a) sending LTS, comprising telephone set, subscriber's line and feeding bridge, up to JS in Figure 1/P.79;

b) receiving LTS, comprising feeding bridge, subscriber's line and telephone set, from JR in Figure 1/P.79;

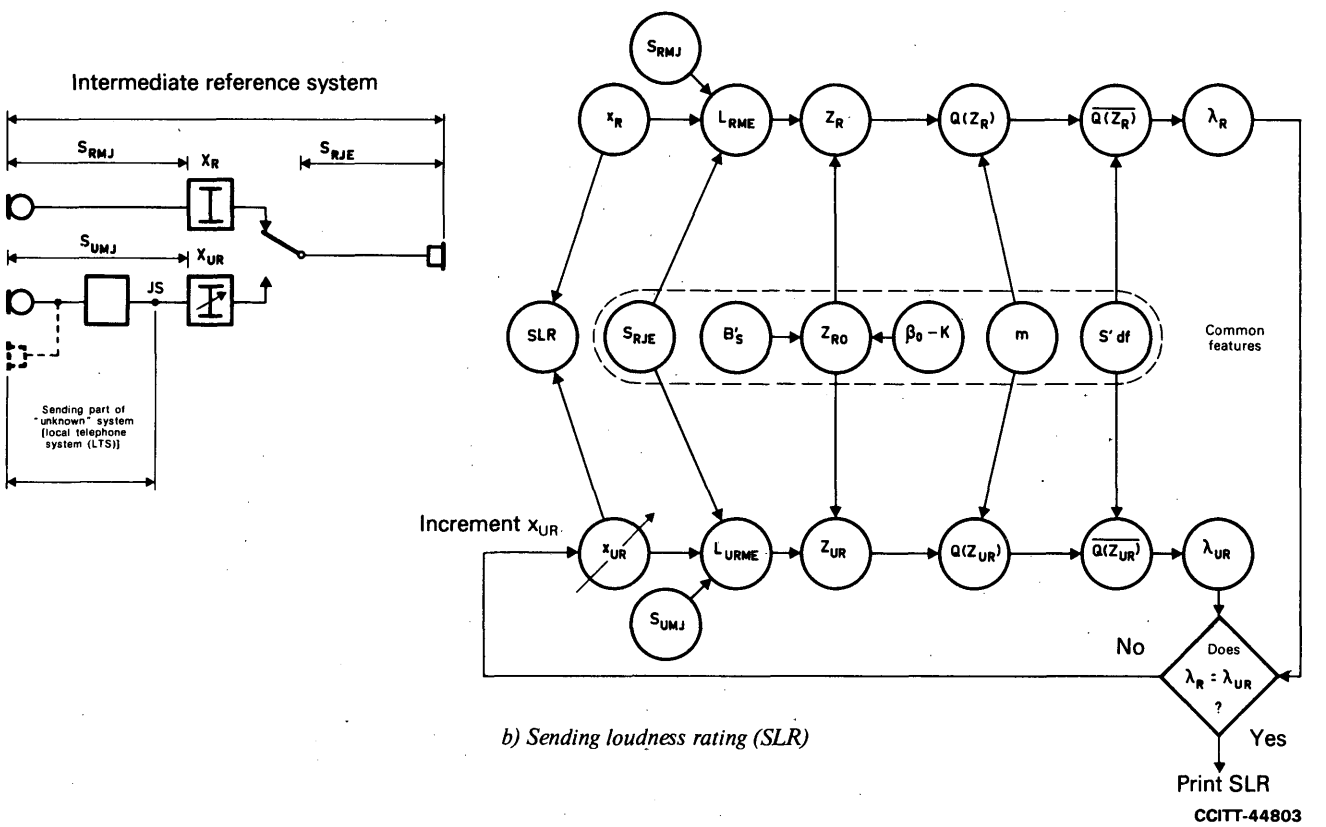
c) the combination of trunk junctions and trunk circuits present in the real connection between JS and JR;

d) additional, adjustable, transmission loss, x_U, introduced in such a manner that it will not disturb the overall frequency response of the complete connection, but will only increase the transmission loss equally at all frequencies.

If the section of the real connection between JS and JR has an image impedance of 600 ohms $\underline{/0^\circ}$, there is no problem either in defining x_{JJ} or in introducing the additional loss, x_U. Where this is not so, the image attenuation of a virtual network having 600 ohms resistance image impedances has to be determined (and a network constructed if actual subjective determinations are to be made). Particularly difficult problems are encountered if the real connection contains no part in the section between JS and JR that has a 600 ohms image impedances (such as in a local call connection), but these can be overcome satisfactorily by calculation. Provided that a part is present having at least about 7 dB attenuation and 600 ohms image impedances, the problems can be overcome fairly easily.

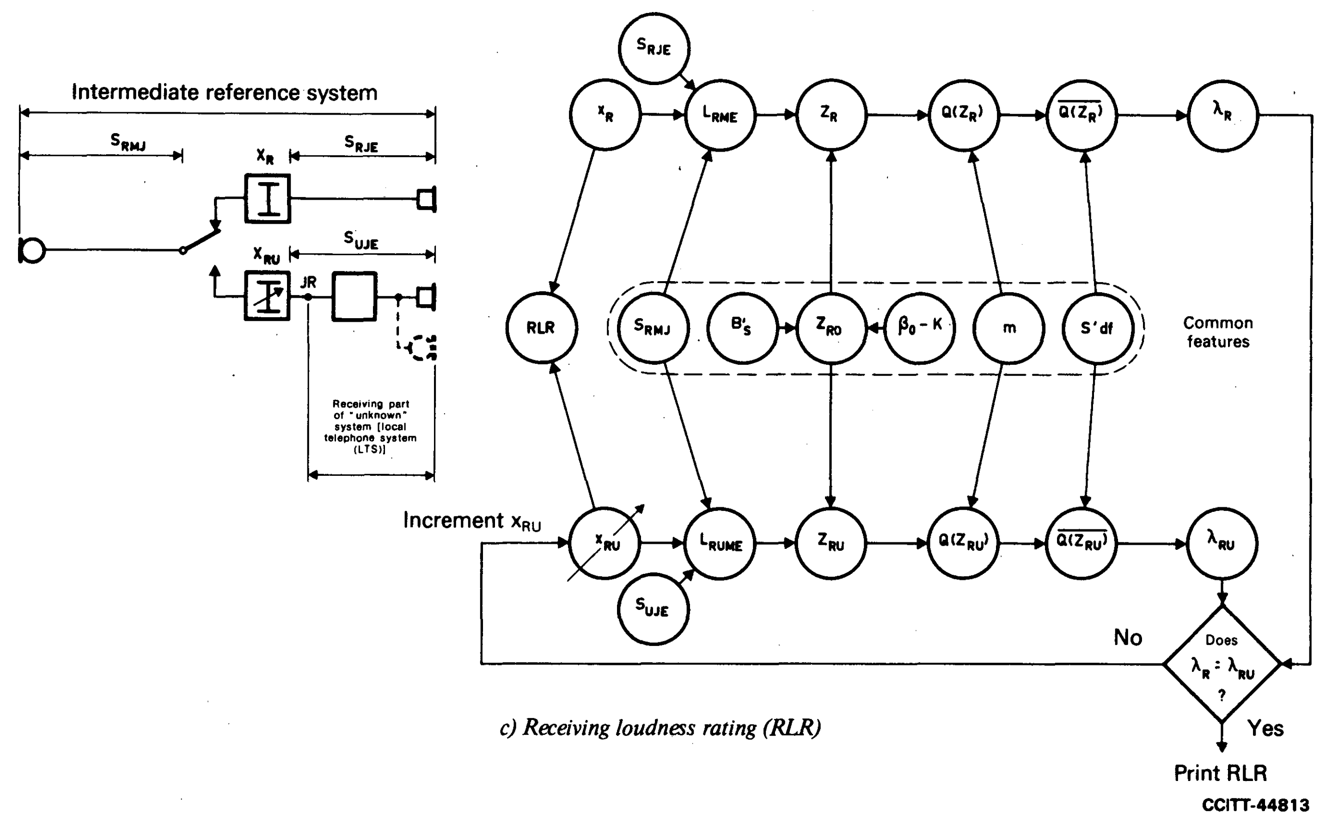
FIGURE 5/P.79

Flow diagrams illustrating determination of loudness ratings



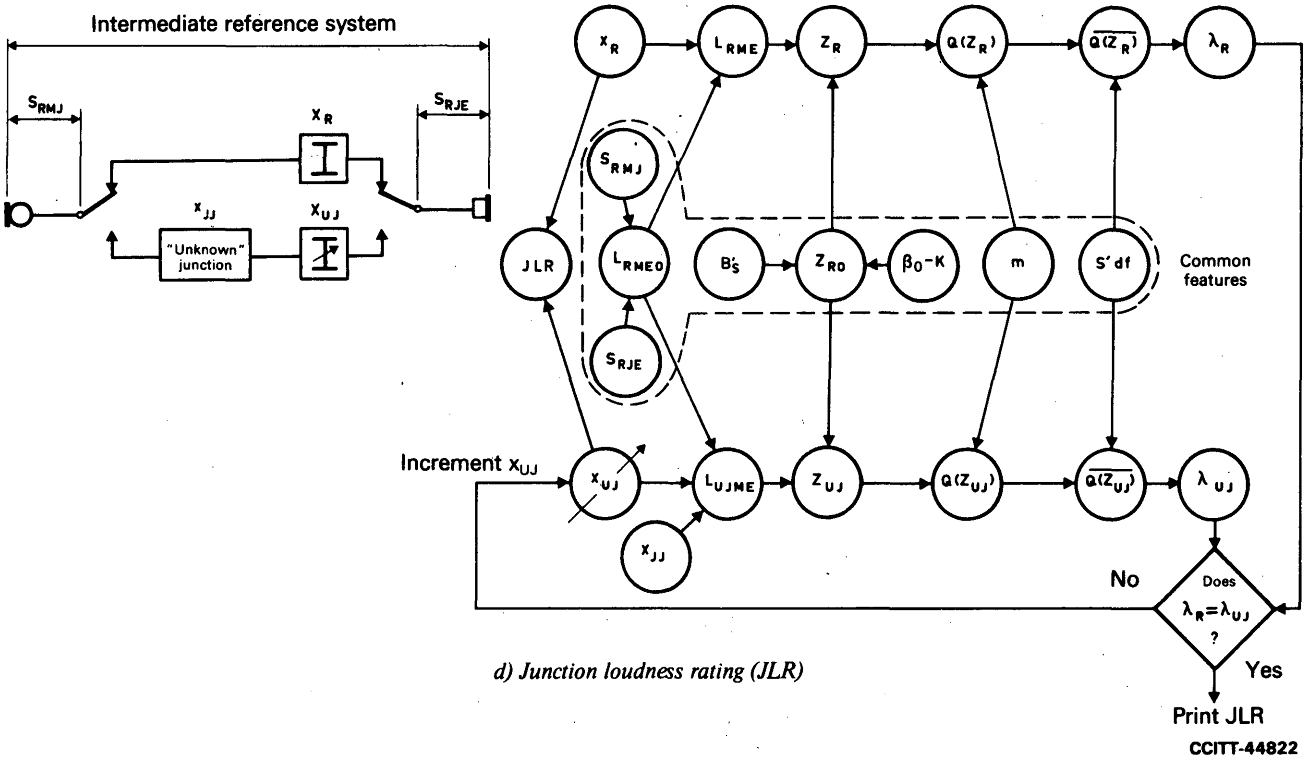
b) Sending loudness rating (SLR)

FIGURE 5/P.79 *(continued)*



c) Receiving loudness rating (RLR)

FIGURE 5/P.79 *(continued)*



d) Junction loudness rating (JLR)

FIGURE 5/P.79 *(end)*

Taking m as constant with the value 0.175, use can be made of the substitution:

$$W_i = -57.1 \log_{10} G \Delta f \tag{4-13}$$

Equation (4-11) can then be simplified in appearance to:

$$\overline{L_{UME}} = -57.1 \log_{10} \sum_{i}^{N} 10^{-(1/57.1)(L_{UME} + W_i)} \tag{4-14}$$

For the present purposes, the reference speech path will be taken as the "intermediate reference system" (IRS) defined in Recommendation P.48 and set with its attenuator at 0 dB; having fixed the reference speech path, L_{RME} becomes constant, i.e. independent of *i*. Therefore Equations 4-10 and 4-14 can be combined to form:

$$\text{loudness rating} = -57.1 \log_{10} \sum_{i}^{N} 10^{-(1/57.1)(L_{UME} - \overline{L_{RME}} + W_i)} \tag{4-15}$$

When rating commercial local telephone circuits, the values of $\overline{L}_{UME}$ can be obtained for any given "unknown" speech path combining appropriate sending and receiving sensitivites, S_{MJ} and S_{JE}, in appropriate combinations.

For determining an "overall loudness rating" (OLR),

$$L_{UME} = -(S_{UMJ} + S_{UJE}) \tag{4-16a}$$

For determining a sending loudness rating (SLR) of a local telephone circuit,

$$L_{URME} = -(S_{UMJ} + S_{RJE}) \tag{4-16b}$$

For determining a receiving loudness rating (RLR) of a local telephone circuit,

$$L_{RUME} = -(S_{RMJ} + S_{UJE}) \tag{4-16c}$$

and for determining a "junction" loudness rating (JLR)

$$\begin{aligned} L_{UJME} &= -(S_{RMJ} + S_{RJE}) + x_{JJ} \\ \text{and} \quad & \\ L_{RMEO} &= -(S_{RMJ} + S_{RJE}) \end{aligned} \tag{4-16d}$$

Substituting these in Equation 4-15:

$$\text{OLR} = -57.1 \log_{10} \sum_{i}^{N} 10^{(1/57.1)(S_{UMJ} + S_{UJE} + \overline{L_{RME}} - W_i)} \tag{4-17a}$$

$$\text{SLR} = -57.1 \log_{10} \sum_{i}^{N} 10^{(1/57.1)(S_{UMJ} + S_{RJE} + \overline{L_{RME}} - W_i)} \tag{4-17b}$$

$$\text{RLR} = -57.1 \log_{10} \sum_{i}^{N} 10^{(1/57.1)(S_{UJE} + S_{RMJ} + \overline{L_{RME}} - W_i)} \tag{4-17c}$$

$$\text{JLR} = -57.1 \log_{10} \sum_{i}^{N} 10^{(1/57.1)(-x_{JJ} - L_{RMEO} + \overline{L_{RME}} - W_i)} \tag{4-18}$$

The terms $\overline{L_{RME}}$ and W_i are common to each of the Equations 4-17 and so further computational simplification is possible by making the following substitutions:

$$W_O = W_i - \overline{L_{RME}} \tag{4-18a}$$

$$W_S = W_i - S_{RJE} - \overline{L_{RME}} \tag{4-18b}$$

$$W_R = W_i - S_{RMJ} - \overline{L_{RME}} \tag{4-18c}$$

$$W_J = W_i + L_{RMEO} - \overline{L_{RME}} \tag{4-18d}$$

When the substitutions are made, the equations become:

$$\text{OLR} = -57.1 \log_{10} \sum_{i}^{N} 10^{(1/57.1)(S_{UMJ} + S_{UJE} - W_O)} \tag{4-19a}$$

$$\text{SLR} = -57.1 \log_{10} \sum_{i}^{N} 10^{(1/57.1)(S_{UMJ} - W_S)} \tag{4-19b}$$

$$\text{RLR} = -57.1 \log_{10} \sum_{i}^{N} 10^{(1/57.1)(S_{UJE} - W_R)} \tag{4-19c}$$

$$\text{JLR} = -57.1 \log_{10} \sum_{i}^{N} 10^{(1/57.1)(x_{JJ} - W_J)} \tag{4-19d}$$

Table 2/P.79 shows the values for these "weighting" factors which have been derived from the information in Table 1/P.79 with $m = 0.175$.

TABLE 2/P.79

Weighting factors for calculating loudness ratings

Band No.	Mid-frequency (Hz)	Send W_s	Receive W_R	Junction W_J	Overall W_o
1	100	154.5	152.8	200.3	107.0
2	125	115.4	116.2	151.5	80.1
3	160	89.0	91.3	114.6	65.7
4	200	77.2	85.3	96.4	66.1
5	250	62.9	75.0	77.2	60.7
6	315	62.3	79.3	73.1	68.5
7	400	45.0	64.0	53.4	55.6
8	500	53.4	73.8	60.3	66.9
9	630	48.8	69.4	54.9	63.3
10	800	47.9	68.3	52.8	63.4
11	1000	50.4	69.0	54.1	65.3
12	1250	59.4	75.4	61.7	73.1
13	1600	57.0	70.7	57.6	70.1
14	2000	72.5	81.7	72.2	82.0
15	2500	72.9	76.8	71.1	78.6
16	3150	89.5	93.6	87.7	95.4
17	4000	117.3	114.1	154.5	76.9
18	5000	157.3	144.6	209.5	92.4
19	6300	172.2	165.8	245.8	92.2
20	8000	181.7	166.7	271.7	76.7

5.2 *Loudness rating calculations over a reduced bandwidth*

In practical cases the complete information for all 20 bands may not be available, or, for some extreme bands, may not be reliable. In such cases it will be desirable to restrict the frequency range over which calculations of loudness are made.

This may be done quite simply by using only those bands for which reliable figures exist and making an allowance equal to the loudness rating of the overall IRS connection calculated over the same reduced bandwidth. This allowance may conveniently be incorporated into the calculations by reducing the W weights uniformly by an appropriate figure, or by simply reducing (subtracting from) the resulting loudness rating by the allowance.

Table 3/P.79 gives some examples of the allowance to be applied for various reduced bandwidths.

Other allowances may be calculated by determining the IRS overall loudness rating for the required bandwidth.

TABLE 3/P.79

Allowance to be subtracted from W weights for reduced bandwidths

Bands	Allowance
3 – 18 (inclusive)	0.1 dB
3 – 17 (inclusive)	0.1 dB
4 – 18 (inclusive)	0.3 dB
4 – 17 (inclusive)	0.3 dB
6 – 16 (inclusive)	2.1 dB

6 Sensitivity and transmission loss data required

The sending sensitivity of the local telephone system, S_{MJ}, should be determined in principle using real mouths and real speech but it is usually sufficient to make these measurements using an artificial mouth and suitable test signal. See Recommendation P.64 for particulars.

The receiving sensitivity of the local telephone system, S_{JE}, should be determined in principle using real ears. The determination of the sensitivity denoted by S_{Je}, using an artificial ear, is explained in Recommendation P.64 but this quantity differs from the quantity required here by the artificial/real ear correction L_E, that is:

$$S_{JE} = S_{Je} - L_E$$

The value of L_E usually depends upon the frequency and upon the manner in which the earphone is held to the ear.

Table 4/P.79 shows values obtained for one type of telephone held fairly closely to the ear. Use of these values for calculation has given reasonably good agreement with receiving and junction loudness ratings determined by subjective measurements in the CCITT Laboratory. Such calculations have used these values of L_E for both the IRS and the "unknown".

The values of S_{RJE} used to determine the values of W_s in Table 2/P.79 include a correction for L_E corresponding to the values of Table 4/P.79. The values of S_{UJE} used in the calculation defined by Equations 4-19a and 4-19c should also include a correction for L_E, using either the values of Table 3/P.79 or other values which might be considered more appropriate for the conditions of use.

Note that the values of L_E used for the IRS have some effect on the calculated values of junction loudness rating. This matter is receiving further study under Questions 8/XII [3] and 12/XII [4].

TABLE 4/P.79

Values of L_E

Frequency (Hz)	L_E (dB)	Frequency (Hz)	L_E (dB)
100	20.0	1000	−2.3
125	16.5	1250	−1.2
160	12.5	1600	−0.1
200	8.4	2000	3.6
250	4.9	2500	7.4
315	1.0	3150	6.7
400	−0.7	4000	8.8
500	−2.2	5000	10.0
630	−2.6	6300	12.5
800	−3.2	8000	15.0

The transmission loss x_{JJ} is the insertion loss between 600-ohms terminations of the chain of transmission elements between JS and JR in Figure 1/P.79. Direct summation (with due respect to sign) of this quantity with S_{UMJ} and S_{UJE} will not, in general, give L_{UME} exactly because there are usually some impedance mismatches. Care must therefore be taken to determine L_{UME} correctly when calculating overall loudness ratings. The inaccuracy will be severe when the transmission loss x_{JJ} is small and when the image impedances of the elements between JS and JR depart considerably from 600 ohms. The correct values for L_{UME} can be obtained by direct measurement or by calculation taking all impedance mismatches properly into account.

7 Restrictions of use

The calculation procedure described here and the values given for the parameters are suitable for calculating sending, receiving and junction loudness ratings. They may also be used for calculating overall loudness ratings and loudness insertion loss provided the complete speech paths concerned are restricted to the telephone frequency band, i.e. nominally to the range 300-3400 Hz.

They are not suitable for making comparisons between speech paths having considerable differences in frequency band.

The values of the parameters have been chosen to give reasonably good agreement with subjective loudness rating determinations by the CCITT Laboratory using the method described in Recommendation P.78. For sending and receiving loudness ratings, the calculated values may be expected to agree fairly well with subjective determinations conducted elsewhere. Certain differences have however been found when junction loudness ratings and loudness insertion losses have been compared with subjective determinations performed by other laboratories.

8 Calculation of sidetone masking rating (STMR)

8.1 *Calculation from first principles*

Recommendation P.76 describes the principles underlying the sidetone masking rating method in which the human sidetone signal L_{MEHS} is treated as a masking threshold against which the telephone sidetone path loss, L_{meST}, is rated. As previously reported the human sidetone path loss, L_{MEHS}, has been determined [5] and is shown graphically in Figure 4/P.76, and in tabular form below in Table 5/P.79. Two sets of values are given in Table 5/P.79 for use depending on whether the conditions of interest are for an earphone coupling that is sealed (column 9) or with a typical leak included (column 10).

The calculation method for STMR makes use of the same underlying principles as described for sending and receiving loudness ratings in §§ 3 and 4. The calculation procedure is summarized by the expression:

$$\mathrm{STMR} = \frac{10}{m} \log_{10} \frac{\sum 10^{\frac{mZ_l + 10 \log_{10} S' \Delta f}{10}}}{\sum 10^{\frac{mZ + 10 \log_{10} S' \Delta f}{10}}} \tag{8-1}$$

where

$$Z = B'_S - L_{meST} - L_E - 10 \log_{10} \left(10^{\frac{\beta_0 - K}{10}} + 10^{\frac{B'_S - L_{MEHS}}{10}}\right) \tag{8-2}$$

and

$$Z_l = B'_S + S_{RmJ} + S_{RJe} - L_E - 10 \log_{10} \left(10^{\frac{\beta_0 - K}{10}} + 10^{\frac{B'_S - L_{MEHS}}{10}}\right) \tag{8-3}$$

where the quantities used are as defined in earlier sections but where, for m, an index:

$$m = 0.225$$

The summations are normally extended over the range 100 Hz to 8 kHz but may be restricted if L_{meST} cannot be satisfactorily determined over the full bandwidth.

Table 5/P.79 lists the values for each of the quantities at the ISO frequencies.

TABLE 5/P.79

Listing of quantities necessary for the calculation of STMR

Band No.	f	B'_S	$\beta_0 - K$	$10 \log_{10} S' \Delta f$	IRS S_{RmJ}	IRS S_{RJe}	L_E	L_{MEHS}	
	Hz	dB	dB	dB	dB	dB	dB	dB	dB
		1 pW/m²/Hz			1 V/Pa	1 Pa/V		Sealed	Unsealed
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	100	57.3	17.5	−19.7	−45.8	−27.5	20	−2.7	11.6
2	125	60.2	14.4	−18.8	−36.1	−18.8	16.5	−4	10.6
3	160	62.0	10	−17.8	−25.6	−10.8	12.5	−5.4	7.1
4	200	63.0	5	−17	−19.2	− 2.7	8.4	−2.7	7.6
5	250	63.0	2.5	−16	−14.3	2.7	4.9	−2.8	7.4
6	315	62.4	− 0.4	−15.1	−10.8	7.2	1.0	−2.6	6.1
7	400	61.1	− 3	−14.4	− 8.4	9.9	−0.7	−0.7	3.5
8	500	59.3	− 5	−13.6	− 6.9	11.3	−2.2	5	5.7
9	630	57.0	− 6.3	−13.3	− 6.1	11.9	−2.6	13.2	8.9
10	800	54.4	− 8	−12.8	− 4.9	12.3	−3.2	19.9	16.2
11	1000	51.5	− 9	−12.4	− 3.7	12.6	−2.3	26.1	23.8
12	1250	48.4	− 8.5	−12.2	− 2.3	12.5	−1.2	23.7	23.7
13	1600	45.4	− 8	−11.9	− 0.6	13	−0.1	22	22
14	2000	42.3	− 9	−11.9	0.3	13.1	3.6	21.1	21.1
15	2500	39.5	−11.5	−12	1.8	13.1	7.4	22.1	22.1
16	3150	36.8	−13.8	−12.1	1.8	12.6	6.7	23.3	23.3
17	4000	34.6	−13	−12.4	−37.2	−31.6	8.8	24.2	24.2
18	5000	32.8	−12.5	−12.5	−52.2	−54.9	10.0	(26)	(26)
19	6300	31.5	−11.1	−13	−73.6	−67.5	12.5	(28)	(28)
20	8000	30.9	− 9	−14	−90	−90	15.0	(30)	(30)

8.2 *Calculation of STMR using W weights*

In § 4 above, the fundamental principles underlying the loudness rating procedure for sending, receiving, overall and junction loudness ratings were further developed, and a simplified equation derived which makes use of the W weights listed in Table 2/P.79 together with simplified equations (4-19a) to (4-19d). The equations (8-1), (8-2) and (8-3), applying to the STMR calculation, may also be reduced to a simplified equation that makes use of a set of W weights and a value of m unique to STMR, thus:

$$\mathrm{STMR} = -\frac{10}{m} \log_{10} \sum_{1}^{N} 10^{(m/10)(-L_{meST} - L_E - W_M)} \tag{8-4}$$

where $m = 0.225$

and W_M take the values given in Table 6/P.79.

In deriving W weights for the unsealed condition, (column 3, Table 6/P.79) values of L_E in accordance with column 8, Table 5/P.79 have been assumed for the reference path (IRS). When calculating STMR unsealed, appropriate values of L_E should be added to the L_{meST} values and inserted in the formula as indicated. In many cases the L_E values of column 8, Table 5/P.79 will be satisfactory.

For the sealed condition the weights of column 2, Table 6/P.79 should be used and the L_E values associated with L_{meST}, set to zero.

TABLE 6/P.79

Weighting factors for calculating STMR

Band No.	W_{MS} sealed	W_{ML} unsealed
(1)	(2)	(3)
1	110.4	94.0
2	107.7	91.0
3	104.6	90.1
4	98.4	86.0
5	94.0	81.8
6	89.8	79.1
7	84.8	78.5
8	75.5	72.8
9	66.0	68.3
10	57.1	58.7
11	49.1	49.4
12	50.6	48.6
13	51.0	48.9
14	51.9	49.8
15	51.3	49.3
16	50.6	48.5
17	51.0	49.0
18	49.7	47.7
19	50.0	48.0
20	52.8	50.7

8.3 *Comments on sealed versus unsealed conditions for the calculation of STMR*

In deriving values of L_{MEHS} for the sealed ear, very stringent measures were taken to eliminate leaks between the earcap of the test receiver and the subjects' ears. For L_{MEHS} unsealed a particular value of L_E was acoustically inserted at the receiver. The difference between the L_{MEHS} sealed and L_{MEHS} with leak can be seen by comparing columns 9 and 10 of Table 5/P.79. Over the most important parts of the frequency range this difference approximates to the value of L_E used at the receiver. In practice, rating differences (sealed-unsealed) are generally less than 1 dB.

This suggests that in practice any leak present will affect L_{MEHS} and L_{MEST} approximately equally, at least over a practical range of acoustic leaks. This in turn suggests that the L_{MEHS} will always have approximately the same masking effect with respect to L_{MEST} irrespective of any leak present and that for purposes of rating sidetone loudness STMR is expected to give better correlation with subjective effects if calculated for sealed ear conditions.

Use of the sealed condition is preferred, but Administrations may continue to use STMR unsealed for experimental purposes or where accumulation of data makes it sensible to do so, e.g. for certain existing specifications. If this is the case it must be clearly stated in the related documentation.

Information on other aspects of sidetone will be found in annexes to Question 9/XII [6] and in Supplement No. 11 at the end of this volume.

References

[1] CCITT – Question 19/XII, Contribution COM XII-No. 1, Study Period 1985-1988, Geneva, 1985.

[2] CCITT – Question 7/XII, Contribution COM XII-No. 1, Study Period 1985-1988, Geneva, 1985.

[3] CCITT – Question 8/XII, Contribution COM XII-No. 1, Study Period 1985-1988, Geneva, 1985.

[4] CCITT – Question 12/XII, Contribution COM XII-No. 1, Study Period 1985-1988, Geneva, 1985.

[5] CCITT – Contribution COM XII-No. 228/AP VII-No.115, Study Period 1977-1980, Geneva, 1980.

[6] CCITT – Question 9/XII, Contribution COM XII No. 1, Study Period 1985-1988, Geneva, 1985.

PAGE INTENTIONALLY LEFT BLANK

PAGE LAISSEE EN BLANC INTENTIONNELLEMENT

PART II

SUPPLEMENTS TO SERIES P RECOMMANDATIONS

PAGE INTENTIONALLY LEFT BLANK

PAGE LAISSEE EN BLANC INTENTIONNELLEMENT

Supplement No. 1

PRECAUTIONS TO BE TAKEN FOR CORRECT INSTALLATION AND MAINTENANCE OF AN IRS

(For this Supplement see Volume V of the *Orange Book*)

Supplement No. 2

METHODS USED FOR ASSESSING TELEPHONY TRANSMISSION PERFORMANCE

(Geneva, 1980; modified at Malaga-Torremolinos, 1984)

(Quoted in Recommendation P.74)

(Contribution from British Telecom)

1 Introduction

This Supplement gives brief descriptions of the methods for assessing telephony transmission performance that are recommended by the CCITT or have been employed over Study Periods 1968 to 1980 in studying Questions assigned to Study Group XII. Some of the methods are already fully described in Recommendations and these will merely be listed here with reference to the appropriate Recommendation. Other methods are also described in detail elsewhere; the essential features of these are given here with a brief description of how they are conducted, with reference to descriptions published elsewhere.

2 List of methods

a) loudness comparison for speech (reference equivalents and loudness ratings);

b) articulation (AEN) ratings;

c) listening opinion tests;

d) conversation opinion tests;

e) quantal-response detectability tests.

3 Brief descriptions and references to more complete descriptions

3.1 Loudness comparisons for speech are intended to quantify the relative level at which speech, transmitted over a given telephone connexion, reaches the ears of customers while they are listening to a person talking at the other end. In order to standardize the measuring procedure, the talking and listening conditions are each controlled in a specified manner. Circuit noise and room noise are excluded from the determination and so the results are governed by the overall mouth-to-ear transmission loss of the speech path being considered. The present recommended method is given in Recommendation P.72 and proposals for new methods are to be found in Question 15/XII [1]. More general information can be found in Reference [2].

3.2 Articulation measurements are based on measurement of the fraction of speech sounds recognized correctly when transmitted and reproduced over the speech path in question. Circuit noise and room noise at specified levels should be present and the result is affected by their levels. Just as for § 3.1 above, talking and listening conditions are controlled. The method recommended by the CCITT is described in Recommendation P.45. Other information will be found in Reference [2].

3.3 Listening opinion tests are conducted using speech material in the form of sentences and the listeners judge the speech received over the path according to a given criterion. The method has been widely used, and further details can be found in Reference [2].

3.3.1 *Method of conducting listening opinion tests*

The speech is usually recorded so that it can be reproduced at a given level. The recordings for this purpose must be carefully made and copied so that uncontrolled degradations do not appear. Circuit noise and room noise may be present, and their effects are taken into account.

Two subjective criteria commonly used are loudness preference and listening effort, for which the following scales are used.

- Loudness preference scale:

  Opinion scale No. 4A

  A Much louder than preferred.

  B Louder than preferred.

  C Preferred.

  D Quieter than preferred.

  E Much quieter than preferred.

- Listening effort scale:

  Opinion scale No. 7: Opinions based on the effort required to understand the meanings of sentences

  A Complete relaxation possible; no effort required.

  B Attention necessary; no appreciable effort required.

  C Moderate effort required.

  D Considerable effort required.

  E No meaning understood with any feasible effort.

Experimental design is usually based on a graeco-latin or hyper-graeco-latin square, in which rows represent listeners, columns represent the order in which conditions are administered, symbols of the first alphabet represent circuit conditions, and symbols of other alphabets represent talkers and lists of sentences. Each cell of the design thus represents a "run", in which a particular list of sentences, recorded by a particular talker, is replayed via a particular circuit condition to a particular listener at a particular position in the sequence of conditions presented to that listener. Within each run the listening level is varied over a number of predetermined values in random order, one value per group of five sentences, and the subject votes on one of the above opinion scales at the end of each group. Rarely some other parameter, such as bandwidth, is varied within each run instead of listening level.

In listening-effort tests, listeners are specially prone to what is known as the "enhancement" effect: that is, their standards of judgement are liable to be strongly influenced by the range of quality and listening level occurring in the same test, and especially within the same run. It is therefore important that the circuit conditions chosen should not include too many bad ones (that is, conditions that will yield a poor listening-effort score even with the best listening levels), that every run should cover a range of listening levels from well above optimum to at least 30 dB below optimum, and that within each run at least one group of sentences should be heard via an "anchor" condition (a good condition with a good listening level). It is also important that groups and lists of sentences should not vary too widely in their intrinsic comprehensibility, and that no subject should hear the same sentence more than once in the same experiment, because the listening effort needed to understand a familiar sentence would obviously be reduced.

The votes using the above scales are scored respectively 4, 3, 2, 1 and 0: the mean of these values for each circuit condition is called "mean opinion score". The opinion scores are processed by analysis of variance in order to verify that the effects due to circuit condition, listening levels, talkers, listeners and other factors are as expected, to determine their significance, and to evaluate confidence intervals. It is usual to express the relationship between listening level and loudness-preference mean opinion score (scale 4A) by fitting an equation describing a straight line or logistic curve, whereas the relationship between listening level and listening-effort mean opinion score (scale 7) is expressed by a fitted quadratic or more complicated equation; other features of the circuit conditions may also enter as parameters into these equations.

Listening tests using sentence material can also be conducted as pair-comparisons, but these should be undertaken with due consideration to ensure that subjects become suitably adapted to each test condition.

3.4 Conversation tests may be conducted either as interviews after real customers have made actual calls or as laboratory tests. Further information regarding methods recommended by the CCITT for the former is given in Question 2/XII [3]. Laboratory conversation tests are intended as far as possible to reproduce under laboratory conditions the actual service conditions experienced by telephone customers: to this end it is necessary to choose the circuit conditions and subjects suitably, and to administer the tests in an appropriate manner. A method intermediate between field observations and laboratory tests is that used by the AT&T Co. and called SIBYL (see also Reference [4]). Particulars of the method used by British Telecom are given below.

3.4.1 *Method for conducting conversation tests*

The need for careful and exhaustive preparations cannot be too strongly emphasized. It will be obvious to all that the connexions must be correctly specified and set up, and measured accurately before and after each experiment; that auxiliary facilities such as dialling and ringing must be provided, so that any of the desired connections can be chosen and established quickly and without error; and that faithful records of the output of each test must be kept. But some other equally important considerations are less obvious. The following gives an outline of a system that takes all these matters into account, and has been found satisfactory in British Telecom.

3.4.1.1 *Experimental design*

The most suitable designs are of the $n \times n$ graeco-latin square type, where each of n pairs of subjects carries out one conversation on each of n circuit conditions. Precision is very low if n is less than 8; at the other extreme it is not practical to expect subjects to attend on more than four occasions, or to carry out more than four conversations per visit. Moreover, the total number of conversations, $n \times n$, increases much more rapidly than n. For this reason, n is normally limited to the range 8 to 15 inclusive: graeco-latin squares (with symbols from two alphabets) exist for all these numbers. In such a design, the convention is that rows denote pairs of subjects; columns denote the order of administering the experiment; symbols of the first alphabet denote circuit conditions (distinguished not only according to properties of the connections by themselves, but also according to room noise levels and any other "treatment" factors); symbols of the second alphabet denote sets of pictures used as the topic of conversation. No further orthogonal factors can be incorporated where n = 10, 12, 14 or 15; but where n = 8, 9, 11 or 13, it is possible to construct hyper-graeco-latin squares with symbols from ($n - 3$) additional alphabets, which may be used to govern further orthogonal factors (such as selection of carbon microphones, choice of calling party, or choice of crosstalk recording), for each conversation. When the square is not hyper-graeco-latin these factors must be allocated by some simple balanced rotation scheme, but this may give rise to biases that cannot be eliminated from the results. For this reason the recommended value of n is now 13 rather than 12 as previously.

To the basic square is added an extra column at the beginning, having the same circuit condition and the same picture set for all pairs of subjects. This column represents a preliminary conversation for each pair of subjects, which serves to accustom them to the procedure, and to some extent stabilizes their standards of judgement. Thus each of the n pairs of subjects carries out ($n + 1$) conversations altogether. The results from the preliminary conversations are not included in the main part of the analysis of results, but are analyzed separately. Using the same preliminary circuit condition in different experiments establishes some common ground between experiments, but if precise comparisons between results from different experiments are desired, care must be taken to include replications of several standard circuit conditions in each such experiment.

3.4.1.2 *Choice of circuit conditions*

Circuit conditions between which particularly precise comparisons are desired must be included within the same experiment.

Besides this it is necessary that all subjects in every experiment should experience more or less the whole range of performance levels: that is, there should be at least one very good circuit condition, one of near average performance, and one very poor one, while the rest should not all cluster too closely about the same mean opinion score value. If one cannot be confident of this beforehand, it is advisable to carry out first a short informal test on the proposed set of circuit conditions, in order to find out whether the range is in fact covered; if not, the selection of conditions should be modified accordingly, otherwise the subjects' opinion scale will be distorted (the "enhancement" effect). Extra circuit conditions, not in themselves of direct interest to the experimenter, may be added to bring up the number to 9, 11 or 13, and to balance the range of performance more effectively.

Subjects generally expect to experience circuit conditions with various values of overall loss or sensitivity, which of course has a very strong influence on performance, and can be varied to provide the required range of circuit conditions. There are also important interactions between overall sensitivity and many other degradations. It is therefore highly desirable, even if overall sensitivity and its interactions are not the main objects of the investigation, to include some conditions differing from each other only in overall sensitivity.

If the investigation cannot be confined to 15 conditions, it is then spread over several experiments, each concentrating on a well defined part of the inquiry but overlapping the others so as to provide common ground.

3.4.1.3 *Eligibility of subjects*

Subjects taking part in the conversation tests are chosen at random from the Research Centre personnel, with the provisons that:

a) they have not been directly involved in work connected with assessment of the performance of telephone circuits; and

b) they have not participated in any subjective test whatever for at least the previous six months, and not in a conversation test for at least one year.

No steps are taken to balance the numbers of male and female subjects unless the design of the experiment requires it. Subjects are arbitrarily paired in the experimental design prior to the test and remain thus paired for its duration.

3.4.1.4 *Environment*

Subjects are seated in separate sound-proof cabinets near the point from which the experiment is controlled. Room noise is fed in with the required spectrum (usually the Hoth spectrum) at the required level (usually 50 dBA), measured with a Bruel and Kjaer Precision Sound Level Meter type 2206, used with the "*A* weighting" and the "slow" meter characteristic. If different conversations in the same experiment require different room noise levels, then care is taken to prevent the transitions from being too obvious to the subjects: ideally, room noise should be changed only when subjects are out of the sound-proof rooms.

3.4.1.5 *Methods of establishing the connection*

The telephone sets used by the subjects are normal in appearance and feel – usually identical to the standard British Telecom Telephone No. 706, unless the experiment specifically concerns handsets of other types. The means of establishing telephone contact between subjects is made as realistic as possible. The calling subject, on lifting the handset, obtains dialling tone, and has to dial or key a prescribed number to obtain the connection. Ringing tone occurs after a suitable fixed delay, and the other party's bell or tone-caller is operated after a further fixed delay. Wrong numbers are rewarded by the "Number Unobtainable" tone.

3.4.1.6 *Conversation task*

Every effort is made to ensure that conversations are purposeful, and that subjects have full opportunity to exploit the transmission capabilities of the test circuit. A task involving sorting pictures into an order of merit has been found suitable for this purpose and sufficiently interesting to the subjects. The pictures, covering a wide variety of topics, are samples of the standard postcard-sized illustrations offered for sale at several different museums, art galleries and similar institutions. These cards are individually numbered on the back, and assembled arbitrarily into sets of six cards each, every set having an exact duplicate.

The subject is instructed to consider these pictures for display in a public place, and, before each conversation, to arrange the cards of a particular set in his personal order of preference for this purpose; the other subject does the same with his copy of the same set. When contact is established via the test circuit, the subjects have to negotiate an agreed order of preference and write this down at the end of the conversation. The duration of each conversation is thus determined by the subjects themselves. Occasionally a conversation may be very long because both subjects are intensely interested in the pictures, or – as happens in less than 1% of cases – very short because both have independently chosen the same order of preference and have little to discuss, but even in these cases it is highly desirable to allow the subjects to decide for themselves how long to converse. After the end of the conversation they express independent opinions of the connection by marking a form provided: this form is reproduced in Annex A.

Some variations of the task (such as numbering the same pictures differently for the two subjects) have been explored, but none has been definitively adopted. No other type of task has been found to have any greater advantages for the purpose, though several types have been tried.

3.4.1.7 *Preparations for an* n × n *experiment*

From a list of all subjects available, the experimenter randomly chooses a sufficient number of those eligible according to the criteria given in § 3.4.1.3 above. He contacts these by telephone to ask whether they are willing to participate at certain times, which have to be arranged in such a way that subjects who converse together on their first visit remain paired for their subsequent visits in the same experiment. A standard letter is sent to each subject, confirming the time and place of each appointment, and explaining in some detail what will be required of the subjects in the experiment: the text of this letter is reproduced in Annex B.

The experimenter prepares schedules, based on the experimental design, showing in what order conditions must be administered to each pair of subjects, with which picture sets, which party initiates the call in each case, and any other necessary details. Space is left for filling in information that becomes available as the experiment proceeds: consecutive conversation number, duration of conversation, identity of tape reel used for recording, comments about faults or unusual events, and so on. Opinion forms (Annex A) are also prepared for each conversation. However, in order to avoid duplicating or altering too many entries, some items are not filled in until they are certain: for example, the actual names of the subjects are liable to change until they actually arrive for their first visit.

Both in the letter and in any discussions with the subjects, great care is taken not to communicate to the subjects any knowledge about the nature of the circuit conditions. The opinion forms do not even carry any number or code identifying the circuit condition – this information is obtained from the schedule and added to the forms after they have been collected from the subjects.

3.4.1.8 *Procedure*

When subjects arrive for their first visit, they are asked whether they have read and understood the letter. Any obscurities are clarified, and opportunity is given for asking questions. The sound-proof rooms and their facilities are demonstrated. Subjects are informed how many calls will be comprised in this visit. Forms are handed to the subjects, and they are then left to prepare for the preliminary conversation. On subsequent visits the subjects are merely informed that the procedure will be the same as before, with possibly a different number of calls.

At the beginning of each conversation, the subjects take out the specified picture set from a box on the desk, arrange the pictures in order of preference, and fill in the appropriate part of the opinion form. When both subjects have done this, the experimenter gives one of them the signal to initiate the call. The subjects are then completely free to determine the course of the conversation, except that they must not discuss their opinions of the connection. When they have written down their agreed order of preference for the pictures, terminated the conversation, and recorded their scores (Excellent, Good, Fair, Poor or Bad) and their answer to the "Difficulty" question (Yes or No), the experimenter contacts each in turn by telephone to ask what answer he has given to the "Difficulty" question; if the answer is "Yes", the experimenter asks the subject to explain briefly (in his own words) the nature of the difficulty. The reply is noted, but neither the subject nor the experimenter is expected to attempt precise formulations: it is essential not to prompt the subjects, and in any case the classification of difficulty has been found far less useful than the undifferentiated percentage "Difficulty" itself.

After this the experimenter requests the subject to put away the form in an envelope provided, and then tells him to start sorting out the next set of pictures, or, as the case may be, to wait to be released from the sound-proof room.

Both the conversations between subjects and the conversations between experimenter and subject are tape-recorded.

3.4.1.9 *Treatment of results*

The results from each conversation comprise two opinions on the scale Excellent-Good-Fair-Poor-Bad (scored respectively 4, 3, 2, 1, 0), two votes on the Difficulty scale (scored 1 = Yes, 0 = No), two speech levels (measured from tape recordings) and one value of duration. In particular cases information may be collected about other variables also; for example, video recordings may be made in order to observe how subjects hold their handsets.

Analysis of variance is applied separately to each variate (opinion score, speech level, etc.) in order to test the significance of circuit-condition features and other effects, and to find confidence intervals for the means. With a binary variate like "Difficulty" this process must be regarded with some reservations. There is usually less scope for curve-fitting than in listening experiments, simply because there are far fewer pairs of coordinate values available.

3.5 *Quantal-response detectability tests*

The best method for obtaining information on the detectability or some analogous property of a sound (such as echo), as a function of some objective quantity (such as listening level), is a quantal-response method similar in principle to that mentioned in § 3.1 above for loudness balancing. The main difference is that the subject's response is not a decision in the form "Reference" or "Test" (the designation of the louder of two circuits), but a vote on a scale such as:

Opinion scale 6A

A Objectionable
B Detectable
C Not detectable

where B is understood to mean "Detectable but not objectionable".

Scales of this sort, usually with three points, may be used in a variety of quantal-response tests; for example the scale as shown above may be used where the stimulus is echo, reverberation, sidetone, voice-switching mutilation, or interfering tones, while crosstalk and perhaps echo in some circumstances may be judged on the scale Intelligible – Detectable – Not detectable.

It is sometimes permissible to regard these votes as opinion scores, with values 2, 1, 0 respectively, and treat them in the same sort of way as one would treat listening or conversation opinion scores. But this is often unsatisfactory because the decisions on such a scale as 6A are not really equivalents of responses on a continuous scale – as votes on such scales as 4A may be legitimately taken to be – but effectively embody two distinct dichotomies (for example detectable/not detectable and objectionable/not objectionable), which though not independent may nevertheless call different psychological processes into action: in other words, Objectionability or Intelligibility differs in kind, not merely in degree, from Detectability. For this reason a more profitable method of analysis is to express the probability of response according to each dichotomy separately, as a function of some objective variable, by fitting probit or logit equations, and then using the quantiles or other parameters as a basis of comparison between circuit conditions, in a manner analogous to that used in applying articulations scores.

The actual conduct of experiments of this type resembles that of listening-effort tests (see § 3.3.1 above), but there are some differences. In particular it is advisable that the first presentation of the signal in each run should be at a high listening level, so that the listener is left in no doubt what kind of signal is a candidate for his decisions. Where sidetone or echo is involved, the subject will be required to talk as well as listen.

Simple audiometric measurements, as described in Recommendation P.78, are usually performed on subjects who participate in these experiments, so that results can be expressed relative to their threshold of hearing.

For examples of the application of these techniques, see References [5] and [6].

Noise and other disturbances are sometimes investigated by means of responses on a scale with many more points; for example, Opinion scale 5 with seven points ranging from "Inaudible" to "Intolerable". These scales are more nearly of the quantized-continuum type, like Opinion scale 4A, and can be treated similarly.

4 References to Recommendations and other CCITT publications relying on Methods a) to e) under § 2 above:

a) Many Recommendations include requirements based on reference equivalents of which Recommendations P.12, G.101 [7], G.103 [8], G.111 [9], G.120 [10] and G.121 [11] are examples.

b) Recommendation P.12 requires certain articulation values to be satisfied but the method is now mainly used for diagnostic purposes. See Recommendation P.45.

c) Study of various Questions, for example see Question 4/XII [12], Question 14/XII [13] and Supplement No. 4 at the end of this fascicle.

d) Study of various Questions, for example see Question 4/XII [12], Question 9/XII [14], Question 14/XII [13] and Supplement No. 4 at the end of this fascicle.

e) Study of various Questions, for example see Question 9/XII [14] and References [15], [16] and [17].

5 General comments on subjective methods used in the laboratory

More detailed information on the conduct of subjective tests and interpretation of their results are given in Recommendation P.74 and Reference [2]. A rather broad survey of the relationship between various methods is given in Reference [18].

When used to provide information to assist in transmission planning of telephone networks, subjective methods should be employed with the following considerations in mind:

a) A clear description must be available of the type of telephone connections to which the results are to be applied. This is provided by formulating appropriate hypothetical reference connections (HRCs) (see Recommendation G.103 [8]).

b) The levels, transmission losses, sending and receiving reference equivalents, etc., of the HRCs must guide the establishment of laboratory arrangements and the conduct of the tests. Speech spectra and levels must be properly chosen to correspond to those at the various points in the HRC.

c) Subjects must be drawn from an appropriate population. For example, if audiograms are obtained from subjects participating in a conversation experiment, this information should not be used to reject any subjects, because the resultant bias in the sample would make the conclusions applicable only to users with a certain range of hearing sensitivity. For this reason it is safest to collect auxiliary information of this type only after the subjects have finished their main task.

d) Subjects must be treated within the experiments so that the results obtained are valid for the desired applications. This is the reason for taking the precautions described above (§ 3.3.1) to ensure that subjects' judgements are not distorted by the range of conditions and levels chosen, or by the order of presentation; and to make the procedure in conversation tests (§§ 3.4.1.5 to 3.4.1.7) natural yet standardized.

e) Suitable experimental designs must be used so that the results can be properly analyzed and confidence intervals estimated.

f) Uncontrolled variation in some feature of the transmission path is sometimes unavoidable: for example the requirement may be to conduct a listening test over a fading radio link, or a conversation test over a TASI link with freeze-out determined by real traffic. In such cases it is advisable to collect not only the subjects' responses but also contemporary information on the values of the related fluctuating quantities: signal strength in the first case, freeze-out fraction or number of channels occupied in the second. The technique known as analysis of covariance (Reference [19]) is the appropriate method for processing this information on concomitant variables, as they are called, in conjunction with the responses (main variables).

g) Even with proper precautions under c), d), e) and f), reliance should not be placed on absolute values of scores unless "control" conditions (e.g. a set of reference conditions) are included within the experiment. However, relativities between scores obtained from different circuit conditions within the same experiment are more reliable.

h) A set of reference conditions will make it possible to express results as ratings in terms of equivalent settings of some reference device – attenuator, noise source, modulated noise reference unit (see Recommendation P.70), etc. This enables much more reliable comparisons to be made with information from other sources.

i) Results of subjective experiments should always be reviewed for internal consistency and compared with expected results (derived from previous experience or from a theoretical model) before being applied.

6 Objective methods

Clearly the ultimate aim must be to attain the capability of assessing telephony transmission performance purely in terms of the objective characteristics of the telephone connections concerned. This aim is partly satisfied by use of tabulated information based on previous laboratory and other tests: an example of such usage appears in Reference [20]. Considerable progress has now been made towards the prediction of assessment scores, speech levels, etc. by use of subjective modelling as described in Supplement No. 4 at the end of this fascicle and Reference [21]. British Telecom is now updating its tabulated information using this method.

The modelling technique makes it possible to treat many other important features like attenuation/frequency distortion and sidetone in a much more general manner. For example, by making due allowance for the part played by high sidetone level – which is a very potent degradation in connections of poor transmission performance – it makes clear why sensible limits for overall loss and noise cannot be fixed without regard to sidetone suppression.

ANNEX A

(to Supplement No. 2)

Opinion form 12A

Test ____________________

Name ____________________ Cabinet ____________________ No. ____________________

1 Before starting your call, please take out picture set ________ and arrange the cards in order of preference. Record this order in the boxes below, using the numbers on the backs of the cards for identification.

	1st	2nd	3rd	4th	5th	6th
Your Order of Preference						

2 If you receive a green GO AHEAD signal, then call your partner on ________. Otherwise wait for your partner to call you.

3 You may enter your partner's picture-card order here if you find this helps you in the discussion.

	1st	2nd	3rd	4th	5th	6th
Partner's Order of Preference						

4 When you have arrived at an agreed order of preference, please enter it here.

	1st	2nd	3rd	4th	5th	6th
Agreed Order of Preference						

Then replace your handset.

5 Please mark, with a cross, your opinion of the telephone connection you have just been using. *N.B.* – Please do not discuss your opinion with your partner.

Excellent	Good	Fair	Poor	Bad

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

YES	
NO	

If the answer is YES, please explain briefly what the difficulty was when the operator contacts you again.

FOR R13.4 USE														

ANNEX B

(to Supplement No. 2)

(Standard letter sent to subjects)

Name ____________________

Group ____________________

R13.4 SUBJECTIVE TEST No. ____________________

Thank you for agreeing to take part in this experiment.

As arranged earlier by telephone, we should like you to come to Room _______________, Floor 3, Main Laboratory Block, at the following times.

Time Day Date

On arrival, ask for _________________, quoting the above subjective test number. You will be reminded by telephone shortly before each visit is due. You may book your time to project _________________. If you cannot keep an appointment, or if you need further information, contact _________________ on Ipswich 64 _________________.

The experiment to which you have been invited forms part of a series concerned with the transmission performance of telephone connections. You will be asked to converse with another volunteer over particular telephone connections, and it is hoped that the tasks we shall give you will lead to vigorous conversations devoted to discussion and negotiation.

In the test room you will be provided with a set of six picture cards. You are asked to imagine that you and your partner are responsible for choosing some of these (enlarged if necessary) to be displayed in a public place such as the Staff Restaurant – either as items of general interest or simply as decoration. Before each call you should arrange all six cards in your order of preference, and write the six identification numbers in this order on the form provided. Your partner, in another room, will have an identical set of pictures, and his order of preference will probably be different from yours. One of you will then be requested to make a telephone call to the other. The aim of the ensuing conversation will be to negotiate with your partner so as to arrive at a compromise order which satisfies you both. At the end of the conversation you should replace the handset and enter the six numbers in the finally agreed order on the form. You must also mark the appropriate box to indicate your opinion of the connection. After this the operator will contact you and tell you what to do next. Subsequent conversations will be similar, but with different sets of pictures.

In the whole experiment there will be a total of _______ calls spread over the _______ visit(s) arranged as above. Full instructions will be given when you arrive. Please bring this letter with you, and also your glasses if you normally wear any.

Thank you once again for your co-operation.

____

____________________ (date)

References

[1] CCITT – Question 15/XII, Contribution COM XII-No. 1, Study Period 1981-1984, Geneva, 1981.

[2] RICHARDS (D. L.): Telecommunication by speech, *Butterworths*, London, 1973.

[3] CCITT – Question 2/XII, Contribution COM XII-No. 1, Study Period 1981-1984, Geneva, 1981.

[4] SULLIVAN (J. L.): Is transmission satisfactory? Telephone customers help us decide, *Bell Labs Record*, March 1974, pp. 90-98.

[5] RICHARDS (D. L.): Telecommunications by speech, *Butterworths*, London, 1973, § 3.5.3.

[6] *Ibid.*, § 4.5.1.

[7] CCITT Recommendation *The transmission plan*, Vol. III, Rec. G.101.

[8] CCITT Recommendation *Hypothetical reference connections*, Vol. III, Rec. G.103.

[9] CCITT Recommendation *Corrected reference equivalents (CREs) in an international connection*, Vol. III, Rec. G.111.

[10] CCITT Recommendation *Transmission characteristics of national networks*, Vol. III, Rec. G.120.

[11] CCITT Recommendation *Corrected reference equivalents (CREs) of national systems*, Vol. III, Rec. G.121.

[12] CCITT – Question 4/XII, Contribution COM XII-No. 1, Study Period 1985-1988, Geneva, 1985.

[13] CCITT – Question 14/XII, Contribution COM XII-No. 1, Study Period 1985-1988, Geneva, 1985.

[14] CCITT – Question 9/XII, Contribution COM XII-No. 1, Study Period 1981-1984, Geneva, 1981.

[15] RICHARDS (D. L.) and BUCK (G. A.): Telephone echo tests, *P.I.E.E.*, 1960, 107B, pp. 553-556.

[16] CCITT – Contribution COM XII-No. 171, Study Period 1977-1980, Geneva, 1979.

[17] CCITT – Contribution COM XII-No. 132, Study Period 1977-1980, Geneva, 1979.

[18] CCITT – Question 7/XII, Annex 1, Contribution COM XII-No. 1, Study Period 1981-1984, Geneva, 1981.

[19] SNEDECOR (G. W.) and COCHRAN (W. G.): Statistical methods, Chapter 14, 6th edition, *Iowa State University Press*, 1967.

[20] CCITT – Contribution COM XII-No. 173, Study Period 1977-1980, Geneva, 1979.

[21] RICHARDS (D. L.): Calculation of opinion scores for telephone connections, *Proc. IEE*, 121, pp. 313-323, 1974.

Bibliography

BRAUN (K.): Die Bezugsdämpfung und ihre Berechnung aus der Restdämpfungskurve (Frequenzkurve) eines Übertragungssystems; *T.F.T.*, Vol. 28, pp. 311-318, August 1939.

BRAUN (K.): Theoretische und experimentelle Untersuchung der Bezugsdämpfung und der Lautstärke; *T.F.T.*, Vol. 29, pp. 31-37, No. 2, 1940.

BLYE (P. W.), COOLIDGE (O. H.) and HUNTLEY (H. R.): A revised telephone transmission rating plan; *B.S.T.J.*, Vol. 34, pp. 453-472, May 1955 (reproduced in the *Red Book*, Vol. I, pp. 636-651, ITU, Geneva, 1957, and Vol. V, pp. 607-624, ITU, Geneva, 1962).

BRAUN (K.): Image attenuations of microphone and receiver insets; *N.T.Z.*, No. 8, pp. 365-370, 1960 (translated in the *Red Book*, Vol. V *bis*, pp. 255-265, ITU, Geneva, 1965).

FRENCH (N. R.) and STEINBERG (J. C.): Factors governing the intelligibility of speech sounds; *J.A.S.A.*, Vol. 19, p. 89, Jan. 1947.

RICHARDS (D. L.) and ARCHBOLD (R. B.): A development of the Collard principle of articulation calculation; *P.I.E.E.*, Vol. 103, Part B, Sept. 1956 (*Red Book*, Vol. I, Question 7 of Study Group 12, Annex 4, ITU, Geneva, 1956).

Contribution by the Italian Administration to the study of objective methods for measuring reference equivalent and articulation reference equivalent, *Red Book*, Vol. I, Question 7 of Study Group 12, Annex 3, ITU, Geneva, 1956.

FLETCHER (H.) and GALT (R. H.): The perception of speech and its relation to telephony; *J.A.S.A.*, Vol. 22, p. 89, March 1950 (reproduced in the following work, Chapters 15-17).

FLETCHER (H.): Speech and hearing in communication, *D. Van Nostrand*, New York, 1953.

Tonality method studied by the U.S.S.R. Administration to determine articulation; *Red Book*, Vol. V, Part II, Annex 31, ITU, Geneva, 1962.

Method used by the Swiss Telephone Administration for the determination of transmission quality based on objective measurements; *Red Book*, Vol. V, Part II, Annex 30, ITU, Geneva, 1962.

LALOU (J.): Calculation of telephone transmission performance by information theory, *Red Book*, Vol. V *bis*, Question 7/XII, Annex 2, ITU, Geneva, 1965.

SIVIAN (L. J.): Speech power and its measurement, *B.S.T.J.*, 8, pp. 646-661, 1929.

LOYE (D. P.) and MORGAN (K. F.): Sound picture recording and reproducing characteristics, *J. Soc. Motion Picture Engineers*, 32, pp. 631-647, 1939.

RICHARDS (D. L.): Some aspects of the behaviour of telephone users as affected by the physical properties of the circuit. Communication Theory, *Butterworths Scientific Publications*, pp. 442-449, 1953.

ZAITSEV (T. E.): Correlation method for determining the fidelity and intelligibility of speech transmitted over telecommunication channels, *Elektrosvyaz*, 10, pp. 38-46, 1958.

LICKLIDER (J. C. R.), BISBERG (A.) and SCHWARZLANDER (H.): An electronic device to measure the intelligibility of speech, *Proc. Nat. Electronics Conf.*, 15, pp. 329-334, 1959.

RICHARDS (D. L.) and SWAFFIELD (J.): Assessment of speech communication links, *P.I.E.E.*, 106B, pp. 77-89, 1959.

RICHARDS (D. L.): Conversation performance of speech links subject to long propagation times, International Conference on Satellite Communication, *Inst. Elec. Engrs*., pp. 247-251, London, 1962.

RICHARDS (D. L.): Transmission performance of telephone connections having long propagation times, *Het PTT-Bedrijf*, 15, pp. 12-24, 1967.

BOERYD, (A.): Subscriber reaction due to unbalanced transmission levels, *ibid*, pp. 39-43.

RICHARDS, (D. L.): Distortion of speech by quantizing, *Electronics Letters*, 3, pp. 230-231, 1967.

GOLDMAN-EISLER (F.): Sequential temporal patterns and cognitive processes in speech, *Language and Speech*, 10, pp. 122-132, 1967.

Supplement No. 3

TRANSMISSION RATING MODELS

(Geneva, 1980; modified at Malaga-Torremolinos, 1984)

(Quoted in § 3 of Recommendation P.11)

(Contribution by the Bell Communications Research, Inc. [1])

1 Introduction

This Supplement describes transmission rating models which can be used to estimate the subjective reaction of telephone customers to the transmission impairments of circuit noise, corrected reference equivalent, talker echo, listener echo, attenuation distortion (including bandwidth), quantizing distortion, room noise and sidetone.

The models for circuit noise, corrected reference equivalent (CRE) and talker echo are based on several conversational tests conducted at Bell Laboratories in the period from 1965 to 1972 to evaluate the subjective assessment of transmission quality as a function of circuit noise, corrected reference equivalent, talker echo path loss and talker echo path delay [1]. These tests involved several hundred subjects and several thousand test calls, Several tests were conducted on normal business calls. Others were conducted in the laboratory. All of the tests employed a 5-category rating scale: excellent, good, fair, poor and unsatisfactory.

The essential features of the models were originally derived in terms of loudness loss of an overall connection *in dB* (as measured by the Electro-Acoustic Rating System, *EARS*) and circuit noise *in dBmp* at the input to a reference receiving system (electric-to-acoustic efficiency as measured by the *EARS*) [2]. The effects of talker echo were later incorporated in terms of loudness loss of the echo path *in dB* (as measured by the *EARS*) and round trip delay of the echo path *in milliseconds.* Experimentally determined correction factors were used to convert the models to a CRE basis [3].

[1] Supplement No. 3 was formerly a contribution by the American Telephone and Telegraph Company. (See CCITT *Yellow Book* Volume V.) Amendment of the supplement reflects in part work performed at AT&T Bell Laboratories prior to 1 January 1984.

The original model for listener echo was based on a series of four listening-type subjective tests conducted at Bell Laboratories in 1977 and 1978 [4]. Subsequent test results led to an alternative form of the model [5], [6]. The subjective tests included conditions in which the listener echo path loss was flat or frequency-shaped by selective filtering. A weighted echo path loss is defined to provide a weighting of the frequency-shaped test conditions so that subjectively equivalent test conditions have the same transmission rating.

The model for quantizing distortion is based on a series of five subjective tests conducted to evaluate the performance of various digital codec algorithms [7], [8], [9].

The model for bandwidth and attenuation distortion is based on tests conducted in 1978 [10].

The model for room noise is based on unpublished tests conducted in 1976. Opinion ratings of transmission quality on a five-category scale were made by 40 subjects for 156 conditions having various combinations of room noise, speech level, circuit noise and sidetone path loss. The samples of room noise were presented from tape recordings made in an airlines reservations office. A model was fitted to the test results in terms of the circuit noise which produced the same quality ratings as given levels of room noise.

The model for sidetone is based on tests conducted in 1980 [11].

All of the tests were conducted with Western Electric 500-type telephone sets or equivalent. The procedures used in the analysis of the subjective tests results and the derivation of the transmission rating scale are outlined in Reference [1]. Although the procedures are somewhat complex for manual calculation, they are easily handled on a digital computer and have been found to provide a convenient and useful representation for a large variety of test data.

The models incorporate the concept of a transmission rating scale. An important reason for the introduction of this scale was the recognition that subjective test results can be affected by various factors such as the subject group, the type of test, and the range of conditions which are included in the test. These factors have been found to cause changes in both the mean opinion score of a given condition and in the standard deviation. Thus, there are difficulties in trying to establish a unique relationship between a given transmission condition and subjective opinion in terms of mean opinion score or percent of ratings which are good or excellent. The introduction of a transmission rating scale tends to reduce this difficulty by separating the relationship between transmission characteristics and opinion ratings into two parts. The first part, the transmission rating as a function of the transmission characteristic, is anchored at two points and tends to be much less dependent on individual tests. The second part, the relationship between the transmission rating and subjective opinion ratings, can then be displayed for each individual test.

The transmission rating scale for corrected reference equivalent and circuit noise was derived such that it is anchored at two points as shown in Table 1.

TABLE 1

Overall corrected reference equivalent (dB)	Circuit noise (dBmp) a)	Transmission rating
16	−65	80
31	−50	40

a) The circuit noise values are referred to a receiving system with a receiving CRE = +1 dB.

These anchor points were selected to be well separated but within the range of conditions which are likely to be included in a test. The rating values are such that most connections will have positive ratings between 40 and 100. Transmission ratings for other combinations of reference equivalent and circuit noise are relative to those for these two anchor points.

This Supplement presents the transmission rating models in terms of corrected reference equivalent of an overall connection in dB, circuit noise in dBmp referred to the input of a receiving system with a receiving CRE = +1 dB, corrected reference equivalent of the talker echo path in dB, and round-trip delay of the talker echo path in milliseconds. Annex A illustrates representative opinion results.

2 Transmission rating models

2.1 *Overall corrected reference equivalent and circuit noise*

The transmission rating model for overall corrected reference equivalent and circuit noise is

$$R_{LN} = -34.88 - 2.257\sqrt{(L'_e - 8.2)^2 + 1} - 2.0294\ N'_F + 1.833\ L'_e + 0.02037\ L'_e N'_F \qquad (2\text{-}1)$$

L'_e is the CRE of an overall telephone connection (in dB).

N'_F is the total effective noise (in dBmp) referred to a receiving system with a +1 dB receiving CRE. The total effective noise is obtained by the power addition of the circuit noise, N'_c, the circuit noise equivalent, N'_{Re}, of the room noise and the circuit noise equivalent, N'_{Qe}, of the quantizing noise.

N'_c is the circuit noise (in dBmp) referred to a receiving system with a +1 dB receiving CRE.

N'_{Re} is the circuit noise equivalent (in dBmp) of the room noise referred to a receiving system with a +1 dB receiving CRE. (See § 2.2.)

N'_{Qe} is the circuit noise equivalent (in dBmp) of the quantizing noise referred to receiving system with a 1 dB receiving CRE. (See § 2.3.)

Transmission rating as a function of overall CRE and circuit noise is shown in Figure 1. This figure uses a value of $N'_{Re} = -62.63$ dBmp. Bandwidth factor, k_{BW}, defined in § 2.4 is equal to unity.

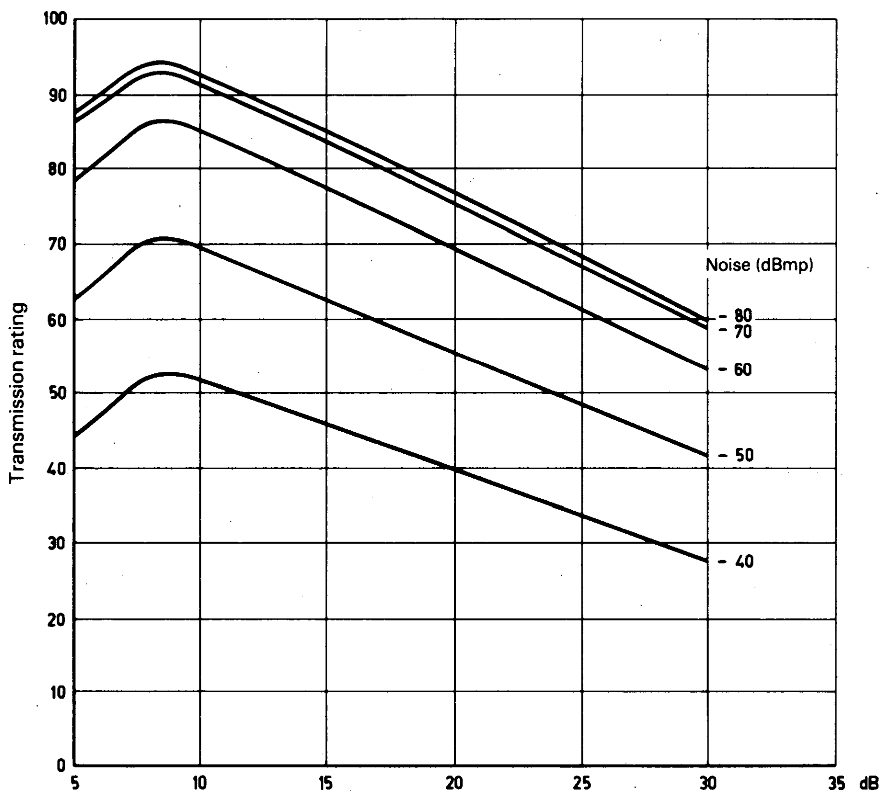


FIGURE 1

Transmission rating for overall CRE and circuit noise

2.2 *Circuit noise equivalent of the room noise*

The transmission rating model for the circuit noise equivalent, N'_{Re} (in dBmp), of the room noise is

$$N'_{Re} = N_R - 125 + 0.0078 (N_R - 35)^2 + 10 \log_{10} \left[1 + 10^{\frac{7 - L'_s}{10}}\right] \tag{2-2}$$

where

N_R is the room noise in dB(A) at the listening end

L'_s is the sidetone reference equivalent (in dB) of the listening end telephone set sidetone path

The circuit noise equivalent, N'_{Re}, is plotted as a function of room noise in Figure 2.

Note – The transmission rating model for reference equivalent and circuit noise is normally used with

$$N'_{Re} = -62.63 \text{ dBmp.} \tag{2-3}$$

This value was determined from analysis of the conversational tests results from which the transmission rating model for the overall corrected reference equivalent and circuit noise was originally formulated.

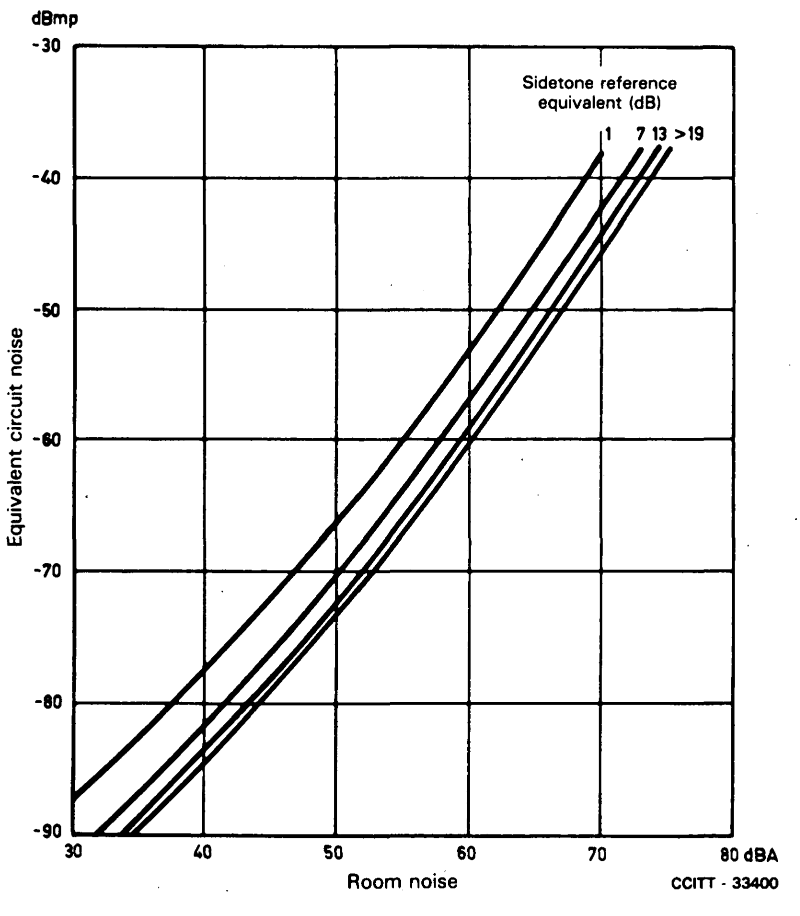


FIGURE 2

Equivalent circuit noise for room noise

2.3 *Circuit noise equivalent of quantizing noise*

The transmission rating model for the circuit noise equivalent N'_{Qe} (in dBmp) of quantizing noise is

$$N'_{Qe} = V_0 - 1 - \text{SNR} \tag{2-4}$$

where

V_0 is the received speech level (in VU) referred to a receiving system with a +1 dB receiving CRE[2)],

and

SNR is the signal-to-circuit noise ratio (in dB) which is judged to provide speech quality equivalent to the speech-to-speech correlated noise ratio, Q (in dB), as determined by a Modulated Noise Reference Unit (see Recommendation P.70).

SNR can be approximated by

$$\text{SNR} = 2.36\ Q - 8.34 \tag{2-5}$$

from which

$$N'_{Qe} = V_0 - 2.36\ Q + 7.34. \tag{2-6}$$

Based on a 1975-1976 Speech Level Survey, [12] the speech level for domestic North American connections can be approximated by

$$V_0 = -14 - L'_e. \tag{2-7}$$

Estimates of Q for single codec pairs are given below for Pulse Code Modulation (PCM), Nearly-Instantaneous Compandored modulation (NIC), Adaptive Differential Pulse Code Modulation (ADPCM) and Adaptive Delta Modulation (ADM). They apply to the particular algorithms described in References [7] and [9].

PCM: $Q = 0.78\ L - 12.9$ (2-8)

NIC: $Q = 0.74\ L - 2.8$ (2-9)

ADM: $Q = 0.42\ L + 8.6$ (2-10)

ADPCM: $Q = 0.98\ L - 5.3$ (2-11)

ADPCM-V: $Q = 1.04\ L - 4.6$ (2-12)

where

L is the line bit rate in kbit/s.

Note – The ADPCM algorithm with fixed predictor is described in Reference [13]. The ADPCM-V algorithm with adaptive predictor is described in Reference [9].

For connections with tandem codec pairs, the total Q can be estimated as follows:

$$Q = -15 \log_{10} \left[\sum_{i=1}^{n} 10^{-\frac{Q_i}{15}} \right] \tag{2-13}$$

2.4 *Bandwidth and attenuation distortion*

The transmission rating model for corrected reference equivalent and circuit noise can be modified to include the effects of bandwidth (and attenuation distortion). The transmission rating, R_{LNBW}, for corrected reference equivalent, circuit noise and bandwidth is

$$R_{LNBW} = (R_{LN} - 22.8)\ k_{BW} + 22.8 \tag{2-14}$$

where

$$k_{BW} = k_1\ k_2\ k_3\ k_4 \tag{2-15}$$

with

$$k_1 = 1 - 0.00148\ (F_l - 310) \tag{2-16}$$

$$k_2 = 1 + 0.000429\ (F_u - 3200) \tag{2-17}$$

---

2) V_0 is typically equal to the speech power while active, plus 1-2 dB.

$$k_3 = 1 + 0.0372\,(S_l - 2) + 0.00215\,(S_l - 2)^2 \qquad (2\text{-}18)$$

$$k_4 = 1 + 0.0119\,(S_u - 3) - 0.000532\,(S_u - 3)^2 - 0.00336\,(S_u - 3)\,(S_l - 2) \qquad (2\text{-}19)$$

and

F_l, F_u is the lower and upper band limits (in Hz) at which the acoustic-to-acoustic response is 10 dB lower than the response at 1000 Hz. (For $F_u > 3200$ Hz, a value of 3200 Hz should be used.)

S_l, S_u is the lower and upper inband response slopes (in dB/octave) below and above 1000 Hz, respectively, which would have the same loudness loss as the actual response shapes.

Figures 3 and 4 illustrate the effect of the band limits, F_l and F_u, and inband slopes, S_l and S_u, on the Bandwidth Factor, k_{BW}.

Note – The functions for the bandwidth factor, k_{BW}, have been selected such that $k_{BW} = 1$ when $F_l = 310$ Hz, $F_u = 3200$ Hz, $S_l = 2$ dB/octave and $S_u = 3$ dB/octave. These response characteristics are representative of those used in the tests to formulate the transmission rating model for corrected reference equivalent and circuit noise.

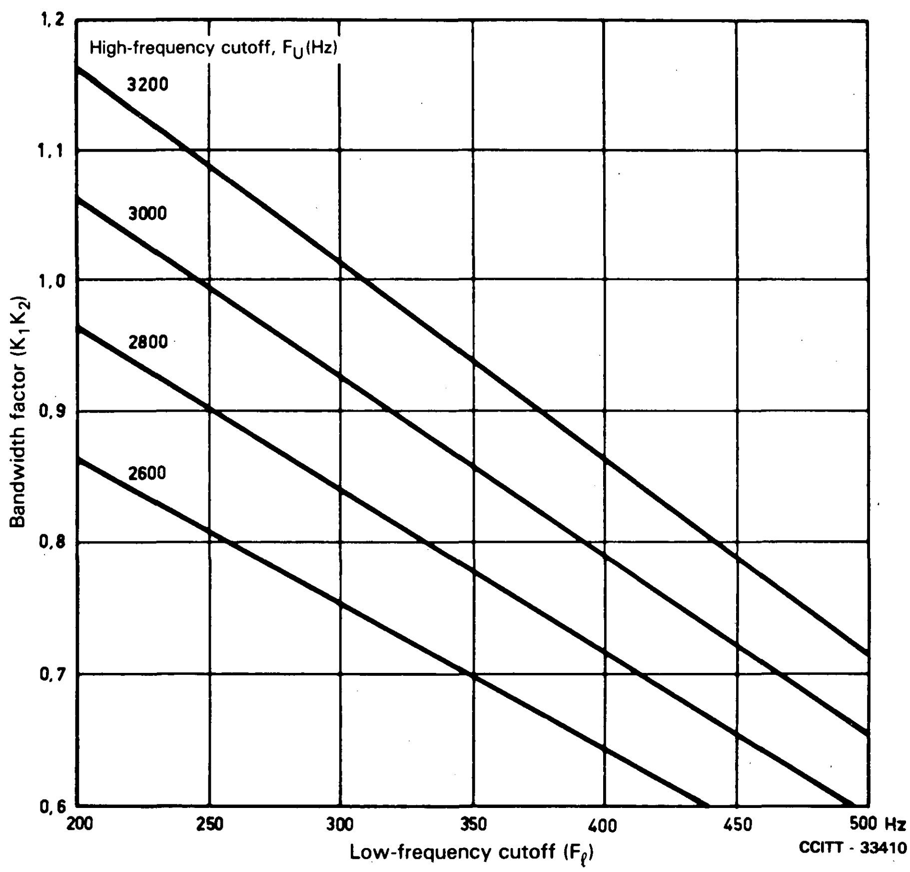


FIGURE 3

Bandwidth model factor

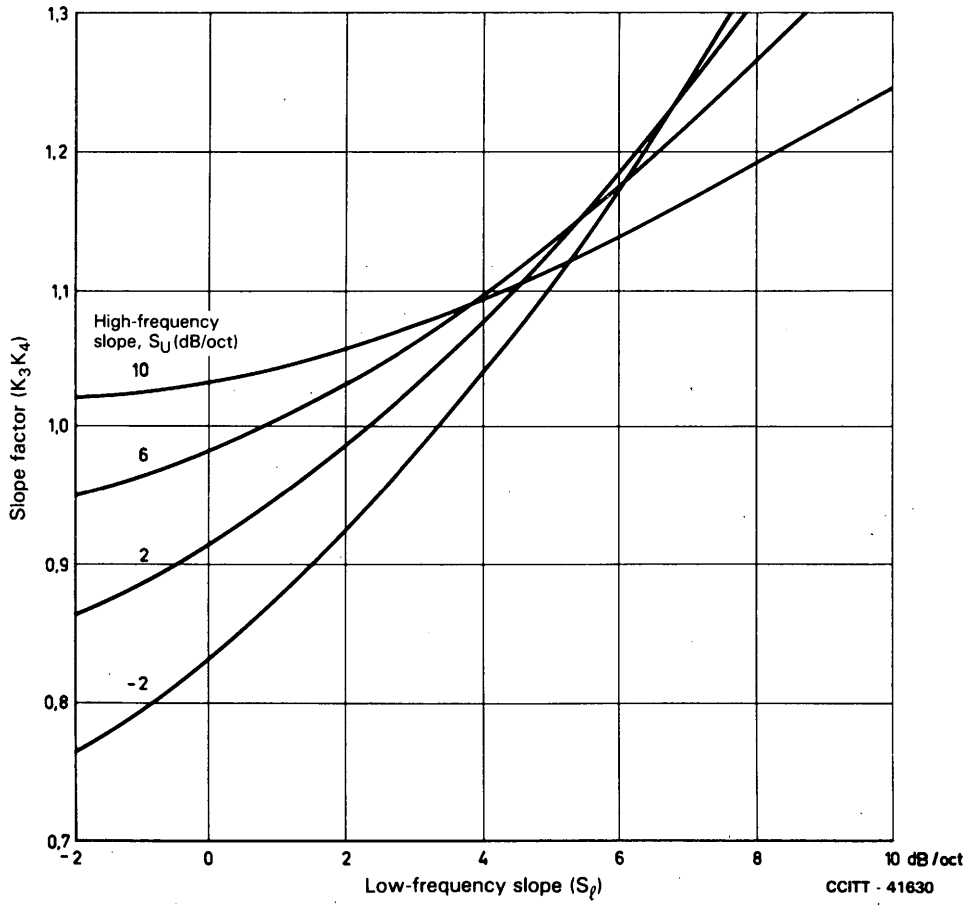


FIGURE 4

Attenuation-distortion model slope factors

2.5 *Listener echo*

The transmission rating model for listener echo is

$$R_{LE} = 9.3\ (WEPL + 7)\ (D_L - 0.4)^{-0.229} \tag{2-20}$$

where

WEPL is the Weighted Listener Echo Path Loss (in dB) and

$$WEPL = -20 \log_{10} \frac{1}{3200} \int_{200}^{3400} 10^{-\frac{EPL(f)}{20}}\, \mathrm{d}f \tag{2-21}$$

EPL(f) is the echo path loss (in dB) as a function of frequency in Hz.

D_L is the round-trip listener echo path delay in milliseconds.

Transmission rating, R_{LE}, as a function of the weighted echo path loss and listener echo-path delay is shown in Figure 5.

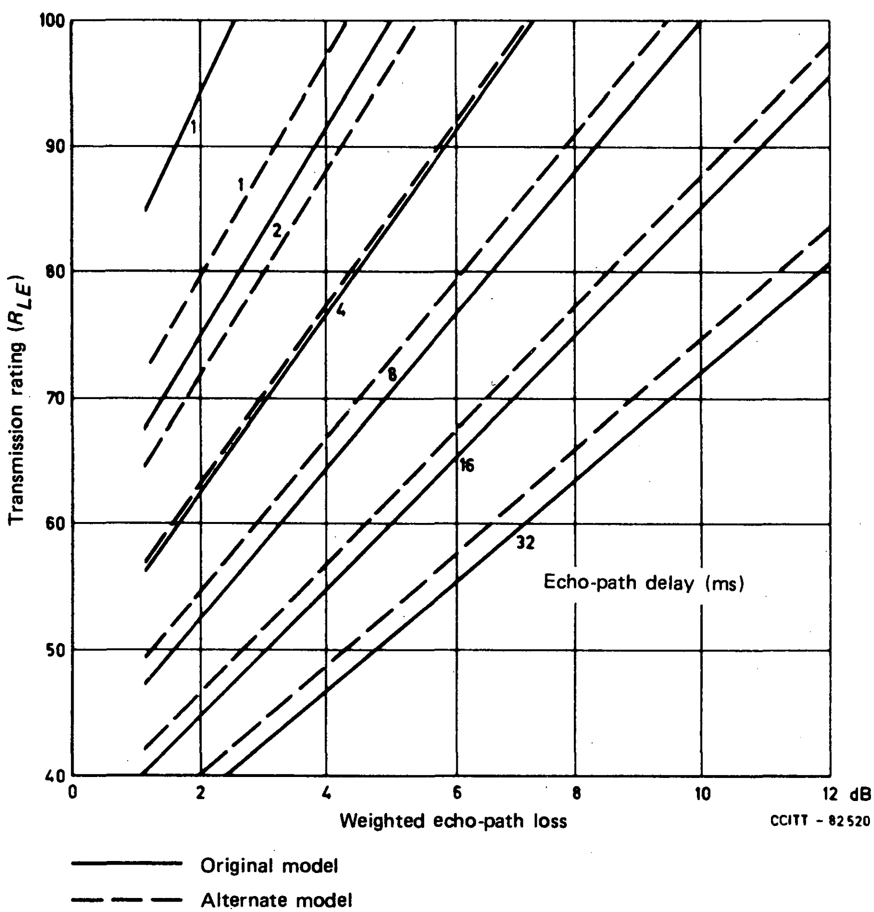


FIGURE 5

Transmission rating for listener echo

The transmission rating for listener echo, R_{LE}, can be combined with the transmission rating for corrected reference equivalent and circuit noise to give an overall transmission rating as follows:

$$R_{LNLE} = \frac{R_{LN} + R_{LE}}{2} - \sqrt{\left[\frac{R_{LN} - R_{LE}}{2}\right]^2 + 13^2} \tag{2-22}$$

Figure 6 provides curves generated by means of the above relationship for transmission rating as a function of weighted listener echo path loss and listener echo path delay in a connection with an overall corrected reference equivalent of 16 dB and a circuit noise of -60 dBmp referred to a receiving CRE of $+1$ dB.

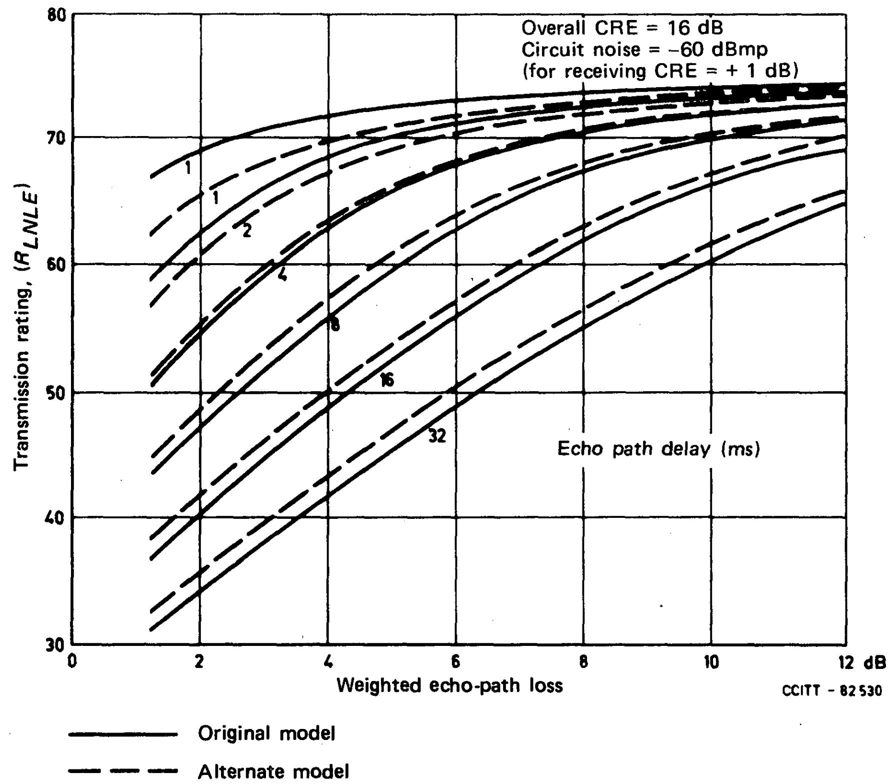


FIGURE 6

Transmission rating for overall CRE, circuit noise and listener echo

Note – The preceding material is based on the use of a specific set of test results and the listener echo model of Reference [4]. Subsequently, new test results were reported in References [5] and [6] which also described studies of the two sets of tests results to see if a single model could be recommended. In general, the agreement between the two sets of results was good. However, the newer results had lower opinion ratings at delays less than about 3 ms. A conservative approach was to revise the original model to provide lower ratings at low delays while retaining the more critical predictions at higher values of delay. The following equation (2-20a) provides a satisfactory replacement for equation (2-20) which accomplishes this goal.

$$R_{LE} = 10.5\,(WEPL + 7)\,(D_L + 1)^{-0.25} \tag{2-20a}$$

Reference [6] also proposed that Weighted Echo Path Loss (WEPL) in the original model be replaced by Scaled Weighted Echo Path Loss (SWEPL). The proposal defined

$$WEPL = SM + SF$$

where

SM is the singing margin,

SF is the shape factor

and then defined

$$SWEPL = SM + SF\,\frac{SM}{1 + SM}$$

Hence, like $WEPL$,

$$SWEPL = SM, \text{ if } SF = 0.$$

Also,

$$SWEPL \approx WEPL, \text{ for } SM \gg 1.$$

The effect of the shape factor is reduced as SM approaches zero. Thus, the shape effect is cut in half when SM is equal to unity, and approaches zero as SM approaches zero. This avoids the possibility of a positive $SWEPL$ when singing margin has become negative. Although the use of $SWEPL$ instead of $WEPL$ will cause little change in most practical situations with typical values of SM, the concept is attractive in forcing the singing margin to be specifically taken into account and is easily accomplished by replacing $WEPL$ by $SWEPL$ in equation (2-20a).

2.6 *Talker echo*

The transmission rating model for talker echo is

$$R_E = 83.62 - 53.45 \log_{10} \left[\frac{1 + D}{\sqrt{1 + \left(\frac{D}{480}\right)^2}} \right] + 2.277\ E \tag{2-23}$$

where

E is the CRE (in dB) of the talker echo path

D is the round-trip talker echo path delay in milliseconds.

Transmission rating as a function of talker echo path loss and delay is shown in Figure 7 and has been derived to exclude the effects of circuit noise and overall CRE. Transformation of the talker echo test results, which included selected values of overall CRE and circuit noise, to the transmission rating scale, was accomplished using the R_{LN} model.

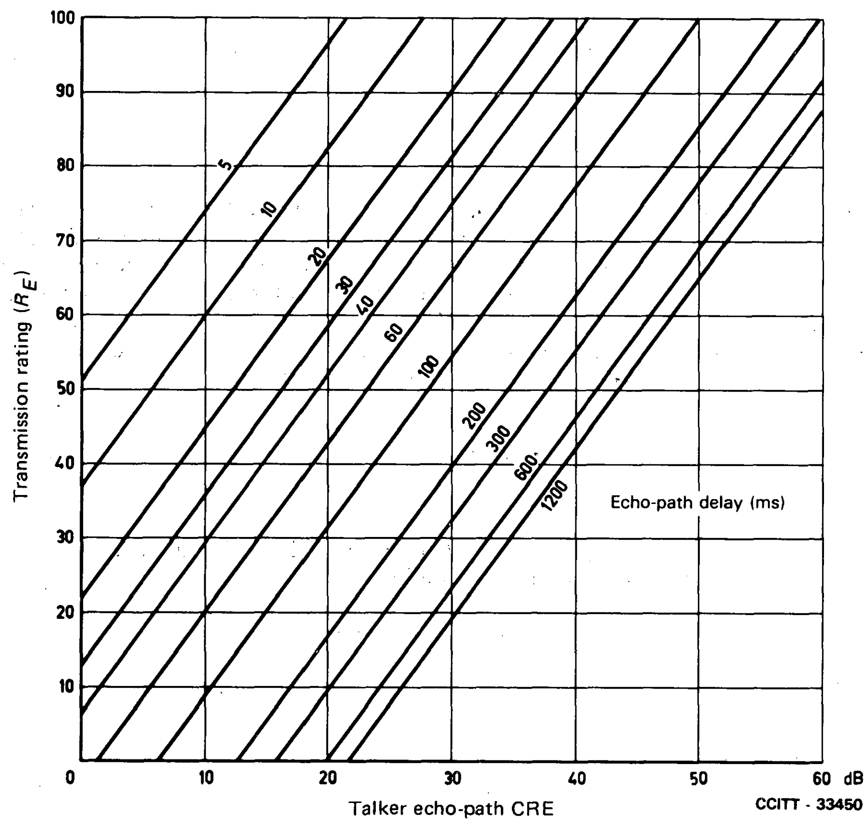


FIGURE 7

Transmission rating for talker echo

The transmission rating model for the combined effects of overall CRE, circuit noise, echo path loss and echo path delay is

$$R_{LNE} = \frac{R_{LN} + R_E}{2} - \sqrt{\left(\frac{R_{LN} - R_E}{2}\right)^2 + 100} \tag{2-24}$$

Figure 8 shows curves generated by means of the above relationship for the transmission rating as a function of talker echo path loss and delay in a connection with an overall CRE of 16 dB and circuit noise of −60 dBmp.

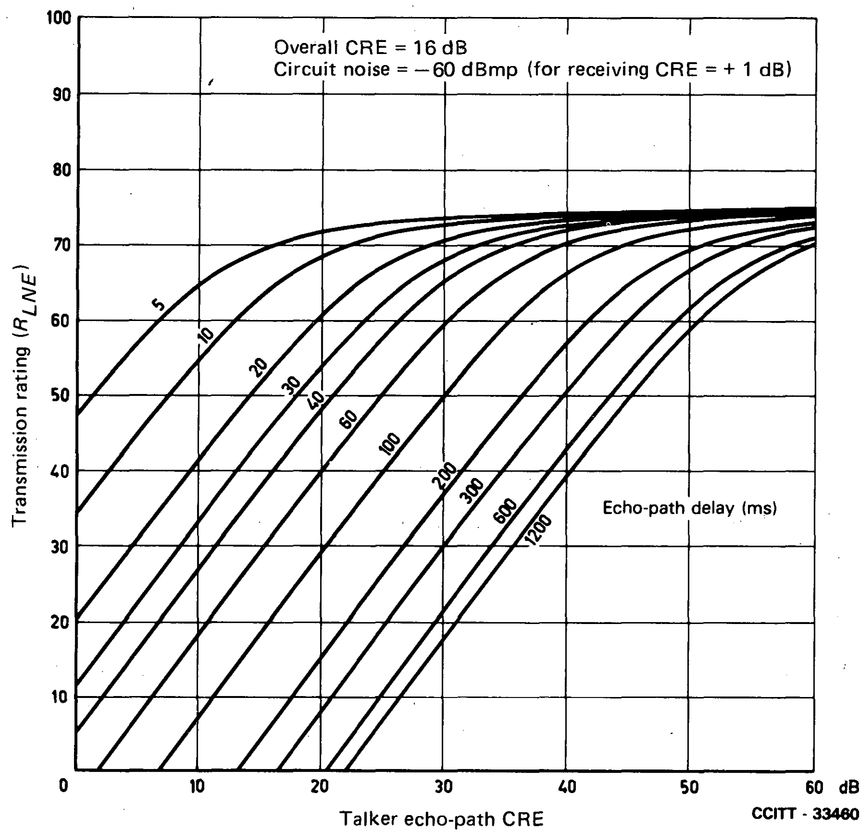


FIGURE 8

Transmission rating for overall CRE, circuit noise and talker echo

2.7 *Sidetone*

The transmission rating model for CRE, total effective noise and talker echo can be modified to include the effects of sidetone. The transmission rating, R_{LN-ST}, for CRE, total effective noise and sidetone is

$$R_{LN-ST} = K_{ST} R_{LN} \tag{2-25}$$

and for talker echo and sidetone is

$$R_{E-ST} = R_E + 2.6(12 - SL) - 1.5(4.5 - SR)^2 + 3.38. \tag{2-26}$$

The sidetone factor, K_{ST}, is calculated from

$$K_{ST} = 1.021 - 0.002(SL - 15)^2 + 0.001(SR - 2)^2(SL - 15). \tag{2-27}$$

SL is the $EARS$ [2] sidetone path loss (in dB), SR is the sidetone response (in dB/octave) below 1 kHz. (The sidetone response above 1 kHz is 1.5 times greater.[3])

Figure 9 shows curves obtained by determining R_{LN-ST} and R_{E-ST}, then substituting these values for R_{LN} and R_E respectively in equation (2-24).

3 Subjective opinion models

Subjective opinion in terms of the proportion of ratings in each of the five categories (E, G, F, P, U) for a condition having a given transmission rating has been found to depend on various factors such as the subject group, the range of conditions presented in a test, the year in which the test was conducted, and whether the test was conducted on conversations in a laboratory environment or on normal telephone calls. The proportion of comments Good plus Excellent (G + E) or Poor plus Unsatisfactory (P + U) can be computed from the following equations:

$$\mathrm{G} + \mathrm{E} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\mathrm{A}} e^{-\frac{t^2}{2}} \, \mathrm{d}t \tag{2-28}$$

$$\mathrm{P} + \mathrm{U} = \frac{1}{\sqrt{2\pi}} \int_{\mathrm{B}}^{-\infty} e^{-\frac{t^2}{2}} \, \mathrm{d}t \tag{2-29}$$

where A and B are given below for data bases of primary interest.

For each data base listed below, the relationship between the subjective judgements and transmission rating is shown in Figure 10.

Data Base[4]	*A*	*B*
1965 Murray Hill SIBYL Test	(R-64.07)/17.57	(R-51.87)/17.57
CCITT Conversation Tests	(R-62)/15	(R-43)/15
Long Toll Interviews	(R-51.5)/15.71	(R-40.98)/15.71

3) *Sidetone Response:*

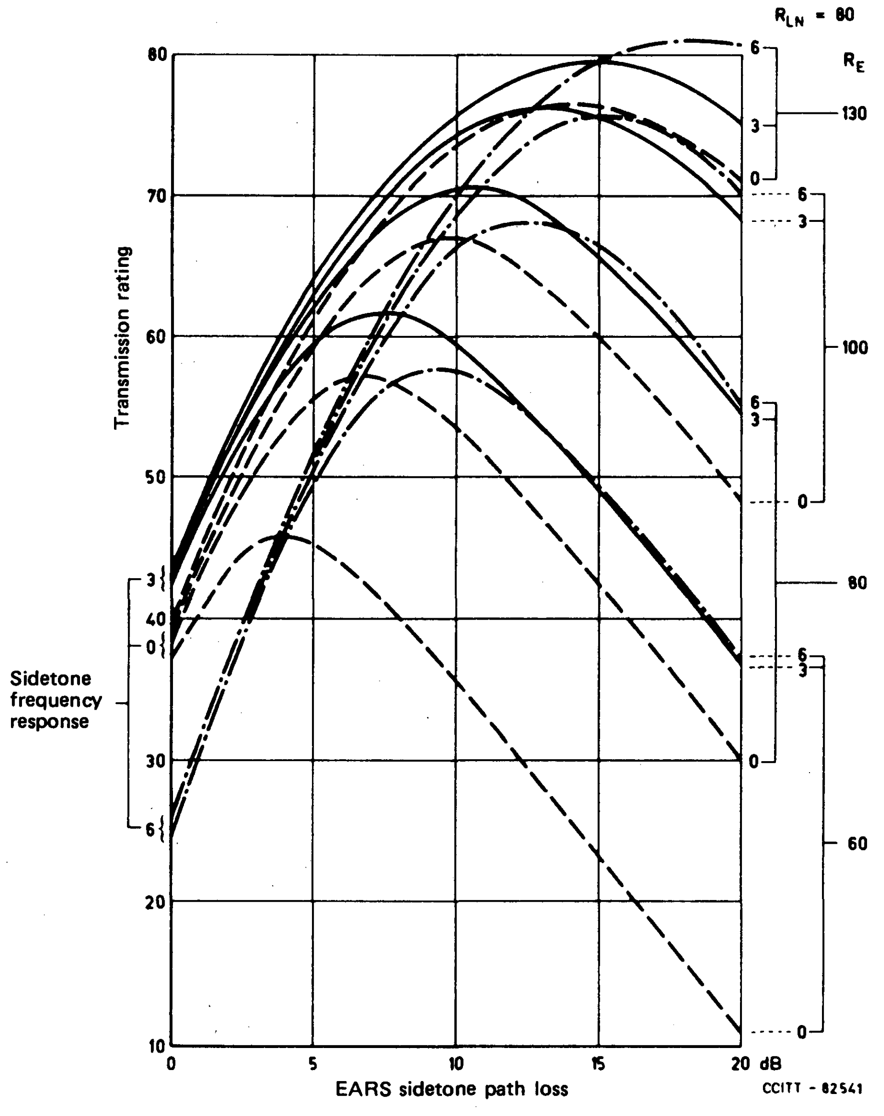
Below 1 kHz	*Above 1 kHz*
0	0
+3.0	+4.5
+6.0	+9.0

4) The three data bases reflect different relationships between the transmission rating scale and opinion ratings as determined in different tests as indicated below:

1965 Murray Hill SIBYL Test – Opinions on actual intra-building business calls.

CCITT Conversation Tests – Composite model of opinion in laboratory conversation tests reported to the CCITT in the 1972-1976 Study Period (see [3]).

Long Toll Interviews – Opinions expressed by North American Telephone customers when interviewed following a call on a long toll connection.



Sidetone path loss		
EARS	*NORE*	*STMR* a)
0	1	3
5	6	8
10	11	13
15	16	18
20	21	23

a) STMR values with seal for SR = 0.
For SR = +3, subtract 1 dB;
for SR = +6, subtract 1.5 dB.

FIGURE 9

Transmission rating for overall CRE, circuit noise, talker echo and sidetone

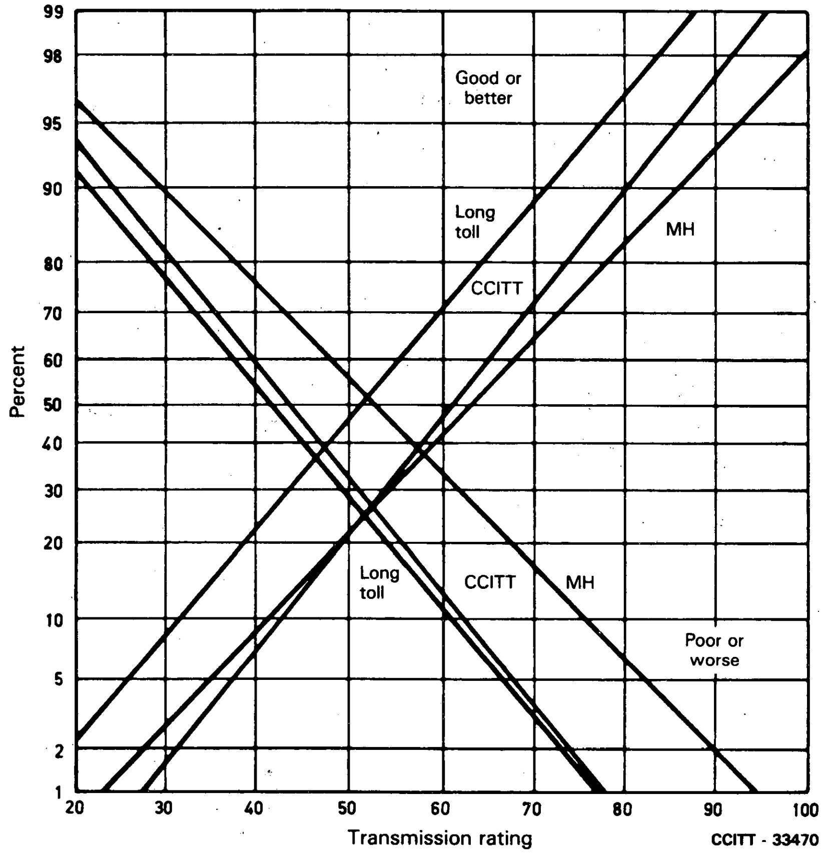


FIGURE 10

Comparison of opinion ratings as a function of transmission rating

ANNEX A

(to Supplement No. 3)

Opinion ratings of transmission impairments

A.1 *Introduction*

The figures in this annex illustrate the relative effect of typical transmission impairments on opinion ratings. They are based on the transmission rating models described above. The opinion ratings assume a five-category rating scale (excellent, good, fair, poor and bad or unsatisfactory) and the results are presented in terms of the percent of ratings which are good or better (good plus excellent) and poor or worse (poor plus bad). Three equations for the conversion from transmission rating to the opinion ratings are described above in the text of the Supplement. The one which is used in this annex is representative of conversational test results reported to the CCITT by several Administrations during the Study Period 1973-1976.

A.2 *Overall CRE and circuit noise*

Opinion ratings for the combined effects of overall CRE (L'_e in dB) and circuit noise (N'_c in dBmp) are shown in Figures A-1 and A-2. The circuit noise is referred to a receiving system with a receiving CRE of +1 dB. In these figures the circuit noise equivalent for room noise N'_{Re} is -62.63 dBmp and the bandwidth/slope factor (k_{BW}) is 1; quantization noise, listener echo, talker echo and sidetone are not included.

A.3 *Quantization noise from PCM processes*

Opinion results for the effect of quantization noise from tandem 7 bit and 8 bit μ-law and A-law PCM processes are shown in Figures A-3 and A-4. These results assume an overall CRE (L'_e) of 16 dB and a circuit noise (N'_c) of -60 dBmp. Room noise, bandwidth/slope and sidetone assumptions are the same as for § A.2. The speech level at the output of a telephone set with a +3 dB sending CRE is assumed to be -10 VU.

A.4 *Bandwidth*

The effect on opinion rating as a function of bandwidth between frequencies having 10 dB of loss relative to 1000 Hz is shown in Figures A-5 and A-6. These results assume an overall CRE (L'_e) pf 16 dB, a circuit noise (N'_c) of -60 dBmp, a circuit noise equivalent for room noise (N'_{Re}) of -62.63 dBmp, and lower (S_l) and upper (S_u) slope factors of 2 and 3 respectively. Listener echo, talker echo and sidetone effects are not included.

A.5 *Listener echo*

The effect of listener echo on opinion ratings is illustrated in Figures A-7 and A-8. In these figures the opinion is plotted (from both the original and alternate models of the supplement) as a function of the weighted listener echo path loss (*WEPL*) in dB and round-trip listener echo path delay (D_L) in milliseconds. The curves were calculated assuming an overall CRE (L'_e) of 16 dB, a circuit noise (N'_c) of -60 dBmp, a circuit noise equivalent for room noise (N'_{Re}) of -62.63 dBmp, and a bandwidth/slope factor of 1. Talker echo and sidetone effects are not included.

A.6 *Talker echo*

Opinion ratings for talker echo are presented in Figures A-9 and A-10 as a function of the CRE of the talker echo path (E) in dB and the round-trip talker echo path delay (D) in milliseconds. Again, the overall CRE (L'_e) was taken as 16 dB, the circuit noise (N'_c) as -60 dBmp, the circuit noise equivalent of room noise (N'_{Re}) as -62.63 dBmp and the bandwidth/slope factor as 1. Listener echo and sidetone effects are not included.

A.7 *Sidetone*

Opinion ratings for sidetone are presented in Figures A-11 and A-12 in terms of the sidetone path loss (*EARS*) in dB and the sidetone response shape in dB/octave. For these curves, impairment levels were selected to provide a constant R_{LN} value typical of toll calls in North America and a range of R_E values which might be encountered on toll calls in North America.

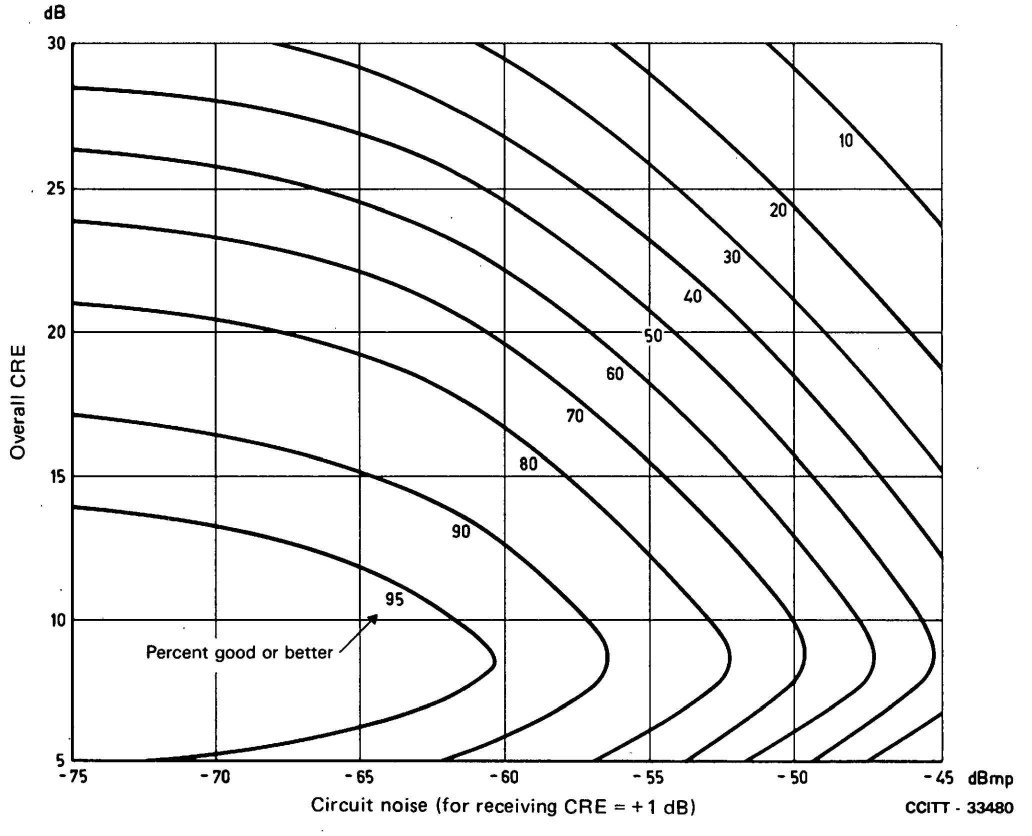


FIGURE A-1

Opinion rating for overall CRE and circuit noise

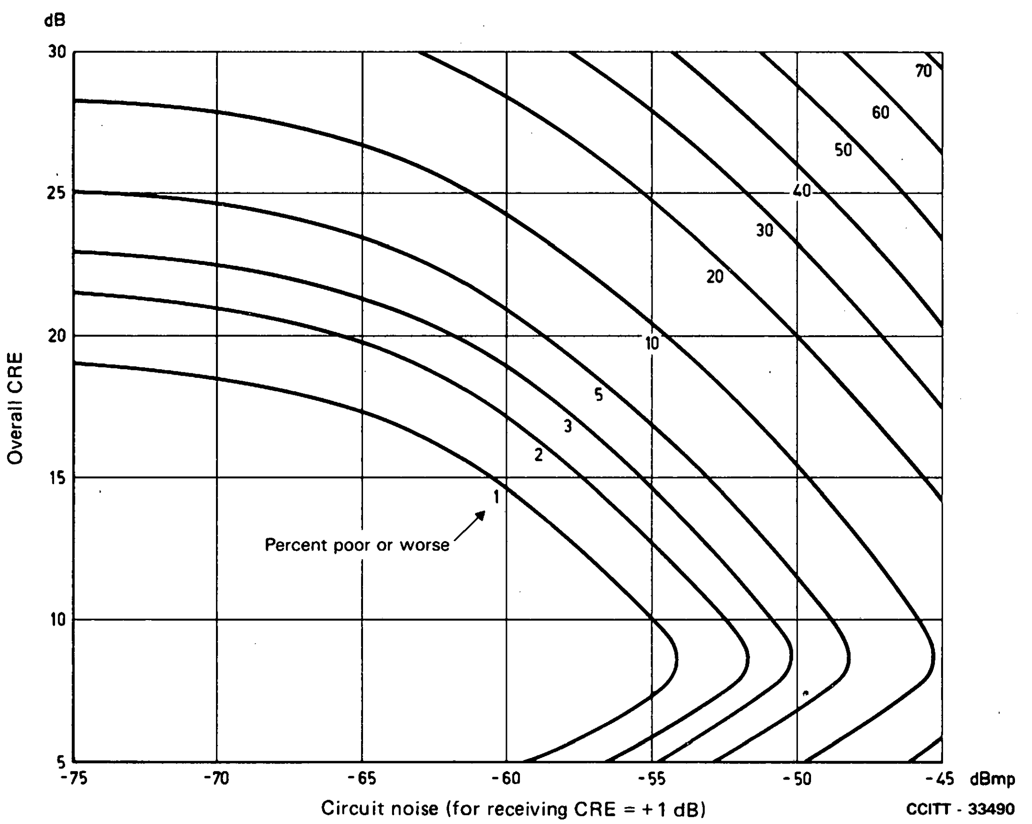


FIGURE A-2

Opinion rating for overall CRE and circuit noise

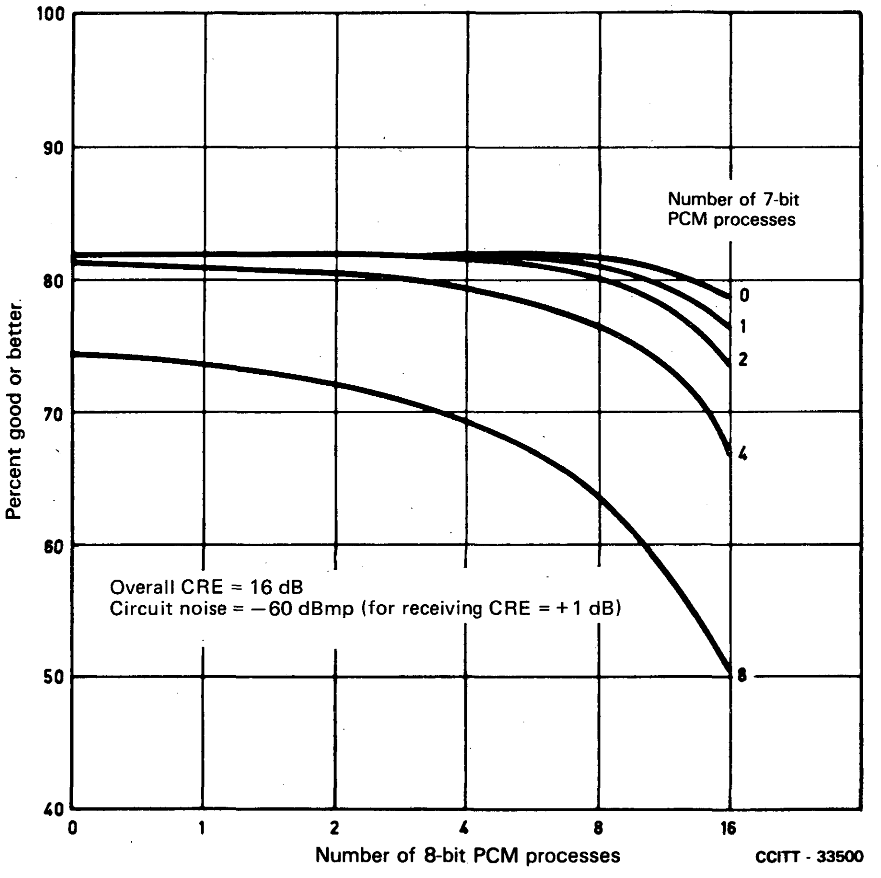


FIGURE A-3

Opinion rating for tandem PCM processes

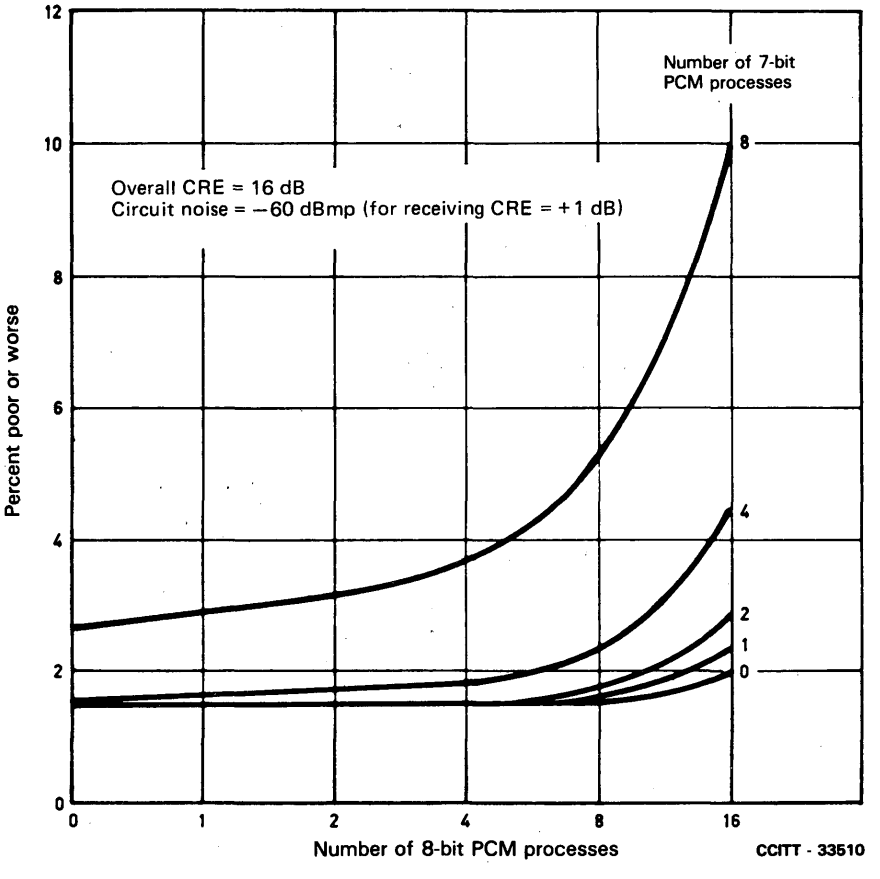


FIGURE A-4

Opinion rating for tandem PCM processes

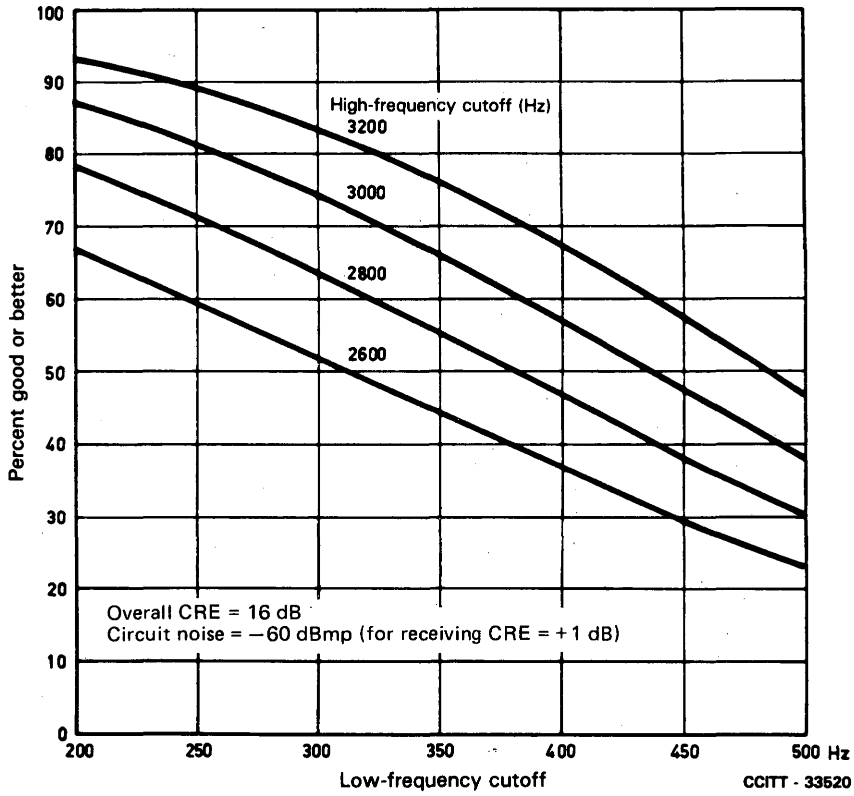


FIGURE A-5

Opinion rating for bandwidth

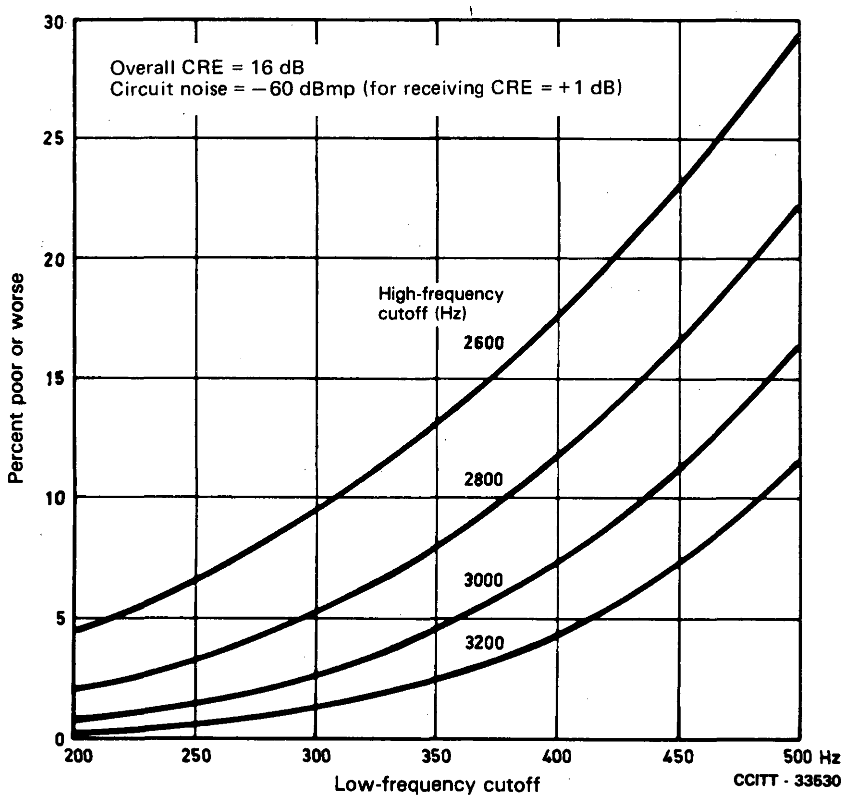


FIGURE A-6

Opinion rating for bandwidth

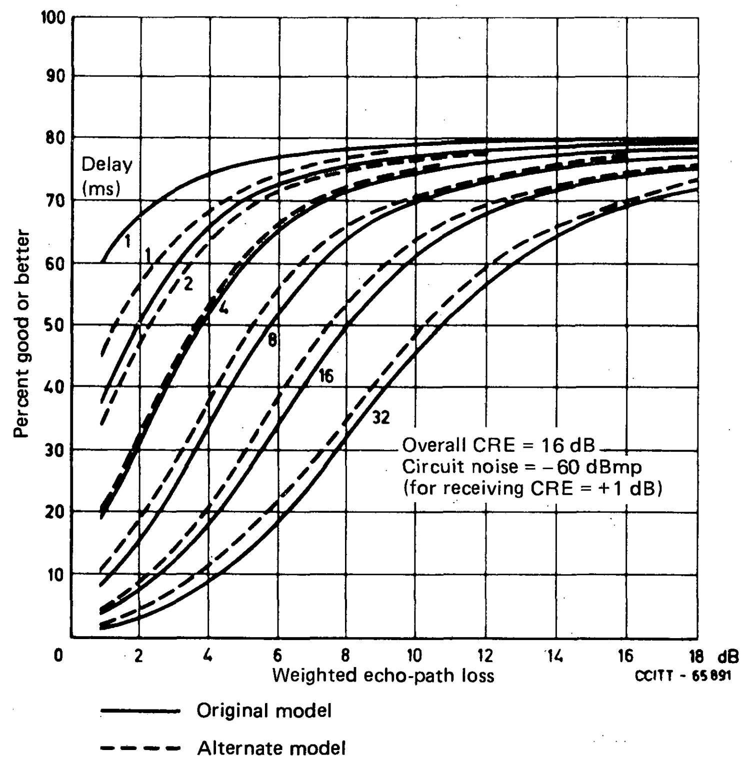


FIGURE A-7

Opinion rating for overall CRE, circuit noise and listener echo

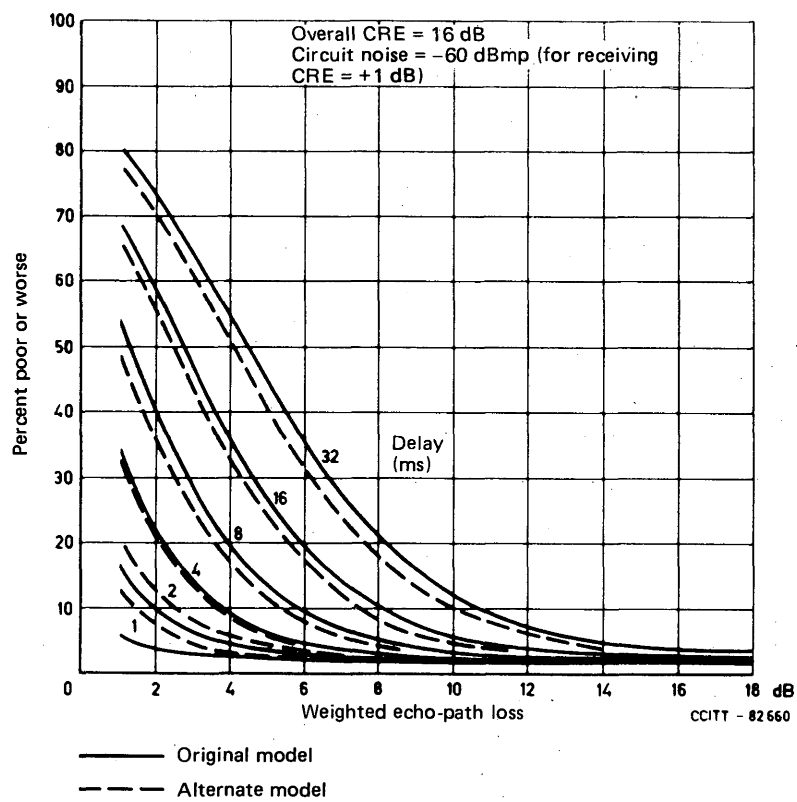


FIGURE A-8

Opinion rating for overall CRE, circuit noise and listener echo

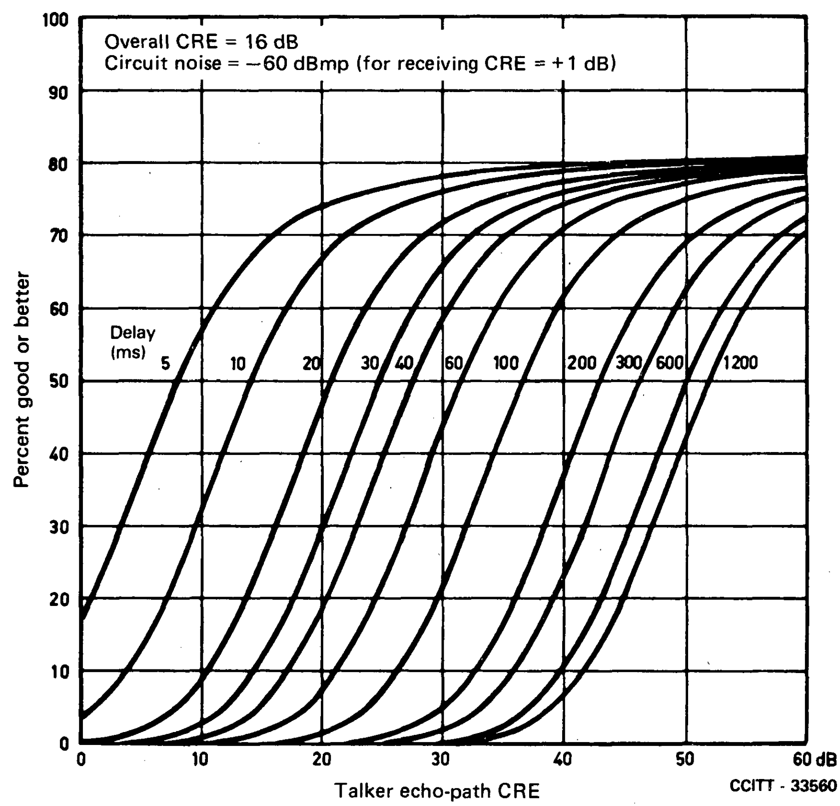


FIGURE A-9

Opinion rating for overall CRE, circuit noise and talker echo

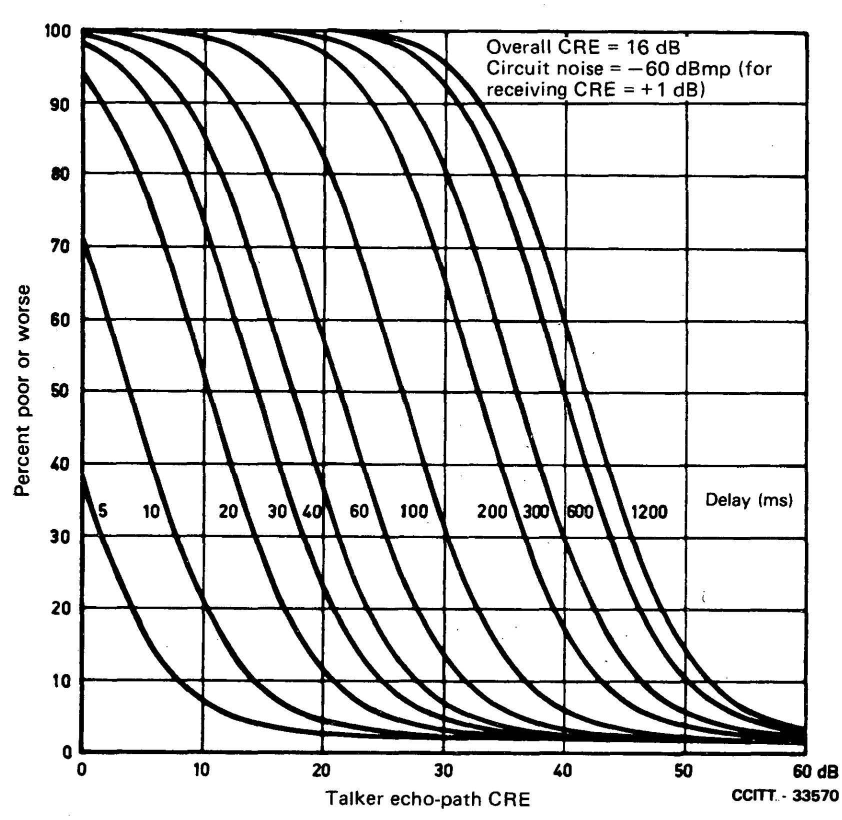
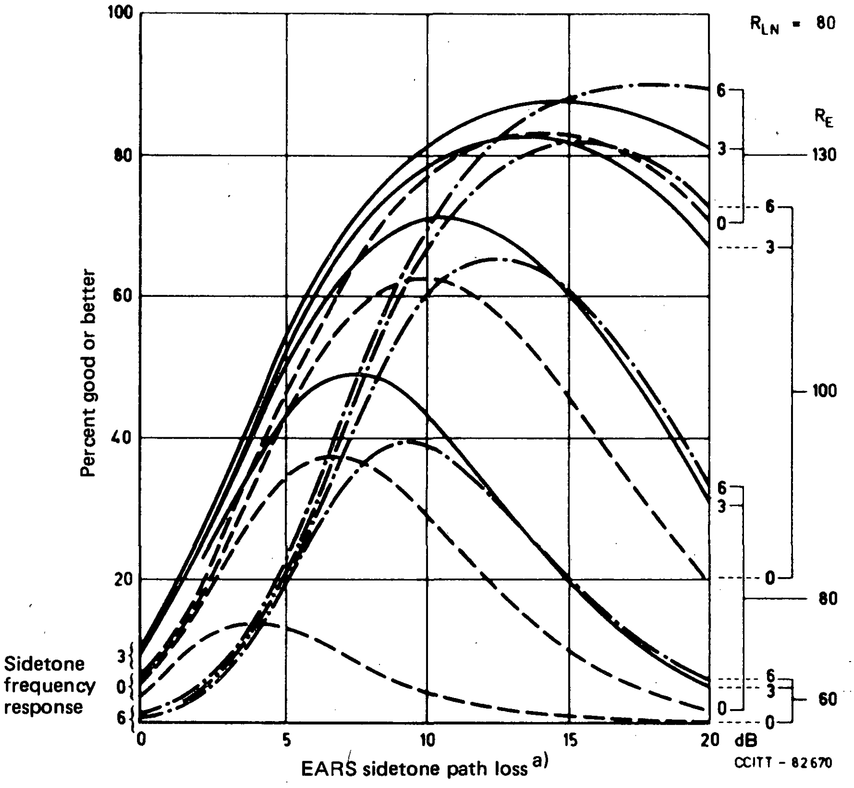


FIGURE A-10

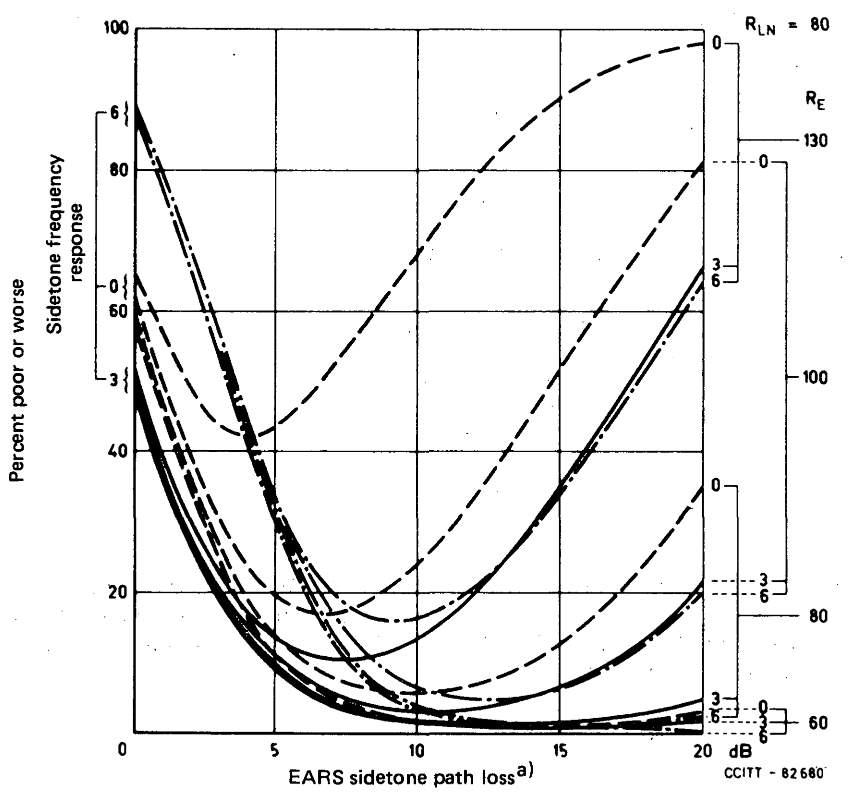
Opinion rating for overall CRE, circuit noise and talker echo



a) See Figure 9.

FIGURE A-11

Opinion rating for overall CRE, circuit noise, talker echo and sidetone



a) See Figure 9.

FIGURE A-12

Opinion rating for overall CRE, circuit noise, talker echo and sidetone

References

[1] CAVANAUGH (J. R.), HATCH (R. W.) and SULLIVAN (J. L.): Models for the subjective effects of loss, noise and talker echo on telephone connections, *Bell System Technical Journal*, Vol. 55, No. 9, pp. 1319-1371, November, 1976.

[2] SULLIVAN (J. L.): A laboratory system for measuring loudness loss of telephone connections, *Bell System Technical Journal*, Vol. 50, No. 8, pp. 2663-2739, October 1971.

[3] CCITT – Question 4/XII, Contribution COM XII-No. R4, Study Period 1981-1984, Geneva, 1982.

[4] CAVANAUGH (J. R.), HATCH (R. W.) and NEIGH (J. L.): A model for the subjective effects of listener echo on telephone connections, *Bell System Technical Journal*, Vol. 59, No. 6, pp. 1009-1060, July-August 1980.

[5] CCITT – Contribution COM XII-No. 13, (Bell-Northern Research), Study Period 1981-1984, Geneva, 1981.

[6] CCITT – Contribution COM XII-No. 14, (Bell-Northern Research), Study Period 1981-1984, Geneva, 1981.

[7] DAUMER (W. R.) and CAVANAUGH (J. R.): A subjective comparison of selected digital codecs for speech, *Bell System Technical Journal*, Vol. 57, No. 9, pp. 3119-3165, November 1978.

[8] CCITT – Contribution COM XII-No. 173, (American Telephone and Telegraph Company), Study Period 1977-1980, Geneva, 1979.

[9] CCITT – Contribution COM XII-No. 94, (American Telephone and Telegraph Company), Study Period 1981-1984, Geneva, 1982.

[10] CCITT – Contribution COM XII-No. 159, (American Telephone and Telegraph Company), Study Period 1977-1980, Geneva, 1979.

[11] CCITT – Contribution COM XII-No. 158, (American Telephone and Telegraph Company), Study Period 1981-1984, Geneva, 1983.

[12] AHERN (W. C.), DUFFY (F. P.) and MAHER (J. A.): Speech signal power in the switched message network, *Bell System Technical Journal*, Vol. 57, No. 7, pp. 2595-2726, September 1978.

[13] CUMMISKEY (P.), JAYANT (N. S.) and FLANAGAN (J. L.): Adaptive quantization in differential PCM coding of speech, *Bell System Technical Journal*, Vol. 52, No. 7, pp. 1105-1118, September 1973.

Supplement No. 4

PREDICTION OF TRANSMISSION QUALITIES FROM OBJECTIVE MEASUREMENTS

(Geneva, 1980; modified in Malaga-Torremolinos, 1984)

(Quoted in Recommendation P.11)

(Contribution from British Telecom)

Summary

British Telecom makes extensive use of a theoretical model for predicting the transmission performance of telephone connections. A brief description is here given of the structure of this model, and of the computer program CATNAP, which embodies a simplified form of the model for routine use, together with facilities for specifying connections in a convenient practical way.

1 Types of model

Question 7/XII [1] recognises two types of "model" for predicting the performance of complete telephone connections in conversation. The first kind, mentioned in [2], involves purely empirical treatment of basic observations, and might lead to a set of tables, graphs or relatively simple formulae, representing performance as a function of certain objective quantities. In a model of this type, where attention is focussed entirely on the correspondence between input (objective quantities) and output (subjective performance), the *form* of the functions employed has no significance in itself. For convenience, simplicity is usually sought, but is obtained at the expense of generality. Interactions between different degradations are often difficult enough to treat in any case, but besides this a purely empirical model must usually be completely revised when a new degradation is brought in; for example, suppose relationships have been established between loss, noise and opinion score for one particular bandwidth: changing that bandwidth to a new constant value will necessitate a redetermination of the functions – not just a constant adjustment of the output. In short, it is unreasonable to expect that a purely empirical model could have more than a limited success in predicting performance.

Models of the second type (mentioned in [3]) are intended to overcome these disadvantages by making the structure of the evaluation process reflect the cause-and-effect relationships which lead from the input (properties of the connection; acoustic environment; characteristics of the participants' hearing, speech sounds and language systems, etc.) to the output (participants' satisfaction or estimate of performance). Such a model is inherently more complicated, and requires more work to develop initially, but can then be extended and applied with much greater ease and confidence. Numerical parameters may and do require revision as more reliable data become available, but the structure, if well chosen, will only rarely require major alterations. As a research tool, such a model is much more powerful in its capability of generating hypotheses to be tested than a collection of useful but arbitrary formulae. As a planning or application tool, it lends itself easily to being embodied in a computer program, to which readily available data (such as losses and line lengths) can be supplied as input.

2 Model and programs: SUBMOD, CATPASS and CATNAP

The model here described is of the more fundamental type. It is intended to predict loudness judgements, listening-effort scores, conversation-opinion scores and vocal levels from objective information supplied. It is embodied in a program called SUBMOD (mnemonic for SUBJECTIVE MODEL) which makes provision for changing the parameters of the model in order to improve agreement between theory and observation. Reference [4] describes an earlier version of the same model.

In its present state of development the model deals fairly successfully with the subjective effects of circuit loss, attenuation-frequency distortion, circuit noise, quantizing noise, room noise, and sidetone paths, for a reasonably wide range of values of these characteristics in any combination. Effects of some other phenomena can also be approximately estimated, but are not yet incorporated in the model. No attempt has yet been made to cater for features such as voice-switching effects, or vocoding and other sophisticated schemes for reducing information rate. Compare the groups of factors listed in Question 7/XII [1].

The program CATPASS [5] – a mnemonic for COMPUTER-AIDED TELEPHONY PERFORMANCE ASSESSMENT – incorporated the same model in a simplified, fixed-parameter implementation, together with facilities for calculating the sensitivity-frequency response of a complete connection formed by concatenating common pieces of apparatus such as telephones, cables, feeding bridges, junctions, and filters. It was similar to the system described in [6] and [7], but the program was differently organized. However, CATPASS could handle symmetrical connections only – that is, those for which transmission, room noise, sidetone and all other relevant features were the same for both participants. It is now superseded by a program called CATNAP (COMPUTER-AIDED TELEPHONE NETWORK ASSESSMENT PROGRAM), which incorporates an extended form of the fixed-parameter model, allowing asymmetry in the connections, as well as containing facilities for assembling performance statistics on sets of connections. See [8].

3 Situation to be represented

Let A and B denote two "average" participants in a telephone conversation over a link terminated in handset telephones, located in rooms with no abnormal reverberation and with specified levels of room noise. "Average" is intended to convey that the participants have representative hearing and speaking characteristics and a normal attitude towards telephone facilities, so that their satisfaction with the telecommunication link may be measured by the mean Conversation Opinion Score (Y_C) and the Percentage Difficulty (%D) that would be obtained from a conversation test, as described in [9]. Y_C can take any value between 4 and 0, the scale being: 4 = EXCELLENT, 3 = GOOD, 2 = FAIR, 1 = POOR, 0 = BAD. %D can of course take any value between 0 for the best connections and 100% for the worst.

For a given connection, the quantities of chief interest are Y_C, %D and the speech level, for each participant. However, other useful auxiliary quantities are computed in the course of the evaluation, such as the loudness ratings of the various paths (calculated according to Recommendation P.79), and Y_{LE}, the mean

Listening Effort Score that would result from a listening opinion test conducted as outlined in [9]. In a listening test of this type, lists of sentences at a standard input speech level are transmitted over the connection and the listener votes, at a number of different listening levels, on the "effort required to understand the meanings of sentences" according to the following scale:

A complete relaxation possible; no effort required

B attention necessary; no appreciable effort required

C moderate effort required

D considerable effort required

E no meaning understood with any feasible effort.

The votes are scored A = 4, B = 3, C = 2, D = 1, E = 0, and the mean taken over all listeners is called the Listening Effort Score, Y_{LE}, for each particular listening level and each circuit condition.

More detailed information about both conversation tests and listening tests may be found in [10], and also in [9].

4 Outline of the model

The model requires the following inputs:

1) overall sensitivity-frequency characteristic of each transmission path (talker's mouth to listener's ear via the connection) and sidetone path (each talker's mouth to his own ear). These sensitivities may be either measured by the method described in Recommendation P.64 or calculated as explained in Reference [5];

2) noise spectrum and level at each listener's ear, composed of noise arising in the circuit, room noise reaching the listening ear direct, and room noise reaching the listening ear via the sidetone path. In the absence of specific measurements, standard noise spectra and levels are taken; e.g. room noise with Hoth spectrum at 50 dBA, circuit noise with bandlimited spectrum at a specified psophometrically weighted level;

3) average speech spectrum and average threshold of hearing, as given for example in [11].

From these data the loudness ratings are calculated. With speech level fixed, Y_{LE} and a provisional value of Y_C are evaluated for each participant. The relationships between Y_C and speech level at each end are then used to refine the values of both, so that the final estimates represent performance at realistic conversational speech levels.

5 Calculation of loudness and loudness ratings

The model starts by setting the speech level emitted from each talker to a standard value and calculating the resultant spectrum and level of both speech and noise at each listener's ear. The loudness of received speech is calculated as a function of signal level, noise level and threshold of hearing, integrated over the frequency range extending normally from 179 to 4472 Hz (14 bands, the lowest centred at 200 Hz and the highest at 4000 Hz). The loudness of the sidetone speech is calculated similarly, but with an allowance for the additional masking effect of speech reaching the ear naturally (via the air path and the bone-conduction path). By comparison with the loudness of speech transmitted over an IRS (Intermediate Reference System), the loudness ratings of the various paths are evaluated: SLR, RLR and STMR for each end, and OLR in each direction.

The method is described in [12], but is not given in detail here. The loudness part of the model is important in its own right [for example in the study of Question 19/XII [15]], but not closely connected with the rest of the model. The program outputs loudness ratings calculated according to Recommendation P.79, but also calculates a set of loudness ratings according to the earlier method [16] which are used for subsequent calculations.

6 Calculation of listening effort score

This part of the model is intended to reproduce the result that would be obtained from a Listening Opinion Test.

It has been found possible to estimate Y_{LE} by a process similar to those already well known in calculating loudness and articulation score. An intermediate quantity, Listening Opinion Index (LOI), is first calculated as follows. Each elementary band in the frequency range contributes to LOI an amount proportional to the product

$B'_f P(Z_f)$, where B'_f is a frequency-weighting factor expressing the relative importance of that elementary band for effortless comprehension, and P is a growth function applied to the sensation level Z (which has already been evaluated for the loudness calculation). The actual values of the frequency-weightings differ somewhat from those used in loudness and articulation calculations; the growth function is limited to the range 0 to 1 as in articulation, but the form used is:

$$P(Z) = 10^{\frac{Z + 3.8}{10}} \qquad \text{if } Z < -11,$$

$$P(Z) = 1 - 10^{\frac{-0.3\,(Z + 14)}{10}} \qquad \text{otherwise.}$$

LOI is proportional to $\int B'_f P(Z_f)\,\mathrm{d}f$, but in practice the integral is replaced by a summation over a number of bands (normally 14), within each of which Z_f and B'_f are reasonably constant, just as in the loudness evaluation. The formula actually used is:

$$\mathrm{LOI} = AD \sum_i B'_i\, P(Z_i)$$

where

B'_i is the frequency weighting for the ith band, (shown diagrammatically in Figure 1),

Z_i is the mean Z in the i th band,

P is the appropriate growth function (illustrated in Figure 2),

A is a multiplier depending on the received speech level, with the value 1 for a small range of levels around the optimum but decreasing rapidly outside this range (see Figure 3), and

D is a multiplier depending on the received noise level (ICN-RLR) with a value decreasing slowly from 1 at negligible noise levels towards 0 at very high levels (see Figure 4).

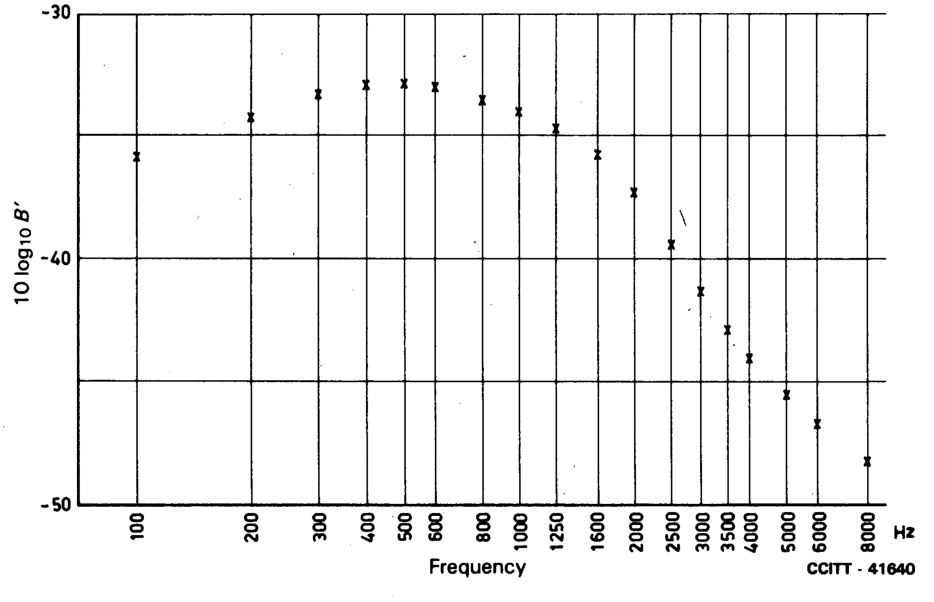


FIGURE 1

Frequency-weighting factor B' for Listening Opinion Index

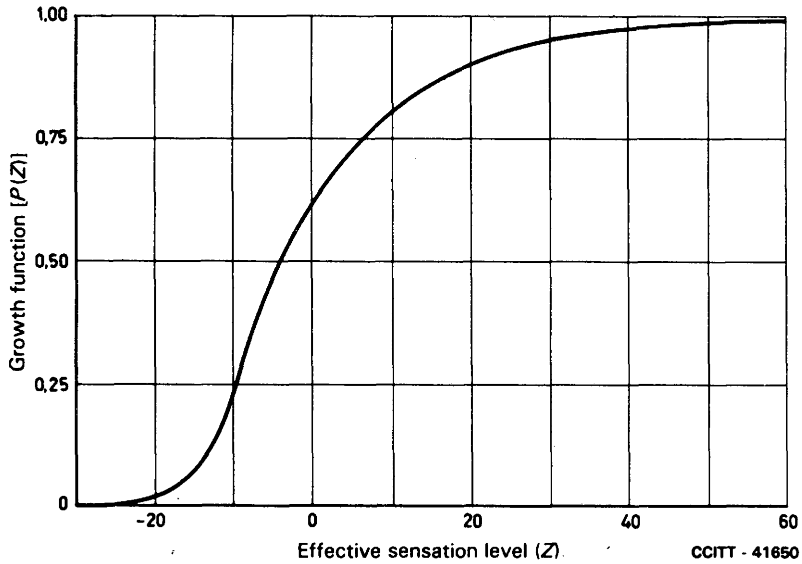


FIGURE 2

Growth function *P*(*Z*)

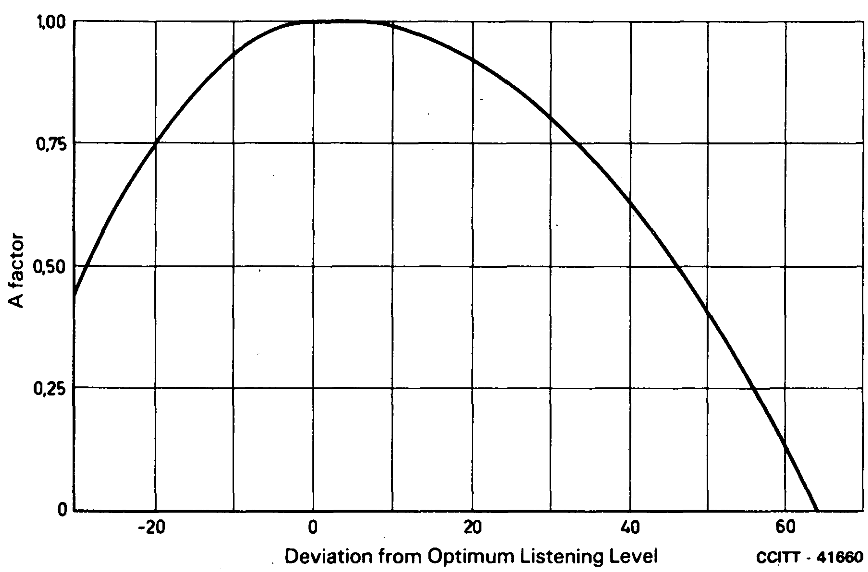


FIGURE 3

Effect of listening level on Listening Opinion Index

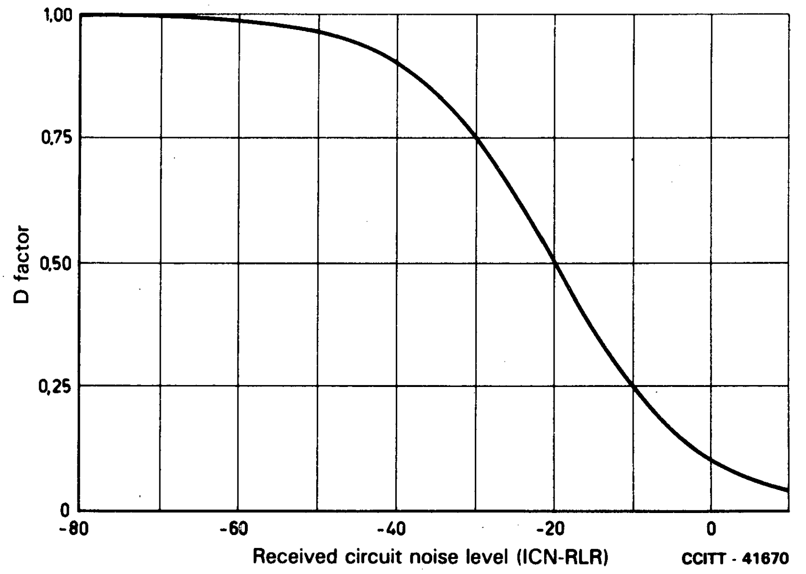


FIGURE 4

Effect of received noise level on Listening Opinion Index

Thus it is only for wide-band, noise-free, distortion-free speech at optimum listening level that LOI attains its maximum value of unity.

The Listening Opinion Index is related to Y_{LE} in a manner which depends on the standard of transmission to which listeners have been accustomed in their recent experience. It is found that the subjects' standard of judgement is influenced mostly by the best circuit condition experienced in the current experiment, or, in real calls, by the quality of the best connections normally experienced. For example, a circuit condition which earns a score of almost 4 in an experiment where it is the best condition, would earn a score of perhaps only 3 if a practically perfect condition were included in the same experiment, and about 3.5 if the best condition in the same experiment were equivalent in performance to the best connection that can normally occur in the British Telecom system. A parameter $\mathrm{LOI_{LIM}}$, introduced to cater for this effect, specifies the value of LOI that corresponds to maximum Y_{LE}; it is generally set equal to 0.885 when connections are being judged against a background of experience with the British Telecom network. The relationship in general terms is

$$\ell n\left(\frac{Y_{LE}}{4 - Y_{LE}}\right) = 1.465 \left[\ell n\left(\frac{\mathrm{LOI}}{\mathrm{LOI_{LIM}} - \mathrm{LOI}}\right) - 0.75\right]$$

as shown in Figure 5. This brings us to the point where Y_{LE} has been evaluated for each participant as a function of listening level – in particular, at the listening level established for each participant when the other speaks at Reference Vocal Level (RVL), defined in [14].

7 Calculation of Conversation Opinion Score

In order to convert a value of Y_{LE} at the appropriate listening level to the corresponding value of Conversation Opinion Score (Y_C), it is necessary to take account of deviations of mean vocal level from RVL.

The symbol V_L is used to denote the electrical speech level in dBV at the output of a sending end when the acoustic level at the input (mouth reference point) is RVL. During conversation, a different level (V_C) will generally prevail at the same point, because participants tend to raise their voices if incoming speech is faint or poor in quality and to lower them if incoming speech is loud. In other words, V_C at end A depends on Y_{LE} at end A, which depends on V_C at end B, which depends on Y_{LE} at end B, which depends in turn on V_C at end A. Thus there is a circular dependence or feedback effect.

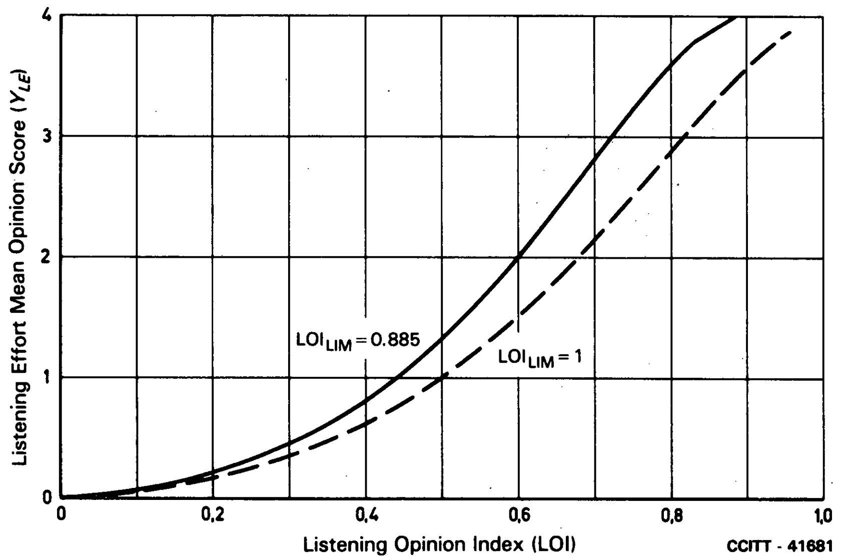


FIGURE 5

Listening Opinion Score as a function of Listening Opinion Index

The sidetone paths introduce complications when STMR < 13 dB (besides contributing noise from the environment to the receiving channel as already explained). Other things being equal, each talker's vocal level goes down by almost 1 dB for every 3 dB decrease in STMR below 13 dB, and this of course further modifies the opinion scores and speech levels at both ends by virtue of the feedback effect.

In addition to this, very high sidetone levels are experienced as unpleasant *per se*, particularly when the connection is poor for other reasons.

This complex interrelationship is found to be reasonably well represented by the following equations.

Y'_C is an intermediate quantity explained below.

$$\ell n \left(\frac{Y'_C}{4 - Y'_C} \right) = 0.7 \left[\ell n \left(\frac{Y_{LE}}{4 - Y_{LE}} \right) + 0.5 - \frac{\mathrm{K}\,(13 - \mathrm{STMR})}{20} \left(\frac{4 - Y_{LE}}{Y_{LE}} \right)^2 \right] \tag{7-1}$$

$$V_C - V_L = 4.0 - 2.1\ Y'_C - 0.3\ \mathrm{K}\,(13 - \mathrm{STMR}) \tag{7-2}$$

$$\ell n \left(\frac{Y_C}{4 - Y_C} \right) = 0.8451\ \ell n \left(\frac{Y'_C}{4 - Y'_C} \right) - 0.2727 \tag{7-3}$$

where

K = 1 if STMR < 13,

K = 0 otherwise.

By substituting in equation (7-1) the value of Y_{LE} already found for end A – which would apply for $V_C = V_L$ at end B – one obtains a first approximation to Y'_C, then from equation (7-2) an approximation to V_C at end A. The earlier calculations are repeated with this speech level to find a new value of Y_{LE} at end B, hence an approximation to Y'_C and V_C at end B. This process is repeated cyclically until each Y'_C converges to a settled value, and then equations (7-1) and (7-2) are simultaneously satisfied.

Figure 6 shows the form of the resultant relationship between Y_{LE} and Y'_C, for two different values of STMR, with V_C at its proper value. The transformation [equation (7-3)], illustrated in Figure 7, is then applied to the intermediate score Y'_C, to give the estimated Conversation Opinion Score Y'_C, which is shown as a function of Y_{LE} in Figure 8.

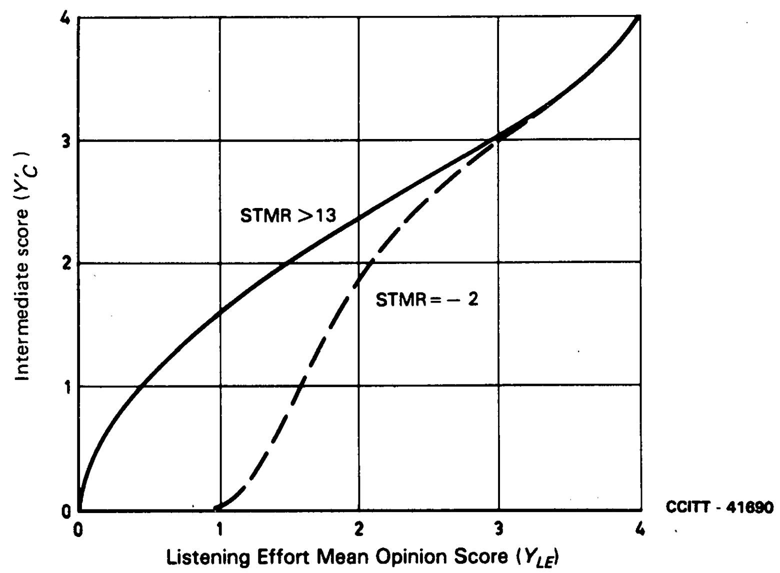


FIGURE 6

Intermediate score as a function of Listening Opinion Score

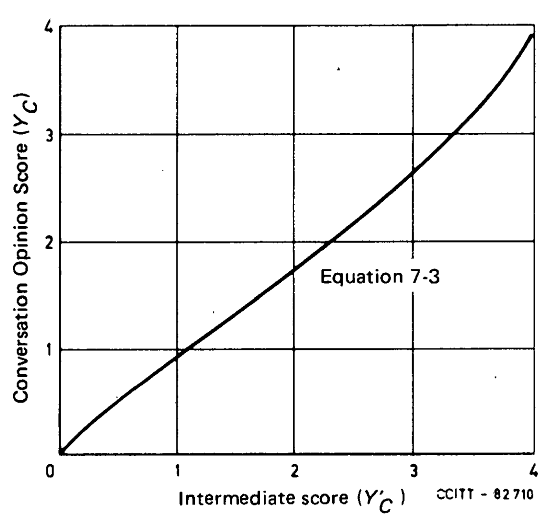


FIGURE 7

Conversation opinion score as a function of intermediate score

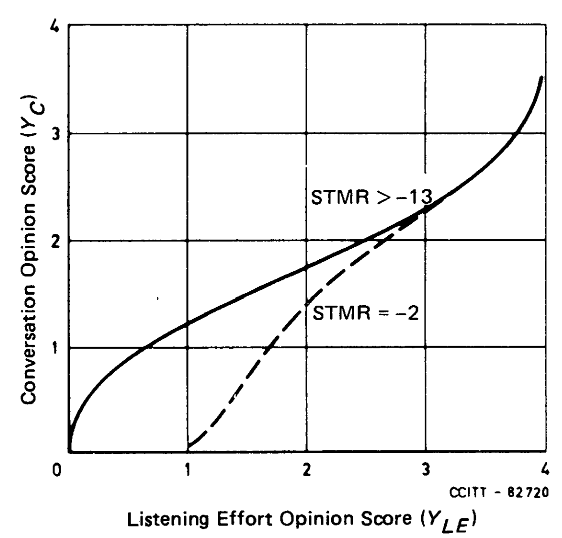


FIGURE 8

Conversation Opinion Score as a function of Listening Opinion Score

8 Evaluation of other subjective measures of performance

Relationships have been developed for various dichotomies of the opinion scale – such as proportion of votes greater than 2 (i.e. votes "Excellent" or "Good") – and for the percentage of positive replies to the "Difficulty" question (Reference [9]).

For example, percentage "Difficulty" is represented by the equation

$$\ln\left(\frac{D}{1 - D}\right) = -2{,}3 \ln\left(\frac{Y_C}{4 - Y_C}\right)$$

where

$D \times 100 = \%D$.

However, these relationships are satisfactory only for certain kinds of degradation and are still under review.

9 Correspondence between calculated and observed values

For symmetrical connections, provided very high sidetone levels and very high room noise levels are excluded, the model reproduces fairly well the results of laboratory conversation tests carried out in the U.K. In the most recent laboratory tests there is a tendency for speech levels and hence opinion scores to be somewhat lower than those observed earlier, but the relativities between circuit conditions are not much disturbed by this. It is believed, but not yet fully established, that approximately the same relativities hold good for other populations of subjects – in particular, for the population of ordinary telephone users accustomed to the British Telecom system – even though different absolute values of scores may be obtained from other populations of subjects or by using different experimental procedures.

Comparatively few results are available from experiments on asymmetrical connections, but such evidence as there is indicates that the model predicts too much divergence between the two ends of the connection – especially in respect of V_C, less so in respect of Y_C. It is proposed to introduce a feedback feature to reduce the divergence between the two V_C values, but care will be needed not to reduce the Y_C divergence too far as a result of this. HRC 4 in Annex A gives an example of CATNAP calculations for a set of connections with asymmetrical losses: compare these predictions with Reference [18] there quoted.

10 Incorporating miscellaneous degradations

10.1 *PCM quantizing distortion*

Reference [15] describes a method for handling the effects of quantizing distortion in PCM systems. It is there established that a quantity Q, effective speech-to-quantization-noise ratio in dB, can be evaluated for any specified type of PCM system as a function of input speech level. It has been found that the subjective effect of a given value of Q can be approximated by that of a level of continuous circuit noise G dB below the speech level, where

$$G = 1.07 + 0.285\ Q + 0.0602\ Q^2.$$

Thus for a connection involving PCM links, one must include an evaluation of equivalent noise level in the iterative process that determines V_C: each successive approximation to V_C leads to a new value for Q, hence to a new value for G, and hence to a new contribution to the circuit noise to be taken into account in calculating the new value of Y_{LE}. In practice these modifications have negligible effect unless the speech level at the input to the PCM system falls below about -25 dBV, or the circuit noise at the same point is very high, or the speech input level is so high (say > -5 dBV) that appreciable peak limiting occurs.

10.2 *Syllabic Companding*

The case of a 2:1 syllabic compandor can be simply handled by finding a subjectively equivalent continuous noise level.

Let S be the speech level at the input to the compressor, and N be the noise level (psophometrically weighted) arising between the compressor and expander, both in dB relative to unaffected level. The resultant levels at the output of the expander will then be as given in Table 1.

TABLE 1

	Speech	Noise while speech present	Noise while speech absent
Level at compressor input	S	—	—
Gain of compressor (dB)	–S/2	—	—
Level at compressor output and expander input	S/2	N	N
Gain of expander (dB)	S/2	S/2	N
Level at expander output	S	N + S/2	2N
Level at same point in absence of compandor	S	N	N
Improvement	—	–S/2	–N

Note that S and N are both normally negative, so that the improvements are positive. Any noise present at the compressor input will be present at the same level at the expander output, and will combine by power addition with the other noise at the same point.

Subjectively equivalent performance is obtained by omitting the compandor and substituting a continuous noise level satisfying the condition:

Total improvement = 1/3 (improvement in presence of speech) +
+ 2/3 (improvement in absence of speech)
= – S/6 – 2N/3.

Hence

equivalent noise level = N – improvement
= N + S/6 + 2N/3 = S/6 + 5N/3.

This noise level is recalculated from V_C on each iteration and used to calculate the next value of Y_{LE}.

10.3 *Delay and echo*

The audibility and objectionability of echo can be expressed as a reasonably simple function of the delay and loudness rating of the echo path, but the wider effects of echo and main-path delay in disrupting conversation can at present only be treated by *ad hoc* estimation from the known performance of circuit conditions in neighbouring parts of the range.

10.4 *Crosstalk*

The loudness part of the model may be used to estimate the audibility of crosstalk, at various attenuations, and hence to find the attenuation required to reduce it to an inaudible level or to an acceptable level.

11 Practical use of the model

At the academic or research level, the chief use of a model of this kind is in promoting an understanding of the fundamentals of telecommunication between human beings, and in finding potential improvements in the techniques of telecommunication systems.

At the practical level, the chief advantage of having the model available is that it encodes the knowledge of the performance of telephone connections in a very economical manner, obviating the need for large and complex tabulations or graphs. For connections containing only the "natural" degradations, the program CATNAP greatly facilitates routine use of the model. The user of this program need not know anything about the theory beyond the meaning of the terms and symbols used, and need not normally make any special measurements. Connections are specified in terms of standard items and quantities, such as noise levels, telephones of particular types, lengths of cable with stated resistance and capacitance per kilometre, and attenuators with stated loss. Starting from these data, the program performs all the necessary calculations and prints out loudness ratings, speech levels, and opinion scores (Y_{LE} and Y_C). More detail can be printed on request.

It would of course be possible to construct a large table of results covering a wide range of connections, but the table would have to be either too large to be practical or else limited by making arbitrary fixed choices for many of the variables. In either case the advantage of having the model – that it holds the information in an economically coded form and releases only the required part on demand – would be lost.

CATNAP may also be used inversely. Suppose it is desired to find what value of some variable in a connection (the independent variable) will yield a given value of one of the dependent variables. By performing runs at different values of the independent variable one identifies a region within which the required value lies; one can then repeat the calculation at ever smaller intervals until the required value is located with sufficient accuracy. For example, where all features except the local line remain fixed, one can find the line length (for the type of cable in question) that will yield values of OLR below some specified maximum, or values of Y_C above some specified minimum. More than one independent variable could of course be adjusted, but correspondingly more work would then be needed in order to find the combinations that satisfied the criterion.

The usefulness of these facilities is evident.

ANNEX A

(to Supplement No. 4)

Calculated Transmission Performance of Telephone Networks

A.1 *Introduction*

This annex is intended to give examples of results from the subjective model which is incorporated in the BT CATNAP (Computer-Aided Telephone Network Assessment Program) program. CATNAP comprises this model and a transmission calculation section which enables elements of a connection to be entered as readily identifiable items, e.g. lengths of cable, feed bridges etc. These results are examples of calculations for various "hypothetical reference connections" (HRCs) which might arise in the network or would be of use to planners.

The loudness ratings quoted are calculated according to Recommendation P.79, using the frequency bands from 200 Hz to 4 kHz. The opinion scores, Y_{LE} and Y_C, are on a scale of 0 to 4, representing the listening effort and conversation opinion scales (see [9]). The values of line current shown with the results are determined by the program which decides from the characteristics of the local telephone system which of a number of standard line currents is appropriate, and hence which values of the telephone instrument characteristics should be used. The program also gives speech levels for controlled talking conditions (V_L) and under conversational conditions (V_C). These and the loudness ratings are referred to the interfaces (NI and FI) shown in the figures below.

These results are for the model as it stands at present (1983 version). Research is continuing to improve the correlation of calculated and experimental results, so the model is liable to modification.

A.2 *HRC 1 – Own exchange call* (see Figure A-1)

This is a symmetrical connection, with average length customers' lines. The sidetone suppression is fairly good, and room noise and circuit noise levels are low. The conversation opinion score is good, but the small overall loss means that the connection is louder than preferred. A slightly quieter connection would give a better opinion score.

A.3 *HRC 2 – Limiting national call* (see Figure A-2)

These two HRCs are both symmetrical and comprise BT limiting local lines of 1000Ω/10 dB, 4.5 dB local junctions and two 4-wire junctions each with 3.5 dB loss, which are the limits set by the BT transmission plan (given in [17]).

HRC 2 (a) uses 0.5 mm copper local lines, which provide much better sidetone matching than the 0.9 mm copper lines of HRC 2 (b). The change in sidetone level (> 10 dB) causes a drop in the conversation opinion score from 1.9 to 0.8 (from fair to poor).

A.4 *HRC 3 – Long distance call with a PCM junction* (see Figure A-3)

The overall loss of this connection (OLR = 13.4 dB) is much less than for HRC 2. The local lines are average length of 0.5 mm copper which give reasonably good sidetone matching, and there is now only one local junction. This is a 4-wire 3 dB PCM junction. This is entered as a single item, characterised by the terminating and balance impedances of the 2/4-wire terminating sets, the matched loss in each direction and the phase delay round the loop. Quantizing noise is negligible for the input speech levels calculated by CATNAP for this connection.

The connection is symmetrical in transmission loss but a small difference in the sidetone level has given slightly different conversation opinion scores at the two ends.

A.5 *HRC 4 – Asymmetry of transmission loss* (see Figure A-4)

A number of calculations have been done for this HRC to show the effect of varying the degree of asymmetry. The curves shown are not fitted curves, but simply join the marked points on the graph. They show the effect on the conversation opinion score and conversational speech voltage of varying the transmission loss in one direction only (from near end to far end). The loss from far to near is kept constant, so the opinion of the near end customer is much less affected. It is suspected that the speech voltage curves are too divergent and further research is needed in this area, but the opinion curves show similar trends to the results produced by Boeryd [18].

The sidetone level was virtually unaffected by the change in transmission loss.

A.6 *HRC 5 – Effect of room noise* (see Figure A-5)

The calculations done for this HRC demonstrate the effect of changing the level of room noise for a customer with a loud sidetone path (near end) and one with a quiet sidetone path (far end). As for HRC 4, the computed points are simply joined to form the line.

A.7 *HRC 6 – Effect of circuit noise and bandlimiting* (see Figure A-6)

This is a connection using 4-wire reference telephones, enabling sidetone to be controlled. The STMR is kept at 20 dB, at which level most customers would not detect it.

Such a connection can be used to investigate the effects of particular transmission impairments varied independently. Here it has been used to demonstrate the effect on the listening effort and conversation opinion scores of the level of injected circuit noise and band limiting (lowpass) over a range of losses likely to occur in telephone networks.

As for the previous curves the computed points are simply joined to form a line.

A.8 *HRC 7 – Multiple calculations with random selection of items* (see Figure A-7)

CATNAP is intended to help assess telephone network proposals rather than single connections. The program can perform multiple calculations on a group of connections or on a single connection with random selection of elements from a database.

Here random selection is made of the customers' lines out of a database derived from a survey of 1800 existing lines. This enables the performance of a particular element to be tested for a range of conditions which would arise in the actual network. Since the survey reflects the distribution of lengths and gauges in the actual network, this method of assessment gives a more accurate picture of the performance in the existing network.

For this example only a few calculations have been done to demonstrate the facility and so the results have been printed. This is not practical for large numbers of calculations, when the results are stored and can be processed as desired, e.g. by plotting the distribution or by statistical analysis.

The line number and radial distance have been given for both ends of each calculation.

A.9 *HRC 8 – Example of the use of CATNAP to meet a design criterion* (see Figure A-8)

This is intended to give an example of the use of CATNAP in the design of individual network components to meet design targets.

With the introduction of electronic telephones the designer has a freer choice of values for the telephone instrument characteristics, e.g. the value of the line impedance which must be connected to the telephone instrument to give full sidetone suppression (Z_{so}).

An iterative procedure can lead to preferred values for Z_{so}. As examples, calculations have been done for a standard BT 706 and a 706 with some trial values for Z_{so} on BT limiting lengths of local copper cable of standard gauges, and an average length of 0.5 mm cable. For one of the trial sets of values which looks possible from these results and for a standard 706 instrument, a set of 40 calculations was done with random selection of local lines from the database of 1800 used for HRC 7. These results are given in terms of the mean and standard deviation of the distribution of STMRs. From this it can be seen that the trial values do give a better performance on average, although the performance is worse on 0.63 mm and 0.9 mm limiting lines, since these are less common in the local network than 0.5 mm.

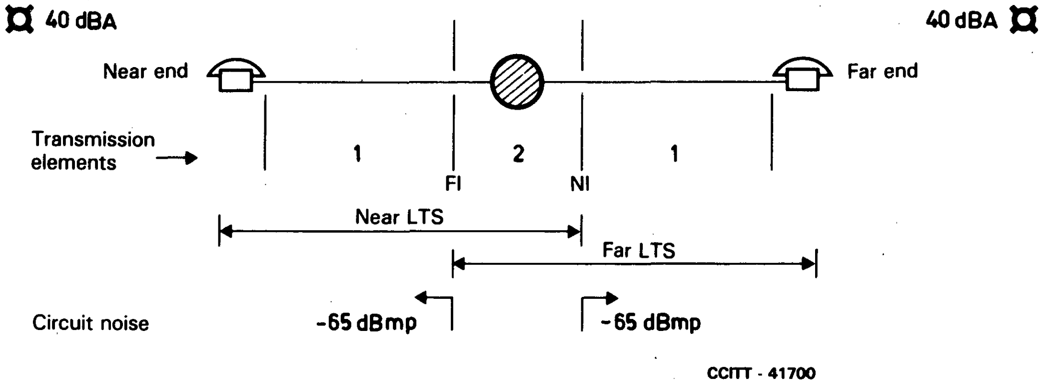
As a design tool, the program could be used further to verify the improvement in performance, to check the effects of tolerances and to consider possible improvements to these values.

A.10 *HRC 9 – Effect of varying line length* (see Figure A-9)

This HRC is identical to HRC 2 except for the gauge of cable. In this case 0.63 mm copper cable is used. Its length is varied from zero to 10 km, which is beyond the BT limiting length (7.2 km).

The results are shown as curves of conversation opinion score, OLR and conversational speech voltage against line length. As before, the computed points are simply joined to form a line.

The calculations on this HRC have been included to demonstrate the "inverse" use of CATNAP. The limits on OLR are known (from the transmission plan) and so these runs could be used to show what range of cable lengths are acceptable. The facility for calculating the performance in terms of conversation opinion score makes it possible to specify performance limits in terms of this, which is closer to the real performance than limits set in terms of loudness ratings.



Transmission elements

Telephone instruments are BT Type No. 706

1 Unloaded cable 1.6 km of 0.5 mm (168 ohms/km, 50 nF/km)

2 Stone feed bridge (2 × 200 Ω, 2 + 2 μF, 50 V)

Near end		IL =	64
STMR =	9.02	SLR =	4.76
RLR =	−5.15	OLR =	−0.27
Y_{LE} =	3.48	V_L =	−18.24
Y_C =	3.15	V_C =	−22.69
RN =	40.00	ICN =	−65.00
IL =	64	*Far end*	
OLR =	−0.27	RLR =	−5.15
SLR =	4.76	STMR =	9.02
V_L =	−18.24	Y_{LE} =	3.48
V_C =	−22.69	Y_C =	3.15
ICN =	−65.00	RN =	40.00

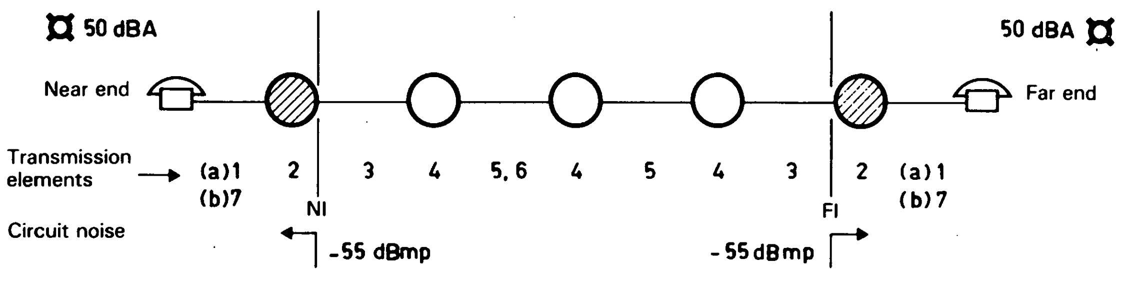
IL	Line current (mA)
SLR	Sending loudness rating (dB)
RLR	Receiving loudness rating (dB)
OLR	Overall loudness rating (dB)
STMR	Masked sidetone loudness rating (dB)
Y_{LE}	Listening effort score
Y_C	Conversation opinion score
V_L	Speech voltage at interface (dBV) under controlled talking conditions
V_C	Speech voltage at interface (dBV) under conversational conditions
RN	Level of room (environmental) noise (dBA), Hoth spectrum
ICN	Level of injected circuit noise referred to a 0 dB RLR receiving end
NI	Near interface
FI	Far interface
LTS	Local telephone system

Note 1 – Room noise has Hoth spectrum.

Note 2 – The OLR printed in the left column is for near to far and the OLR in the right column is for far to near.

FIGURE A-1

HRC 1 Own exchange call



Transmission elements

Telephone instruments are BT Type No. 706

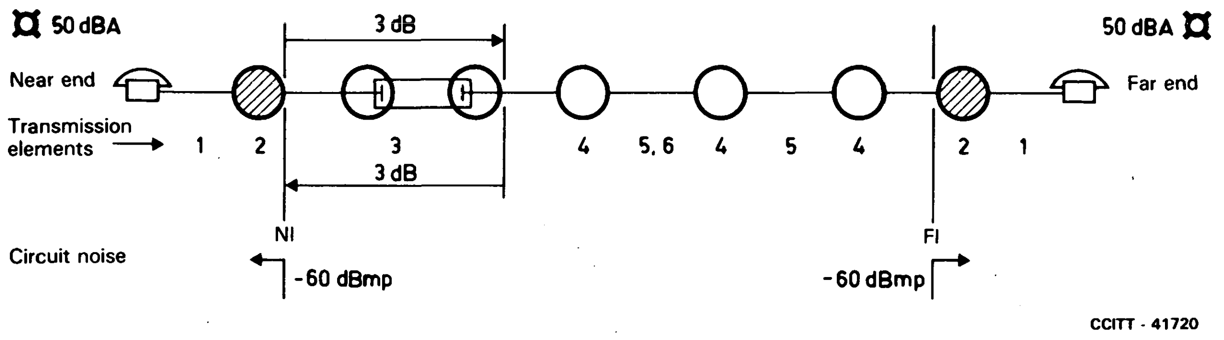
1 Unloaded cable 6 km of 0.5 mm (168 ohms/km, 50 nF/km)
2 Stone feed bridge (2 × 200 Ω, 2 + 2 μF, 50 V)
3 Loaded junction 19.6 km of 0.9 mm, 88 mH @1.83 km
4 Transformer feed bridge (50 V)
5 Attenuation 3.5 dB, frequency independent, 600 Ω
6 Channel filtering 300 Hz-3.4 kHz, 600 Ω
7 Unloaded cable 10 km of 0.9 mm (55 ohms/km, 50 nF/km)

Near end		*IL* =	32	*IL* =	32	*Far end*	
STMR =	11.19	SLR =	8.21	OLR =	25.07	RLR =	−1.32
RLR =	−1.32	OLR =	25.07	SLR =	8.21	STMR =	11.19
Y_{LE} =	1.98	V_L =	−21.40	V_L =	−21.40	Y_{LE} =	1.98
Y_C =	1.86	V_C =	−22.46	V_C =	−22.46	Y_C =	1.86
RN =	50.00	ICN =	−55.00	ICN =	−55.00	RN =	50.00

Near end		*IL* =	50	*IL* =	50	*Far end*	
STMR =	−0.14	SLR =	6.62	OLR =	24.04	RLR =	−2.05
RLR =	−2.05	OLR =	24.04	SLR =	6.62	STMR =	−0.14
Y_{LE} =	1.72	V_L =	−19.75	V_L =	−19.75	Y_{LE} =	1.72
Y_C =	0.81	V_C =	−21.52	V_C =	−21.52	Y_C =	0.81
RN =	50.00	ICN =	−55.00	ICN =	−55.00	RN =	50.00

FIGURE A-2

HRC 2 – Limiting national call



Transmission elements

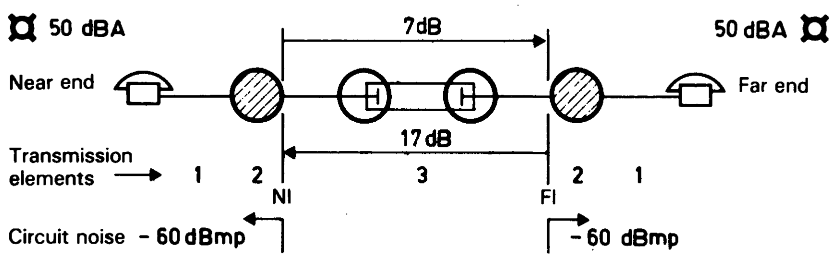
Telephone instruments are BT Type No. 706

1 Unloaded cable 1.6 km of 0.5 mm (168 ohms/km, 50 nF/km)
2 Stone feed bridge (2 × 200 Ω, 2 + 2 μF, 50 V)
3 PCM system 3 dB up to 3.4 kHz, 600 Ω
4 Transformer feed bridge (50 V)
5 Attenuation 3.5 dB, frequency independent, 600 Ω
6 Channel filtering 300 Hz-3.4 kHz, 600 Ω

Near end		*IL* = 64		*IL* = 64		*Far end*	
STMR =	9.31	SLR =	4.95	OLR =	13.38	RLR =	−4.95
RLR =	−4.95	OLR =	13.38	SLR =	4.95	STMR =	8.55
Y_{LE} =	3.34	V_L =	−18.60	V_L =	−18.60	Y_{LE} =	3.34
Y_C =	2.73	V_C =	−22.19	V_C =	−22.45	Y_C =	2.75
RN =	50.00	ICN =	−60.00	ICN =	−60.00	RN =	50.00

FIGURE A-3

HRC 3 – Long distance call with a PCM junction



CCITT - 41730

Transmission elements

Telephone instruments are BT Type No. 706

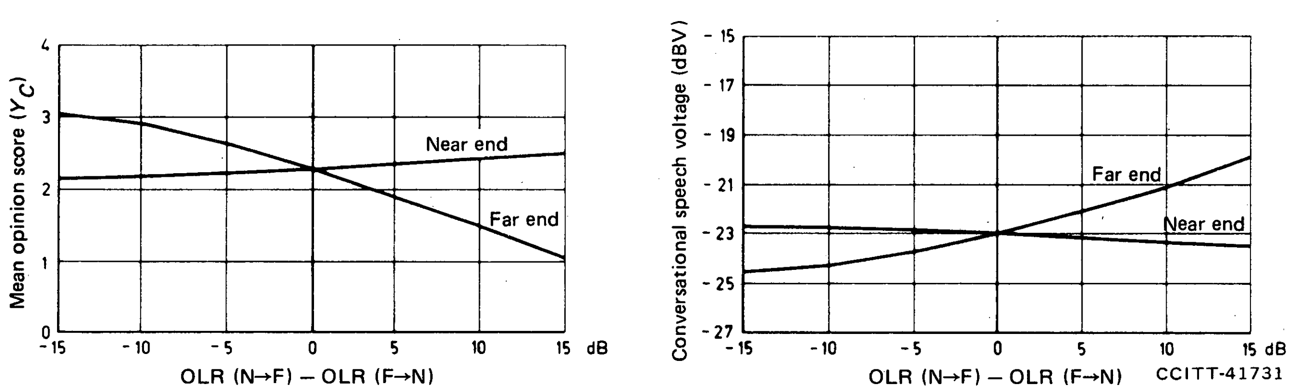
1 Unloaded cable 6 km of 0.5 mm (168 ohms/km, 50 nF/km)

2 Stone feed bridge (2 × 200 Ω, 2 + 2 μF, 50 V)

3 FDM system loss as shown up to 3.4 kHz, 600 Ω

Near end		*IL* =	32	*IL* =	32	*Far end*	
STMR =	13.89	SLR =	8.21	OLR =	14.35	RLR =	−1.31
RLR =	−1.31	OLR =	24.22	SLR =	8.21	STMR =	13.89
Y_{LE} =	2.62	V_L =	−21.40	V_L =	−21.40	Y_{LE} =	3.40
Y_C =	2.18	V_C =	−22.69	V_C =	−24.22	Y_C =	2.90
RN =	50.00	ICN =	−60.00	ICN =	−60.00	RN =	50.00

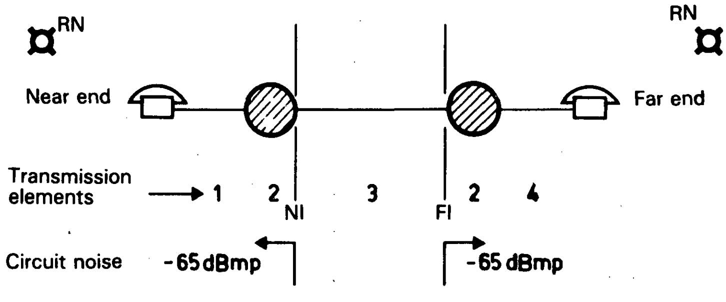
The results shown in the curves below are for the same connection with the loss from near to far in the FDM system varied from 2 dB to 32 dB. The loss from far to near was kept at 17 dB.



CCITT-41731

FIGURE A-4

HRC 4 – Effect of asymmetry of transmission loss



Transmission elements

Telephone instruments are BT Type No. 706

1 Unloaded cable 10 km of 0.9 mm (55 ohms/km, 50 nF/km)
2 Stone feed bridge (2 × 200 Ω, 2 + 2 μF, 50 V)
3 Attenuation 20 dB, frequency independent, 600 Ω
4 Unloaded cable 6 km of 0.5 mm (168 ohms/km, 50 nF/km)

Near end		*IL* =	50	*IL* =	32	*Far end*	
STMR =	0.28	SLR =	6.62	OLR =	24.69	RLR =	−1.32
RLR =	−2.04	OLR =	25.53	SLR =	8.82	STMR =	13.66
Y_{LE} =	2.63	V_L =	−19.75	V_L =	−21.40	Y_{LE} =	2.87
Y_C =	2.17	V_C =	−24.83	V_C =	−22.67	Y_C =	2.17
RN =	40.00	ICN =	−65.00	ICN =	−65.00	RN =	40.00

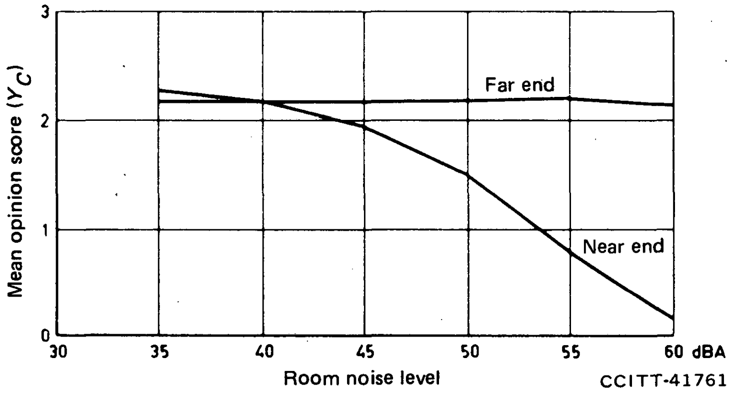
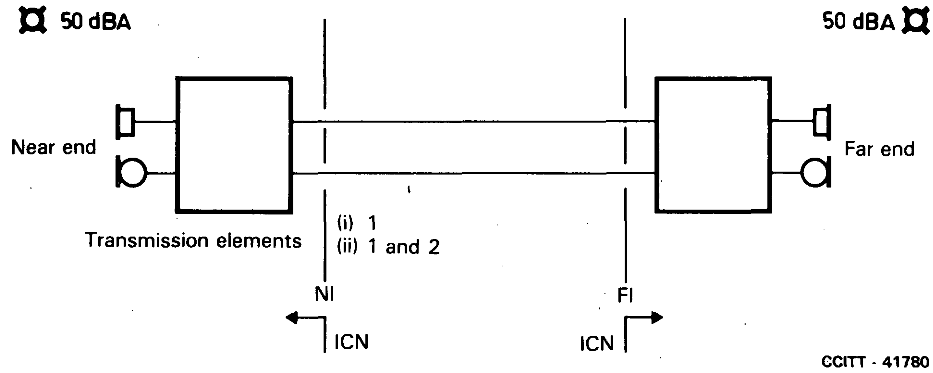


FIGURE A-5

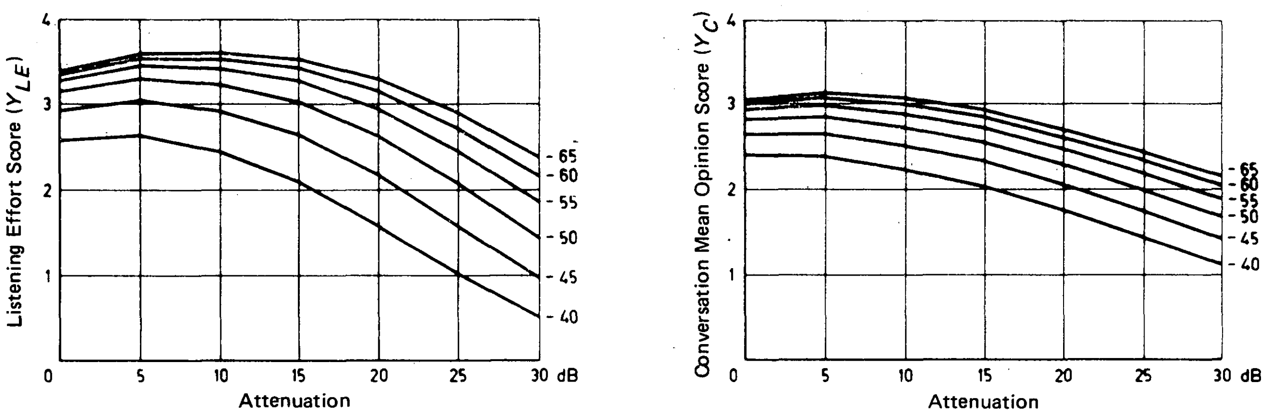
HRC 5 – Effect of room noise level



Transmission elements

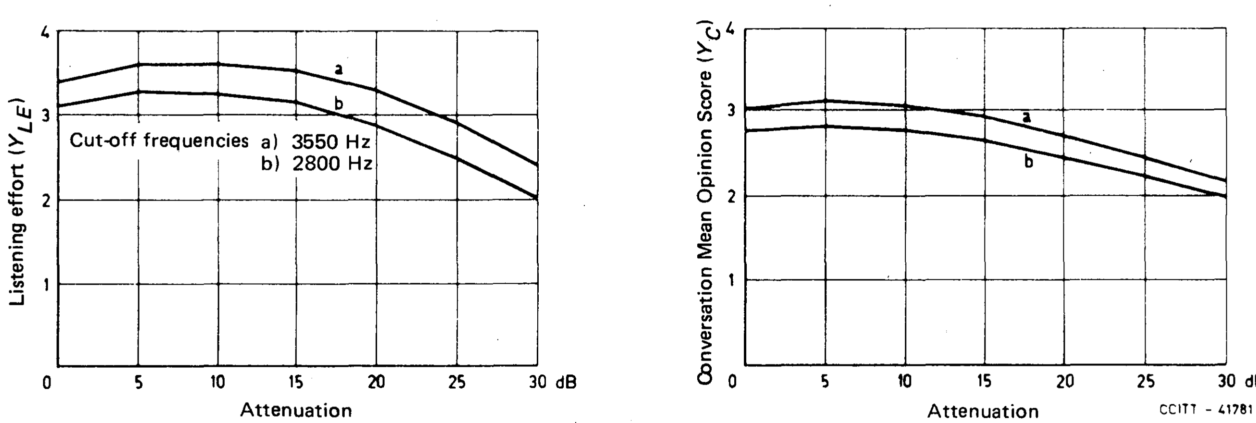
Telephone instruments are Intermediate Reference Systems (see Recommendation P.48) with a sidetone path of STMR = 20 dB.

1 Attenuation 0-30 dB, frequency independent, 600 Ω
2 Filtering 600 ohms, (a) 0-3.55 kHz
(b) 0-2.80 kHz



Note – These curves show the effect on Y_{LE} and Y_C of changing the level of injected circuit noise from −65 dBmp to −40 dBmp, referred to a 0 dB RLR.

a) Effect of injected circuit noise level and overall loss on the listening effort and conversation opinion scores.

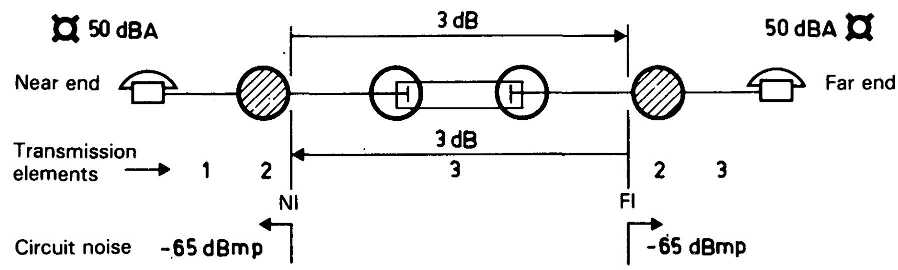


Note – These curves show the effect of bandlimiting with ideal lowpass filters.

b) Effect of bandlimiting (lowpass) and loss on the listening effort and conversation opinion scores.

FIGURE A-6

HRC 6 – Effect of injected circuit noise level and bandlimiting



CCITT-41831

Transmission elements

Telephone instruments are BT Type No. 706

1 Line: random selection from a sample of 1800 existing customers' lines.
2 Stone feed bridge (2 × 200 Ω, 2 + 2 μF, 50 V)
3 PCM system 600 Ω, 3 dB
4 Line: random selection of a line from the same sample of 1800 as in 1 above.

Line 43 (1.3 km)				Line 121 (0.9 km)			
Near end		*IL* =	64	*IL* =	64	*Far end*	
STMR =	10.38	SLR =	5.45	OLR =	4.77	RLR =	−4.44
RLR =	−4.46	OLR =	4.77	SLR =	5.47	STMR =	8.59
Y_{LE} =	3.56	V_L =	−19.07	V_L =	−19.11	Y_{LE} =	3.57
Y_C =	3.06	V_C =	−22.96	V_C =	−23.56	Y_C =	3.07
RN =	50.00	ICN =	−65.00	ICN =	−65.00	RN =	50.00

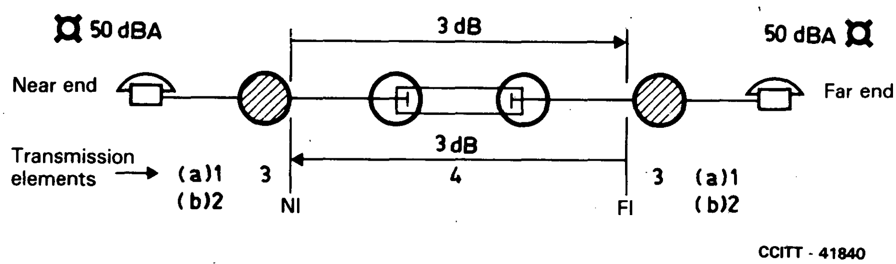
Line 731 (0.3 km)				Line 87 (0.5 km)			
Near end		*IL* =	75	*IL* =	64	*Far end*	
STMR =	7.19	SLR =	4.08	OLR =	2.42	RLR =	−5.41
RLR =	−5.74	OLR =	2.53	SLR =	4.50	STMR =	6.77
Y_{LE} =	3.45	V_L =	−17.46	V_L =	−18.16	Y_{LE} =	3.50
Y_C =	3.05	V_C =	−22.29	V_C =	−23.13	Y_C =	3.06
RN =	50.00	ICN =	−65.00	ICN =	−65.00	RN =	50.00

Line 4 (2.0 km)				Line 776 (0.9 km)			
Near end		*IL* =	50	*IL* =	75	*Far end*	
STMR =	4.33	SLR =	4.05	OLR =	2.45	RLR =	−5.38
RLR =	−4.65	OLR =	3.54	SLR =	4.45	STMR =	7.28
Y_{LE} =	3.53	V_L =	−17.84	V_L =	−17.83	Y_{LE} =	3.47
Y_C =	3.05	V_C =	−23.53	V_C =	−22.60	Y_C =	3.03
RN =	50.00	ICN =	−65.00	ICN =	−65.00	RN =	50.00

Line 1018 (2.2 km)				Line 1647 (2.5 km)			
Near end		*IL* =	50	*IL* =	40	*Far end*	
STMR =	8.95	SLR =	3.41	OLR =	4.37	RLR =	−2.72
RLR =	−5.27	OLR =	4.59	SLR =	6.18	STMR =	8.94
Y_{LE} =	3.54	V_L =	−17.17	V_L =	−19.54	Y_{LE} =	3.56
Y_C =	3.03	V_C =	−21.42	V_C =	−23.88	Y_C =	3.07
RN =	50.00	ICN =	−65.00	ICN =	−65.00	RN =	50.00

FIGURE A-7

HRC 7 – Example with random selection of customers' lines



Transmission elements

Telephone instruments are BT Type No. 706, with the values of Z_{so} modified as required

1 Unloaded cable: as specified below.
2 Line: random selection from a sample of 1800 exisiting customers' lines
3 Stone feed bridge (2 × 200 Ω, 2 + 2 µF, 50 V)
4 PCM system 600 Ω, 3 dB

Note – See also Tables A-1 and A-2.

FIGURE A-8

HRC 8 – Example of the use of CATNAP in design

TABLE A-1

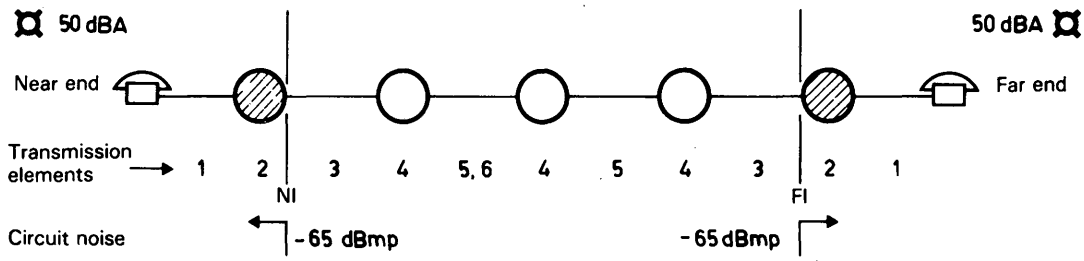
Values of STMR (dB) for specified lines (copper conductors)

Z_{so}	1.6 km 0.5 mm (median)	6 km 0.5 mm (limiting)	3.7 km 0.4 mm (limiting)	7.2 km 0.63 mm (limiting)	10 km 0.9 mm (limiting)
706	9.9	15.7	7.2	7.5	0.0
Conjugate of input Z	1.8	1.1	0.6	−0.2	−0.6
60 ohms	6.6	−0.8	−1.2	−2.0	−3.0
Suggested values	10.2	13.4	13.8	4.4	−1.3

TABLE A-2

Distribution of STMR for a sample of 40 lines for a Standard 706 and the suggested values of Z_{so}

Z_{so}	Mean	Standard deviation	Maximum value	Minimum value
706	8.3	±2.5	14.1	3.8
Suggested values	9.4	±3.1	17.9	4.2



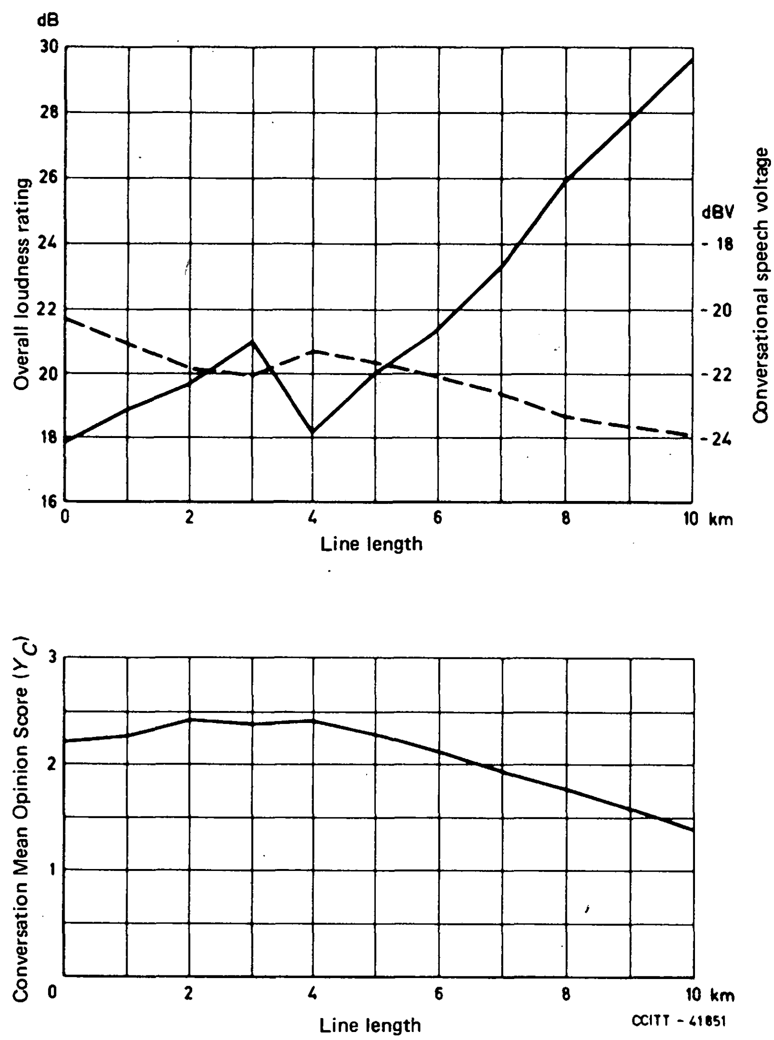
CCITT - 41850

Transmission elements

Telephone instruments are BT Type No. 706

1 Unloaded cable 0-10 km of 0.63 mm (109 ohms/km, 50 nF/km)
2 Stone feed bridge (2 × 200 Ω, 2 + 2 μF, 50 V)
3 Loaded junction 19.6 km of 0.9 mm, 88 mH at 1.83 km
4 Transformer feed bridge (50 V)
5 Attenuation 3.5 dB, frequency independent, 600 Ω
6 Channel filtering 300 Hz-3.4 kHz, 600 Ω

The results are shown in the curves.



CCITT - 41851

FIGURE A-9

HRC 9 – Effect of varying line length

References

[1] CCITT – Question 7/XII, Contribution COM XII-No. 1, Study Period 1981-1984, Geneva, 1981.

[2] *Transmission rating models*, Volume V, Supplement No. 3, ITU, Geneva, 1985.

[3] RICHARDS (D. L.): Calculation of opinion scores for telephone connections, *Proceedings of the I.E.E.*, Vol. 121, No. 5, pp. 313-323, May 1974.

[4] CCITT – Question 7/XII, Annex 2, Contribution COM XII-No. 1, Study Period 1977-1980, Geneva, 1977.

[5] *Some remarks on the structure of a telephone connection assessment model*, CCITT Contribution COM XII-No. 113, Study Period 1977-1980, Geneva, October 1978.

[6] *Use of a telephone connection assessment model in the study of Question 15/XII*, CCITT Contribution COM XII-No. 129, Study Period 1977-1980, Geneva, February 1979.

[7] WEBB (P. K.): The background and philosophy of the telephone network assessment program (CATNAP), British Post Office Research Department Report No. 752, 1979.

[8] *Prediction of transmission qualities from objective measurements*, Yellow Book, Volume V, Supplement No. 4, ITU, Geneva, 1981.

[9] *Methods used for assessing telephony transmission performance*, Vol. V, Supplement No. 2, ITU, Geneva, 1985.

[10] RICHARDS (D. L.): Telecommunication by speech: The transmission performance of telephone networks, Chapter 3, *Butterworths*, London, 1973.

[11] RICHARDS (D. L.): Telecommunication by speech: The transmission performance of telephone networks, Chapter 2, *Butterworths*, London, 1973.

[12] RICHARDS (D. L.): New definitions for loudness ratings, *Proceedings of the I.E.E.*, Vol. 119, No. 10, pp. 1429-1441, 1972.

[13] CCITT – Question 15/XII, Annex 1, § 3, Contribution COM XII-No. 1, Study Period 1981-1984, Geneva, 1981.

[14] CCITT – Question 15/XII, Contribution COM XII-No. 1, Study Period 1977-1980, Geneva, 1977.

[15] CCITT – Question 19/X11, Contribution COM XII-No. 1, Study Period 1981-1984, Geneva, 1981.

[16] RICHARDS (D. L.): Transmission performance of telephone networks containing P.C.M. links, *Proceedings of the I.E.E.*, Vol. 115, No. 9, pp. 1245-1258, September 1968.

[17] CCITT Handbook *Transmission planning of switched telephone network*, Chapter II, Annex 3, ITU, Geneva, 1976.

[18] BOERYD (A.): Subscriber reaction due to unbalanced transmission levels, *Third International Symposium on Human Factors in Telephony*, 1966, pp. 39-43, The Hague, 1967.

Supplement No. 5

THE SIBYL METHOD OF SUBJECTIVE TESTING

(Geneva, 1980)

(Quoted in Recommendation P.74)

(Contribution by the American Telephone and Telegraph Company)

SIBYL is the name given to the facility developed by Bell Laboratories for the study of human factors in communications systems. The use of SIBYL allows an experimenter to control transmission parameters during normal business calls of cooperating Bell Laboratories employees in a manner which ensures maintaining privacy of telephone conversations. (See [1], [2] and [3].)

The test subjects used in any experiment involving SIBYL are obtained by contacting employees before the test to ask if they would be willing to participate in the test. If they agree to participate, their telephone lines are routed through the SIBYL facility.

The present SIBYL facility is available at the Holmdel Bell Laboratories where each telephone set is connected to the serving central office over a 2-wire cable about 3 miles in length. In general, only internal (within Holmdel Bell Laboratories) telephone calls originated by the subjects are included in a SIBYL experiment. In addition to presenting controlled transmission conditions on a test call, SIBYL collects objective data such as timing information (e.g. time length of call) or speech levels.

A typical connection involving SIBYL is depicted in the diagram of Figure 1. At the left is the subject's telephone set connected by less than 1500 feet of 2-wire cable to SIBYL which converts the 2-wire transmission path into a 4-wire path and also separates the signals sent from the subject and the signals received by the subject. This permits the insertion of different transmission parameters in the two directions of transmission and also permits independent measurements of the transmitted and received signals.

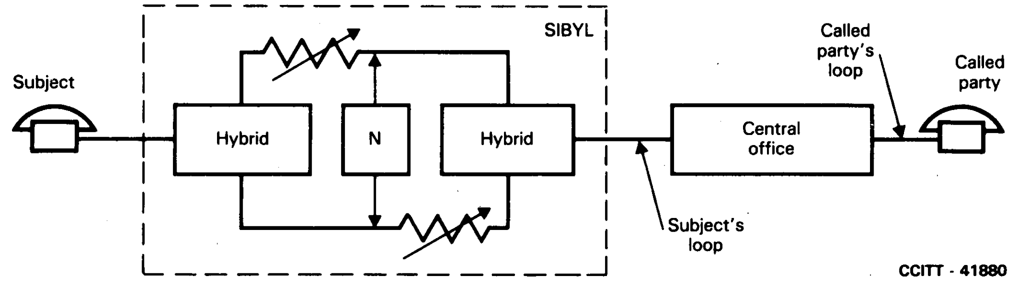


FIGURE 1

Diagram of a telephone connection with SIBYL

Moving from left to right through the diagram of Figure 1, the 4-wire path in SIBYL is converted back to a 2-wire path and connected to the serving central office over the subject's loop. The central office connects the subject's line to the called line dialled by the subject which is terminated with the telephone set of the called party who is another employee at the Holmdel Laboratories. If the called party also happens to be a subject for the experiment no transmission conditions are introduced into the called party's line.

Each subject is provided with instructions at the beginning of an experiment explaining the procedures involved in making a test call. A sample procedure is depicted in Figure 2. The subject initiates the procedure by taking his handset from the telephone set cradle and dialling a number. If the call is not a test call, it completes normally just as if SIBYL were not involved. This would happen if the called number was not that of another employee at Holmdel or if SIBYL was otherwise programmed not to intercept the call. If the call is a test call and the called line is not busy, the selected transmission conditions are inserted while the connection is being completed and the call proceeds. If the transmission condition is unacceptable to the subject at any time during the conversation, the subject can dial a special digit (i.e. 5 in Figure 2) which signals SIBYL to remove the selected transmission degradations from the call.

When the conversation is ended and the subject replaces the handset on the cradle, SIBYL rings the subject's telephone set with a short burst of sound. This alerts the subject to the fact that a test call has just been completed and that it is necessary to rate the transmission quality of the call by dialling one of five single digits indicating his rating on the five-point comment scale: Excellent, Good, Fair, Poor and Unsatisfactory. This rating information is recorded by SIBYL for later processing.

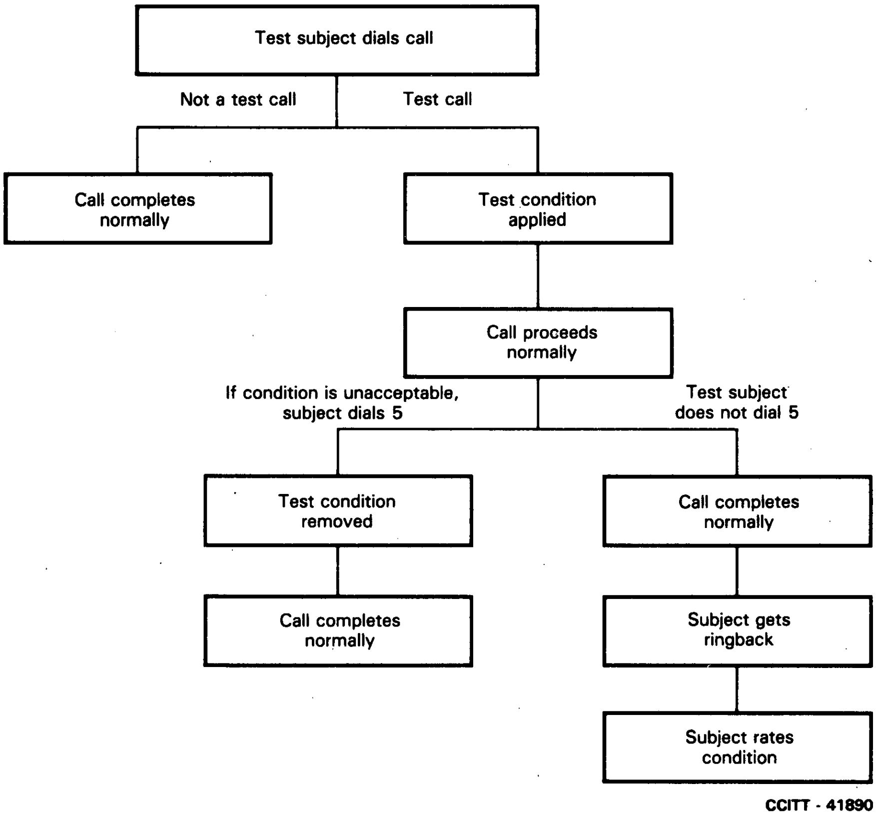


FIGURE 2

Step-by-step procedure for a call with SIBYL

References

[1] IRVIN (H. D.): Studying Tomorrow's Communication... Today, *Bell Labs Record*, pp. 399-402, November 1958.

[2] SULLIVAN (J. L.): Is Transmission Satisfactory? Telephone Customers Help Us Decide, *Bell Labs Record*, pp. 90-98, March 1974.

[3] CAVANAUGH (J. R.), HATCH (R. W.) and SULLIVAN (J. L.): Models for the Subjective Effects of Loss, Noise and Talker Echo on Telephone Connections, *Bell System Technical Journal*, 55, No. 9, pp. 1319-1371, November 1976.

Supplement No. 6

ATTENUATION OF THE ELECTRO-ACOUSTIC EFFICIENCY OF TELEPHONE SETS IN VIEW OF PROTECTION AGAINST ACOUSTIC SHOCKS

(Geneva, 1980)

It is also desirable for the protective devices to reduce efficiently the discomfort liable to result for the user of the telephone set from exceptionally high voltages on the subscriber line. For this purpose the electro-acoustic efficiency of the telephone set, expressed as the ratio between the acoustic pressure produced by the earphone on a suitable measuring device and the sinusoidal voltage applied at its terminals, should decline when the level of the electrical signal increases in relation to its value for the usual speech signals, which is taken as a reference.

Pending the conclusions of a full study of the subject, it is recommended that the values in Table 1 below should be observed for the attenuation, expressed in dB, of that efficiency, relative to the electrical level N applied to the receiver terminals.

Measurements are to be made for frequencies between 200 and 4000 Hz. The reference value of the electro-acoustic efficiency is that recorded when $N = -20$ dBm.

The acoustic pressure considered is that produced by the earphone of the receiver applied to an IEC 318 [1] artificial ear (see Recommendation P.51) with a weighting curve A.

TABLE 1

Voltage level at terminals, N (reference dB 0.775 V)	Attenuation of electro-acoustic efficiency (dB)
−20	0 (reference)
−10	< 0.5
0	⩽ 2
+10	> 6
+20	>12
+30	>18

Reference

[1] International Electrotechnical Commission Recommendation *An IEC artificial ear, of the wide band type, for the calibration of earphones used in audiometry*, IEC publication 318, Geneva, 1970.

Supplement No. 7

GENERATION OF THE ARTIFICIAL VOICE

(Malaga-Torremolinos, 1984)

(Contribution from the Italian Administration)

1 Introduction

This Supplement describes a possible generation process of the artificial voice according to Recommendation P.51.

The artificial voice is produced by alternatively activating two different sources, namely a glottal periodic wave and a random noise (Figure 1). While the glottal wave produces voiced sounds and contains almost all the power of the signal, the random noise reproduces the unvoiced sounds and lasts on the average only for 10% of the time.

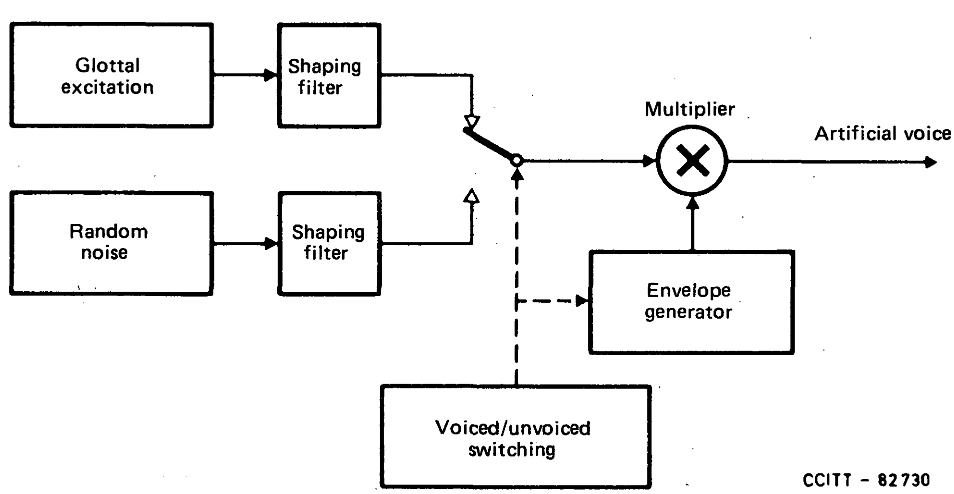
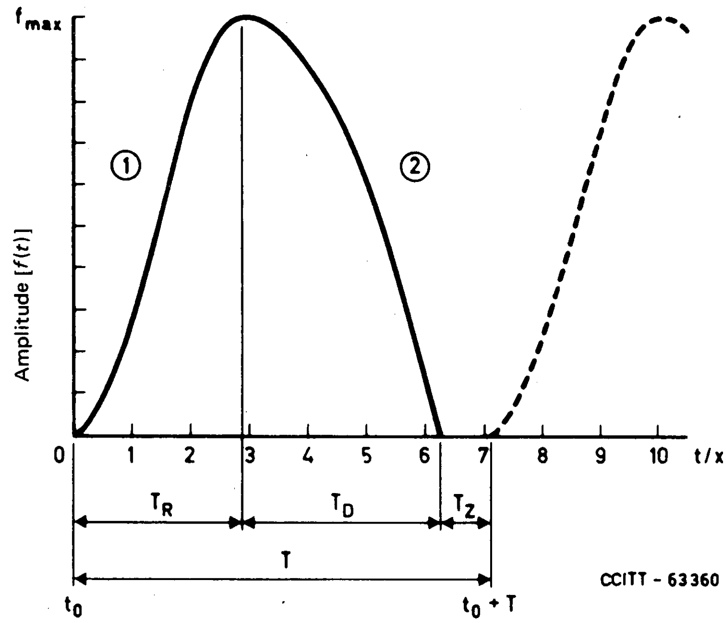


FIGURE 1

2 Glottal excitation

The waveform of the glottal excitation signal is shown in Figure 2. The pitch is time-dependent, according to the time variation of the parameter x: every 200 ms the x value is randomly varied, according to a normal distribution centred around a mean value of 1 and with a standard deviation of 0.11. In addition, it is allowed to vary only within 0.835 and 1.165 in order to reproduce the typical pitch variations of natural voice (120 Hz-168 Hz).



T_R : Raise time = 2.85 x (ms)
T_D : Decay time = 3.42 x (ms)
T_Z : Zero time = 0.855 x (ms)
x : Slowly varying randomly distributed variable (see text)

1: Build-up : $f(t) = f_{max}\left[3\,\dfrac{(t-t_0)^2}{T_R} - 2\,\dfrac{(t-t_0)^3}{T_R}\right]$

2: Decay : $f(t) = f_{max}\left[1 - \dfrac{(t-t_0-T_R)^2}{T_D}\right]$

FIGURE 2

The variation of the x parameter from each value to the next one, randomly generated 200 ms later, is smoothly accomplished following a semiperiod of a cosine waveform.

2.1 *Shaping filter*

The shaping filter processes the glottal excitation signal in order to obtain the desired long-term average spectrum (see Figure 3/P.51). Its response characteristic closely depends on the glottal excitation spectrum and on the long-term spectrum recommended in Recommendation P.51, § 3.4.1.

3 Unvoiced sounds

Unvoiced sounds are generated from a random noise generator, followed by a shaping filter giving to the signal the spectrum specified in Recommendation P.51, § 3.4.5.

4 Syllabic rate and time configuration of the envelope

The envelope of both voiced and unvoiced sounds is created in such a way as to reproduce average real speech characteristics; while the voiced sounds have a sawtooth envelope, where the rise time is much shorter than the decay time, the unvoiced sounds have an envelope characterized by a half-period sinusoidal waveform. Every nine sawtooth vocal envelopes, one half-wave sinusoidal unvoiced envelope is generated. The characteristics of the envelope waveforms are specified as follows:

4.1 *Voiced sounds*

Voiced sounds have an envelope characterized by triangular waves as shown in Figure 3.

The amplitude A is normally distributed between A_{max} and 0.3 A_{max} with standard deviation S_a = 0.25 A_{max}.

The gradient of the rising slope RS is normally distributed between RS_{max} = 64 ms/A_{max} and RS_{min} = 16 ms/A_{max} with standard deviation S_r = 16 ms/A_{max}.

The gradient of decay slope DS is normally distributed between DS_{max} = 576 ms/A_{max} and DS_{min} = 144 ms/A_{max} with standard deviation S_d = 144 ms/A_{max}.

The silence time T_Z between two consecutive triangular envelopes is normally distributed between $T_{Z max}$ = 17 ms and $T_{Z min}$ = 7 ms with standard deviation S_t = 5 ms.

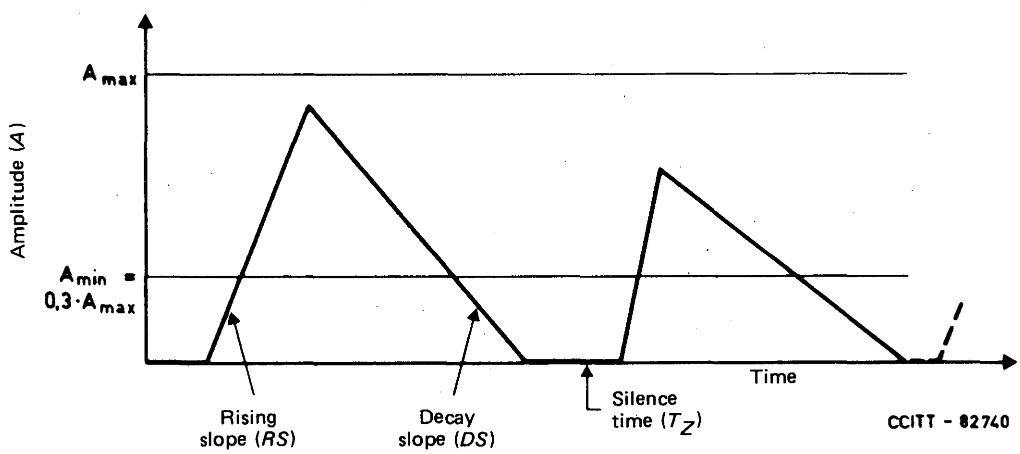


FIGURE 3

Envelope waveform of voiced sounds

4.2 *Unvoiced sounds*

Unvoiced sounds are envelopped by half-period sinewaves whose amplitude and duration are related to those of voiced envelopes. In particular, the duration of each unvoiced burst is equal to the last occurring voiced envelope, while the amplitude is reduced by 15 dB.

OBJECTIVE MEASUREMENT OF ACTIVE SPEECH LEVEL

(Malaga-Torremolinos, 1984)

(quoted in Recommendation P.52)

Preamble

It is considered important that there be a standardized method of measuring speech level in telecommunication links, so that measurements made by different Administrations may be directly comparable. This requires a meter that is not dependent on operator interpretation.

This meter should ideally be such as to maintain maximum comparability and continuity with past work: in particular, that the new method should yield data and conclusions compatible with those that have established the conventional value (22 microwatts) of speech power at the input to the 4-wire point of the international circuit according to Recommendation G.223.

The active speech level should be measured and reported in decibels relative to a stated reference according to the methods described below, namely:

- Method A: measuring a quantity called speech volume, used for the purpose of real-time control of speech level,
- Method B: measuring a quantity called active speech level, used for other purposes.

A suggested form of terminology is as follows:

speech volume: hitherto used interchangeably with speech level, should henceforth be used exclusively to denote a value obtained by Method A;

active speech level: should be used exclusively to denote a value obtained by Method B;

speech level: should be used as a general term to denote a value obtained by any method yielding a value expressed in decibels relative to a stated reference.

The definitions of terms such as these, e.g. see [1], and other related terms such as those for the meters themselves [2], should be adjusted accordingly.

1 General

1.1 *Introduction*

Recent studies [3,6,9] have shown that it is possible to construct devices fulfilling most of the functions listed as desirable in § 1 of Question 26/XII (Study Period 1981-1984). Most of these devices are different embodiments of the same general principle – that of measuring the mean power of speech averaged over the aggregate of time during which the speech is present ("active time"). However, the minimum active time required to give a stable mean is of the order of a few seconds, which is too long for the purpose of giving an immediate direct-reading indication suitable for rapid real-time adjustment of level by a talker or operator. It is possible, though not yet definitely established, that an instrument suitable for this purpose could be based on the same fundamental principle, implemented in some rapidly adaptive form [9]. But because of this uncertainty, this Supplement confines itself to describing the method to be used for three other functions mentioned in Question 26/XII (Study Period 1981-1984), leaving real-time control of speech level to be governed by the devices listed in Recommendation P.52 pending further study.

1.2 *Electrical, acoustic and other levels*

This Supplement deals primarily with electrical measurements yielding results expressed in terms of electrical units, generally decibels relative to an appropriate reference value such as one volt. However, if the calibration and linearity of the transmission system in which the measurement takes place are assured, it is possible to refer the result backwards or forwards from the measurement point to any other point in the system, where the signal may exist in some non-electrical form – acoustical, for example. Power is proportional to squared voltage in the electrical domain, squared sound pressure in the acoustical domain, or the digital equivalent of either of these in the numerical domain, and the reference value must be of the appropriate kind (1 volt, 1 pascal, reference acoustic pressure equal to 20 micropascals, or any other stated unit, as the case may be).

1.3 *Universal requirements*

For speech-level measurements of all types, the information reported should include: the designation of the measuring system, the method used (A, B or B-equivalent as explained in § 4, or other specified method), the quantity observed, the units, and other relevant information such as the margin value (explained below) where applicable.

All the relevant conditions of measurement should also be stated, such as bandwidth, position of the measuring instrument in the communication circuit, and presence or absence of a terminating impedance. Apart from the stated band limitation intended to exclude spurious signals, no frequency weighting should be introduced in the measurement path (as distinct from the transmission path).

1.4 *Averaging*

Where an average of several readings is reported, the method of averaging should be stated. The mean level (mean speech volume or mean active speech level), formed by taking the mean of a number of decibel values, should be distinguished from the mean power, formed by converting a number of decibel values to units of power, taking the mean of these, and then optionally restoring the result to decibels.

Any correction that has been applied should be mentioned, together with the facts or assumptions on which any such correction is based. For example, in loading calculations, when the active levels or durations of the individually measured portions of speech differ widely, $0.115\sigma^2$ is commonly added to the median or mean level in order to estimate the mean power, on the ground that the distribution of mean active speech levels (dB values) is approximately Gaussian.

2 **Method A: Immediate indication of speech volume for real-time applications**

Measurement of speech volume for rapid real-time control or adjustment of level by a human observer should be accomplished in the traditional manner by means of one of the devices listed in Recommendation P.52.

The choice of meter and the method of interpreting the pointer deflexions should be appropriate to the application, as in Table 1.

Values obtained by Method A should be reported as speech volume; the meter employed, the quantity observed, and the units in which the result is expressed, should be stated.

TABLE 1

Application	Meter	Quantity observed
Control of vocal level in live-speech loudness balances	ARAEN volume meter (SV3)	Level exceeded once in 3 s
Avoidance of peak limiting	Peak programme meter	Highest reading
Maintenance of optimum level in making magnetic tape recordings	VU meter	Average of peaks (excluding the most extreme peaks)

3 **Method B: Active speech level for other applications**

3.1 *Principle of measurement*

Active speech level is measured by integrating a quantity proportional to instantaneous power over the aggregate of time during which the speech in question is present (called the active time), and then expressing the quotient, proportional to total energy divided by active time, in decibels relative to the appropriate reference.

The mean power of a speech signal when known to be present can be estimated with high precision from samples taken at a rate far below the Nyquist rate, but the all-important question is what criterion should be used to determine when speech is present.

Ideally, the criterion should indicate the presence of speech for the same proportion of time as it appears to be present to a human listener, excluding noise that is not part of the speech (such as impulses, echoes, and steady noise during periods of silence), but including those brief periods of low or zero power that are not perceived as interruptions in the flow of speech [5]. It is not essential that the detector should operate exactly in synchronism with the beginnings and ends of utterances as perceived: there may be a delay in both operating and releasing, provided that the total active time is measured correctly. For this reason, complex real-time voice-activity detectors depending on sampling at the Nyquist rate, such as those that have been successfully used in digital speech interpolation, are not necessarily the most suitable for this application. Their function is to indicate when a channel is available for transmission of information: this state does not always coincide with the absence of speech. On one hand, it may occur during short intervals that ought to be considered part of the speech, and on the other hand, it may be delayed long after the end of an utterance (for reasons of convenience in the allocation of channels, for example).

The description given below shows how a very simple detection method can be made to meet the requirements satisfactorily. Even though the method involves applying a signal-dependent threshold which cannot be specified in advance, so that accurate results cannot be guaranteed while the measurement is actually in progress, nevertheless by accumulating sufficient information during the process it is possible to apply the correct threshold retroactively, and hence to output a correct result almost as soon as the measurement finishes. Continuous adaptation of the threshold level in real time appears to yield similar results in simple cases, but further study is needed to find out how far this conclusion can be generalized.

3.2 *Details of realization*

The algorithm for Method B is as follows.

Let the speech signal be sampled at a rate not less than f samples per second, and quantized uniformly into a range of at least 2^{12} quantizing intervals (i.e. using 12 bits per sample including sign).

Note – This requirement ensures that the dynamic range for instantaneous voltage is at least 66 dB, but two factors combine to make the range of measurable active speech levels about 30 dB less than this:

1) Allowance must be made for the ratio of peak power to mean power in speech, namely about 18 dB for the value exceeded with probability 0.001 [4].

2) Envelope values down to at least 16 dB below the mean active level must be calculated: these values may be fractional, but will not be accurate enough if computed from a quantizing interval much exceeding twice their own voltage, i.e. it should not be expected that an active speech level less than about 10 dB above the quantizing interval would be measurable.

Let the successive sample values be denoted by x_i where i = 1, 2, 3.... Let the time interval between consecutive samples be $t = 1/f$ seconds.

Other constants required are:

v	(volts/unit) scale factor of the analogue-digital converter
T	time constant of smoothing in seconds
$q = \exp(-t/T)$	coefficient of smoothing
H	hangover time in seconds
$I = H/t$	rounded up to next integer
M	(margin) dB difference between threshold and active speech level.

Let the input samples be subjected to two distinct processes, 1 and 2.

Process 1

Accumulate the number of samples n, the sum s, and the sum of squares sq:

$$n_i = n_{i-1} + 1$$

$$s_i = s_{i-1} + x_i$$

$$sq_i = sq_{i-1} + x_i^2$$

where s_0, sq_0 and n_0 (initial values) are zero.

Process 2

Perform two-stage exponential averaging on the rectified signal values:

$$p_i = g \times p_{i-1} + (1-g) \times |x_i|$$

$$q_i = g \times q_{i-1} + (1-g) \times p_i$$

where p_0 and q_0 (initial values) are zero.

The sequence q_i is called the envelope, p_i denotes intermediate quantities.

Let a series of fixed threshold voltages c_j be applied to the envelope. These should be spaced in geometric progression, at intervals of not more than 2:1 (6.02 dB), from a value equal to about half the maximum code down to a value equal to one quantizing interval or lower. Let a corresponding series of activity counts a_j, and a corresponding series of hangover counts, h_j, be maintained. For each value of j in turn:

if $q_i > c_j$ or $q_i = c_j$, *then* add 1 to a_j and set h_j to 0;

if $q_i < c_j$ and $h_j < I$, *then* add 1 to a_j and add 1 to h_j;

if $q_i < c_j$ and $h_j = I$, *then* do nothing.

In the first case the envelope is at or above the jth threshold, so that the speech is active as judged by that threshold level. In the second case the envelope is below the threshold, but the speech is still considered active because the corresponding hangover has not yet expired. In the third case the speech is inactive as judged by the threshold level in question.

All the a_j and the h_j values are initially set equal to zero.

It should be noted that the suffix $_i$ in all the above cases is needed only to distinguish current values from previous values of accumulated quantities: for example, there is no need to hold more than one value of sq, but this value is continually updated. At the end of the measurement, therefore, the suffixes can be omitted from s, sq, n, p, and q.

Let all these processes continue until the end of the measurement is signalled. Then evaluate the following quantities:

$$\text{Total time} = n \times t$$

$$\text{Long-term power} = sq \times v^2/n$$

Note – If it is suspected that there may be significant d.c. offset, this may be estimated as $s \times v/n$, and used to evaluate a more accurate value of long-term power (a.c.) as $v^2\,[sq/n - (s/n)^2]$. However, in this case, the effect of the offset on the envelope must also be taken into account and appropriate corrections made.

For each value of j, the active-power estimate equals $sq \times v^2/a_j$.

At this stage the powers are in volts squared per unit time. Now express the long-term power and the active-power estimates in decibels relative to the chosen reference voltage r:

Long-term level	$L = 10 \log_{10} (sq \times v^2/n) - 20 \log_{10} r$
Active-level estimate	$A_j = 10 \log_{10} (sq \times v^2/a_j) - 20 \log_{10} r$
Threshold	$C_j = 20 \log_{10} (c_j \times v) - 20 \log_{10} r$

For each value of j, compare the difference $A_j - C_j$ with the margin M, and determine (if necessary, by interpolation on a decibel scale between two consecutive values of A_j and of C_j) the true active level A and corresponding threshold C for which $A - C = M$. If one of the pairs of values A_j and C_j fulfills this condition exactly then the true activity factor is a_j/n, but in all cases it can be evaluated from the expression $10^{(L-A)/10}$.

For simplicity the algorithm has been defined in terms of a digital process, but any equivalent process (one implemented on a programmable analogue computer, for example) should also be considered to fulfill the definition.

3.3 *Values of the parameters*

The values of the parameters in Table 2 should be used. They have been found suitable for the purpose and have stood the test of many years of application by various organizations [5], [6].

TABLE 2

Parameter	Value	Tolerance
f	694 samples/s	Not less than 600
T	0.03 s	± 5%
H	0.2 s	± 5%
M	15.9 dB	± 0.5

Note – The value M = 15 dB might appear to be implied in [5], but the threshold level described there equals the mean absolute voltage of a sine wave whose mean power is 15 dB below the reference. The difference of 0.9 dB is 20 $\log_{10}$ (r.m.s. voltage/mean absolute voltage) for a sine wave.

The result of a measurement made by means of the above algorithm with parameter values conforming to the above restrictions should be reported as *active speech level*, and the system should be described as *using Method B* of this Supplement. Where noise levels are very high, as they are for example in certain vehicles or in certain radio systems, it is often desirable to set the threshold higher (i.e. use a smaller margin) in order to exclude the noise: this may be done provided the margin is also reported. The result of such a measurement should be reported as *active speech level with margin M*, and the measurement system described as *using Method B with margin M.*

The activity factor should preferably be reported as a percentage, with a specification of the margin value if this is outside the standard range.

Note that certain measurement systems with fixed thresholds, instead of the retroactively selected threshold as described above, may still give active speech level according to the above definition in cases when the margin turns out to be within the specified limits. Otherwise the margin must be quoted with each reading. This applies to some of the methods described in reference [3], including the SV5B [5].

4 Approximate equivalents of Method B

Other methods under development, some of which are described in [3], use a broadly similar principle of measurement but depart in detail from the algorithm given above.

One such method uses a threshold based on peaks of the instantaneous waveform, using a larger margin, no smoothing, and a minimum time above threshold for detection of speech. Since peaks are more variable than the mean power, this method seems likely to give less stable results, though parameter values could no doubt be chosen to yield results agreeing on average with those obtained from Method B.

Another method, otherwise similar to Method B, omits the accumulation of x^2 when the detector indicates absence of speech. This procedure excludes not only noise energy but also a substantial part of the speech energy at the beginning of each utterance, because the delay inherent in the envelope-extraction process leads to a time-lag in operating the detector. The resultant levels are too low, especially for speech comprising a large number of short utterances.

A third method [7] fixes the threshold by examining the distribution of amplitudes at the end of the utterance, leaving the temporal features of the speech structure out of account. It is not clear how the results of this method are affected by activity factor and noise level.

It is not the intention of exclude any such method entirely, provided it is convincingly shown by experimental evidence to yield results consistent with those obtained by Method B in a sufficiently wide range of conditions. For this reason a class of methods called *B-equivalent methods* is recognized.

A B-equivalent method of speech-level measurement is defined as any method that satisfies the following test in all respects.

Measurements shall be carried out simultaneously by the method in question and by Method B on two or more samples of speech in every combination of the following variables:

Voices	one male and one female voice;
Speech material	a list of independent sentences, a passage of continuous speech, and one channel of a conversation, each lasting at least 20 seconds (active time);
Bandwidth	300-3400 Hz and 100-8000 Hz;
Added noise	flat within the measurement band at levels (M + 5) dB and (M + 25) dB below the active speech level, where M (the margin) is normally 15.9 dB, but smaller in high-noise applications;
Levels	at intervals of 10 dB over the range claimed for the system in question.

From the results, 95% confidence limits for the difference between the level given by the method in question and the active speech level given by Method B shall be calculated for each of the above 24 combinations.

If for every combination the upper confidence limit of this difference is not higher than +1 dB and the lower confidence limit is not lower than −1 dB, then the method shall be deemed to be a B-equivalent method.

Further, a method qualifies as B-equivalent if it gives results that fall within the specified limits when corrected by the addition of a fixed constant, known in advance of the measurement and not dependent on any feature of the speech signal (except possibly the bandwidth if this is known independently). The EPL method [8] is an example of a method that can be expected to meet this criterion.

The results of measurements by such a method should be reported as *B-equivalent active speech level*, and the activity factor as *B-equivalent activity factor.*

5 Conclusion

Two methods of measuring speech level have been suggested: Method A for real-time human control, and Method B for other uses; where Method B is not practical, allowance is made for the use of demonstrably equivalent methods. Method B gives substantially the same results as the methods that were used to establish the speech contribution to the conventional load.

References

[1] CCITT definition *Volume*, Terms and Definitions, Vol. X, Fascicle X.1.

[2] CCITT definitions *Electrical speech level meter, SFERT speech level meter, ARAEN speech level meter*, Terms and Definitions, Vol. X, Fascicle X.1.

[3] Results of speech-level measurements by various methods on speech material supplied by British Telecom, Contribution of British Telecom to study of Question 26/XII (April 1984).

[4] RICHARDS (D. L.): Telecommunication by speech, *Butterworths*, § 2.1.3.2, pp. 56-69, London, 1973.

[5] BERRY (R. W.): Speech-volume measurements on telephone circuits, *Proc. IEE*, Vol. 118, No. 2, pp. 335-338, February 1971.

[6] CARSON (R.): A digital speech voltmeter – the SV6, *British Telecommunications Engineering*, April 1984.

[7] CCITT Contribution COM XII-No. 43 (A method for speech level measurements using IEC-interface bus and calculation) Study Period 1981-1984, Geneva, 1982.

[8] BRADY (P. T.): Equivalent peak level: a threshold-independent speech level measure, *Journal of the Acoustical Society of America*, Vol. 44, pp. 695-699, 1968.

[9] STEWART (J. M.): A digital speech voltmeter using a microprocessor, M.Sc. Thesis, Imperial College, London, 1978.

PROVISIONAL METHODS OF TESTING HEADSETS SO THAT P-SERIES RECOMMENDATIONS MAY BE USED

(Malaga-Torremolinos, 1984)

1 *Introduction*

This Supplement discusses the precautions recommended when testing headsets so that the procedures developed for analysis of conventional telephone sets can apply also to headsets. CCITT Study Group XII has been studying the subject of testing of headsets [operator telephone system (OTS)] under Question 3/XII.

2 **Definition of a headset; operator telephone set**

For the definition of these terms, see § A.1.

3 **Elements of the testing method**

3.1 In contrast to previous proposals for headset testing procedures, where a complete set of procedures just for headset testing were envisioned, this Supplement highlights the several areas of differences between headsets and conventional handsets so that data can be obtained which can then be analyzed by the established subjective and objective procedures and recommendations.

3.2 There are many similarities between headsets and conventional handsets. Each has a transducer system for converting voice signals into sending electrical signals. Each has a transducer system for converting received electrical signals into sound pressure for the user's ear. Each system is ultimately conditioned to a common 2-wire standard.

3.3 Briefly the differences between headsets and handsets that affect testing are as follows:

a) The majority of headsets are adjustable and have no fixed distance between the receiver and the microphone, as is the case with conventional handsets. Thus headset microphone positioning rules can be different.

b) Headsets with other than conventional on-ear (supra-aural) receivers are normally measured on ear simulators with smaller cavity size than the on-ear receiver ear simulator.

c) Headsets with other than conventional on-ear types of receivers cannot be simply held to the head with a hand with any certainty of acoustic parameters, thus requiring a different subjective test method for these headsets (contra-lateral balance method).

d) Headsets with other than conventional on-ear receivers require a method for converting receiver test results to a reference point to enable analysis. It would simplify analysis if this reference point is the same as the ear reference point (ERP) used for analysis of conventional on-ear (supra-aural) receivers.

e) Headsets are 4-wire devices. A 4-wire to 2-wire network needs to be specified to allow the headset to be tested as a system in the manner of telephone sets.

f) Headsets typically have been leaders in the use of electronic "signal-conditioning" which requires some precautions in testing to achieve correct results.

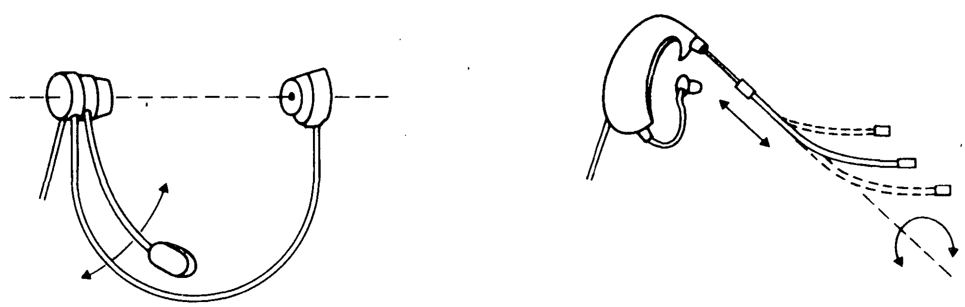
These areas are discussed below in greater detail.

4 Basic testing philosophies of headsets

4.1 *Non-adjustable versus adjustable headsets – mouth to ear distance; two broad categories*

4.1.1 Those headsets which are *not* adjustable in distance from the receiver (or ear centre) to the microphone require that the microphone be adjusted to a position relative to the mouth lip-ring as per established median dimensions based on measurement of a population of user's head sizes. The procedures established for positioning of conventional telephone sets can be used with suitable modification for establishing the equivalent of the receiver cap plane surface to the headsets that do not have supra-aural receivers. [See Figure 1*a)*.]

4.1.2 Those headsets which *are* adjustable in distance from the receiver (or ear centre) to the microphone should be adjusted to the subjective talker's lip-ring or the artificial mouth's lip-ring per the headset manufacturer's or Administration's user wearing instructions. [See Figure 1*b)*.]



a) Headset with non-adjustable ear to mouth distance

b) Headset with adjustable ear to mouth distance

FIGURE 1

4.2 *Converting receiver test results to the ERP (ear reference point) to be used with Series P Recommendations*

Objective calculation methods for receiver rating would be greatly simplified if a fixed reference point for measurement of headset receiver sound pressure could be used, or sound pressure readily related to it. For on-ear (supra-aural) receivers, this point is the ear reference point (ERP) which is defined as the centre of the plane of the circular receiver earcap when it is placed comfortably against the ear. With insert type receivers, it is not possible to directly measure the sound pressure at the ERP because it is outside of the ear insert and ear canal.

It is proposed by use of narrow frequency bands of noise signal and contra-lateral balancing technique to estabish correction factor tables for ear insert type receivers which would be very similar to the "L_E" table now used for handset receiver reference equivalent calculation (see Recommendation P.79) which would correct the data to the ERP [1].

4.3 *Subjective loudness testing of headsets – contra-lateral balance method*

Loudness balance rating of receivers is traditionally done by balancing the unknown receiver system against the NOSFER system by holding the unknown receiver and the NOSFER receiver in one hand and alternately placing the receivers to one ear (typically the left ear) while balancing the loudness. This method is not appropriate to headsets, and particularly headsets with ear-insert type of receivers because of the probability of breaking the headset receiver seal in the process of removing and restoring the headset to one ear.

Thus the contra-lateral method of loudness balancing is one recommended method. With this method, the headset receiver is carefully applied to one ear and the NOSFER receiver is applied to the other ear, the balance made and recorded. The receivers are then interchanged and the balance made. The two balances are then averaged for the loudness balance used.

Data from contra-lateral balance tests typically shows the left ears to be on the average 1 to 2 dB more sensitive than the right ear. This is at present believed to be caused by the fact that test teams typically use the left ear for balancing and thus this ear gets more "practice" than the right ear. If this is true, providing additional practice for the right ear should correct the problem. It is suggested that the average difference of the overall balances be presented with the data, and corrective action be considered if the average difference between the left and right ear balances exceeds 4 dB.

4.4 *Rating of headset receiver sensitivity – acoustic leaks*

The following test philosphy for headset receiver objective testing extends existing practices used for supra-aural receiver testing, with modification for special headset receiver systems.

4.4.1 If the receiver system is of a type with which a seal can easily be made by the user to his ear, then the objective testing is done with the sealed condition.

4.4.2 If the receiver system is of a type in which an acoustic leak is an unavoidable result of the design of the receiver coupling system, then suitable nominal acoustic leakage should be used in the objective testing. The receiver category termed "intra-concha" type generally falls in this category.

4.5 *Rating of receive sensitivity of headsets with dual receivers*

Receive sensitivity of headsets with dual receivers cannot be rated directly with conventional or contra-lateral loudness balancing techniques. The following method for rating dual-receiver sensitivity is recommended for study.

Rate the headset with one receiver of the two in use by the contra-lateral balance method. Establish a subjective improvement factor from adding the second receiver. Use this factor to correct the value of the single receiver sensitivity rating.

5 4-wire to 2-wire coupling networks

5.1 Headsets are 4-wire devices. Although there may be circuitry from the sending side to the receiving side for the purpose of powering the receive-electronics, there is no deliberate audio path between the two sides. To test the headset, it is thus necessary to provide a 4-wire to 2-wire network which is an integral part of conventional telephone sets. This 4-wire to 2-wire network fulfills the following functions:

a) transforms the headset sending impedance and the headset receiver impedance to the correct 2-wire line impedance – 600 ohms in the case of CCITT testing;
b) provides the path for electrical current from the 2-wire circuit to the sending side of the headset;
c) couples the signal from the sending side of the headset to the 2-wire circuit;
d) couples the signal from the 2-wire circuit to the receiving side of the headset;
e) attenuates the signal from the headset sending side to the headset receiving side to the proper level specified ("side-tone" level).

An example of a coupling network from British Telecom is given in Figure 2.

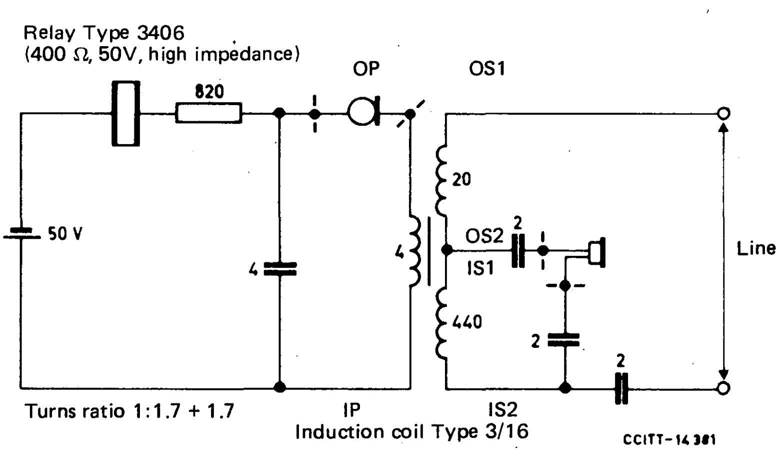


FIGURE 2

4-wire to 2-wire network for testing headsets (British Telecom Operator's network)

5.2 There are two general types of headset applications: those for telephone "operators" such as at telephone Administration offices and those in which the headset is an adjunct of a telephone instrument for "hands-free" usage by reservation agents, for example.

The 4-wire to 2-wire networks of each of the above applications are slightly different and may yield different results. In particular, the network from some telephone sets (the 500 type set in the USA, for example) contains varistors which attenuate the sending signal about 4 dB and shunt some of the line current at higher line currents.

Thus it is possible that a headset may have more than one sensitivity rating depending on the application for which it is tested and the corresponding 4-wire to 2-wire network specified. The nomenclature defining the test specimen should thus include the network reference as well as the headset model reference.

6 Types of headset microphones used

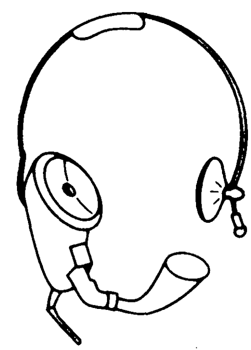
6.1 There are typically four types of microphones used in headsets:

a) carbon type [see Figure 3*a)*] – these are tested basically in the same manner as conventional handset carbon microphones regarding conditioning and in the same position as the conventional handset is normally used (see Recommendation P.75),

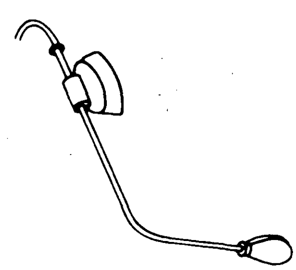
b) miniature microphone on boom [see Figure 3*b)*],

c) voice-tube, coupling sound to a remote transducer [see Figure 3*c)*],

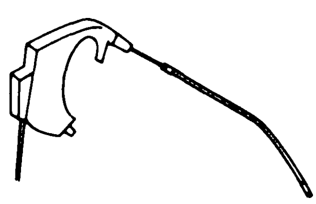
d) proximity, or pressure-gradient microphone (sometimes called "noise-cancelling") [see Figure 3*d)*].



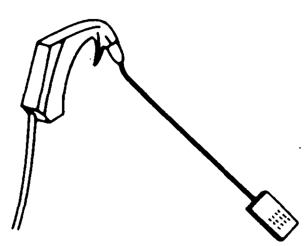
a) Carbon microphone



b) Miniature microphone on boom



c) Voice-tube, coupling to remote transducer



CCITT - 14081

d) Proximity microphone on boom

FIGURE 3

Microphone types used with typical headsets

6.2 All of the above microphone types except the proximity type are of the pressure type and will yield an output voltage proportional to the input sound pressure. The proximity type yields an output voltage proportional to the difference in sound pressure from the front to the back of the microphone. This gives the proximity microphone an output voltage inversely proportional to the *square* of the distance from the microphone to the sound source, and will aggravate any errors in sensitivity testing caused by positioning of the microphone. It is thus very important in subjective testing of headsets with proximity-type microphones to position them carefully to the manufacturer's or the Administration's instructions as to how the user should wear them. If the lip-ring prevents the microphone from being thus positioned, then a correction factor should be introduced in the data which accounts for the sensitivity difference caused by the positioning error.

7 Positioning the headset microphone to real and artificial mouths

7.1 Subjective loudness testing of sending systems involves using real voices as excitation signals for the sending microphone. The amplitude of the talker's voice is controlled by speech level meters so that the sound pressure is at the desired level at a lip-ring in front of the talker's mouth. The headset microphone or voice-tube should be positioned to this lip-ring, either by positional gauges or by simple measurement relative to the lip-ring following the positional dimensions established for the headset type (see § 4.1 above).

7.2 To obtain high values of signal-to-background noise with headsets, most microphones are specified to be close-talking types and thus the user is instructed to adjust the microphone close to his lips. A microphone close to and *directly in front of* the user's lips will produce "blasting" noises from the user's breath. Thus most headset manufacturers will also specify that the user position the microphone or voice-tube at the side of the mouth or below the lips, out of the "blasting" area. Because of the talker's lip-ring used in subjective testing and the need to sometimes position the microphone farther away from the talker's mouth than when in actual use, some "blasting" may still occur. If this happens, it is recommended that the microphone or voice-tube be adjusted vertically or horizontally away from the blasting area until the speech signal is clean.

7.3 Administrations have typically positioned headset microphones to artificial mouth lip-rings in the same way as real-mouth lip-rings. This method is satisfactory providing that the sound distribution around the artificial mouth is essentially the same as with real mouths.

Since there are generally no "blasting" effects associated with artificial mouths, and since the sound pressure versus distance on-axis from an artificial mouth is more consistent than at the sides of the artificial mouth, some Administrations place the headset microphone or voice-tube on-axis from the artificial mouth at a position of calibrated sound pressure level. This method is particularly recommended to avoid errors in testing headsets with proximity microphones.

8 Types of headset receiver systems

There are a variety of receiver coupling methods used with headsets:

a) conventional on-ear (supra-aural) receivers similar to conventional handsets receivers [see Figure 4*a)*],

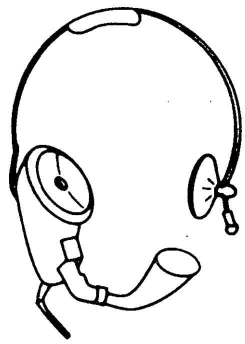
b) headsets with ear insert (eartip or earmould) types of receivers [see Figure 4*b)*],

c) headsets with couplers that mount in the concha portion of the ear, termed "intra-concha" couplers [see Figure 4*c)*].

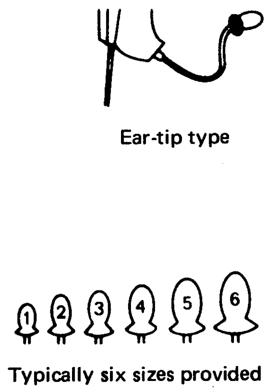
9 Coupling the headset receiver system to real and artificial ears

9.1 Headsets with conventional on-ear (supra-aural) receivers can be tested per conventional handset receiver test procedures and with the artificial ear for these types of receivers (see Recommendation P.51).

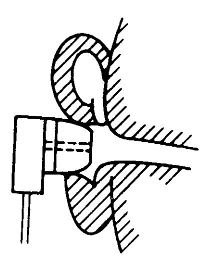
9.2 Headsets with ear insert types of receiver couplers have been tested in the past with simple 2 cm^3 cavity artificial ears. For future testing an ear simulator per IEC-711 is recommended. This is a multi-cavity simulator which is a better simulation of the impedance of a nominal human ear versus frequency. Attention is drawn to an increase of sound pressure of up to 8 dB (at 3 kHz) using the IEC-711 ear simulator compared to the simple 2 cm^3 cavity.



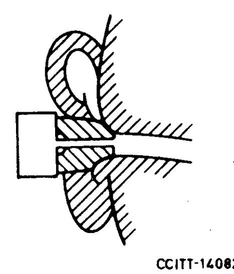
a) Conventional on-ear receiver system



b) Ear insert receiver system



d) "Intra-concha" receiver systems



Ear mould type is custom fitted to concha

d) Ear insert receiver system

FIGURE 4

Receiver types used with typical headsets

9.3 Headsets with intra-concha types of receiver couplers have also been tested in the past with simple 2 cm^3 cavity artificial ears. It is understood that the IEC-711 Recommendation is in the process of being extended to provide an adaptor to couple intra-concha types of couplers to the multi-cavity ear simulator.

10 Precautions in testing of headsets

10.1 Headsets with receive electronics require the transmit side to be powered to provide power to the receiver side electronics.

10.2 Some headsets have sending electronics (termed "threshold amplifiers") which reduce the sending sensitivity when the input sound pressure level is below a "threshold" value. It is important that input sound pressure levels be above this "threshold" value during all testing. Typical threshold sound pressure level values are in the range of -25 to -10 dB Pa. [2].

10.3 Some headsets have receive electronics which "compress" or linearly limit the higher drive signals to the receiver transducer. It is important that the receiver excitation levels be below the compression level when rating the sensitivity of the receiver system [2].

10.4 Receiver coupling systems that use sound pipes comprise an acoustic system in which the length and inside diameter of the sound pipe are important to the acoustic characteristics. Thus the ear coupling system of the headset should not be altered in testing.

ANNEX A

(to Supplement No. 9)

Definitions of terms

A.1 headset; operator telephone set

Telephone apparatus comprising essentially a "hands-free" handset. It is typically secured to the head of the wearer. It includes a telephone microphone and a telephone receiver. It is by definition a 4-wire device having a pair of wires for sending and a second pair for receiving. Also termed "head telephone sets", "speaking sets".

Note – When referring to a headset tested using a 4-wire to 2-wire network, such as tested by CCITT procedures, the term "headset" includes the network.

A.2 intra-concha receiver

A receiver intended to couple directly into the concha of the human ear and essentially filling the concha cavity but without entering the ear canal.

A.3 intra-concha (receiver to ear) coupler

A device which adapts the IEC-711 occluded ear simulator to provide simulation of an extension of the occluded ear canal to its entrance, together with the concha cavity. Its purpose is to facilitate objective performance measurements on intra-concha types of receivers.

A.4 threshold amplifier

An electronic system which will decrease the gain of a sending system (10 to 15 dB) when the user is not talking and the signal peaks are below a certain threshold. It is typically used to increase the signal-to-noise ratio of sending systems. Also termed "switch-gain" amplifier.

A.5 compression limiter

An electronic system which will linearly limit signals above a certain threshold by controlling (lowering) the gain of the system. It is typically used in telephone receiver systems to limit the loudness of high level signals.

A.6 proximity microphone

A transducer converting sound energy into electrical energy, responding to the difference in sound pressure between a front port and a rear port. Also termed "noise cancelling microphone", "pressure gradient microphone".

References

[1] CCITT Contribution COM XII-No. 19 (Determination of receiving sensitivities for some operators' headsets) United Kingdom Post Office, Study Period 1977-1980.

[2] CCITT Contribution COM XII-No. 156 (Testing considerations of operator telephone sets and telephone sets equipped with sending signal conditioners or receiving signal conditioners) Plantronics, Inc., USA, Study Period 1981-1984.

Bibliography

ARCHBOLD (R. C.): Proposals for the measurement of loudness ratings of operator's headsets, *Bruel and Kjaer Technical Review*, 1975-1.

CCITT – Question 3/XII, Contribution COM XII-No. 128 (Proposals for reply to Question 3/XII on loudness ratings of operator's telephone circuits) R.C. Archbold, Study Period 1973-1976.

CCITT – Questions 3/XII and 12/XII, Contribution COM XII-No. 172 (Evaluation of the IEC-711 occluded ear simulator for the objective testing of ear-insert receivers) Plantronics, Inc., USA, Study Period 1981-1984.

CCITT Contribution COM XII-No. 15 (Determination of modal position for some operator's headsets) U.K. Post Office, Study Period 1977-1980.

CCITT Contribution COM XII-No. 18 (Determination of sending sensitivities for some operators' headsets) U.K. Post Office, Study Period 1977-1980.

CCITT Contribution COM XII-No. 20 (Comparison of two loudness balance techniques for subjective measurements of receiving reference equivalents) USSR Telecommunication Administration, Study Period 1981-1984.

CCITT Temporary Document No. 4 (CCITT Laboratory Technical Report 676 System D_3, E_5, F_6, and G_1 – interaural balancing technique, usual balance and calculated objective loudness rating) CCITT Working Party Laboratory, CCITT Laboratory, Geneva, 24-26 January 1983.

CCITT Temporary Document No. 8 [CCITT Laboratory Technical Report No. 722, subjective and objective measurements of headsets (HA-1, HA-2)] CCITT Working Party Laboratory, Geneva, 6-9 April 1984.

CCITT Contribution COM XII-No. 198 (Frequency response curves for earphones, comparison of different subjective methods) Sweden, Study Period 1981-1984.

Supplement No. 10

CONSIDERATIONS RELATING TO CHARACTERISTICS FOR SUBSCRIBER SETS (HANDSET TELEPHONES)

(Malaga-Torremolinos, 1984)

1 Introduction

This Supplement based on reference [10] summarizes available information on how some characteristics for handset telephones can be optimized.

It contains information about sending and receiving sensitivities, frequency responses, sidetone characteristics, influence of impedance and handset dimensions. It must be remembered that there are different ways to make an optimization. For instance the number of degrees of freedom are essential. As there are different opinions in different countries (for instance, the different assumptions made) the results of the optimization will be different. This Supplement touches some of these aspects.

Desirable characteristics for digital telephone sets are still under study. However, what is mentioned about handset dimensions will also be applicable to digital sets.

2 Receiving frequency response

Most Administrations seem to prefer a fairly flat frequency response between 300 Hz and 3400 Hz. This probably derives from the early days of telephone networks, when it was determined that possible pre-emphasis at higher frequencies should be located at the sending end to obtain the best possible overall signal-to-noise performance. If we consider free-field, two-ear listening as a reference (face to face conversation) and assume a frequency-independent (flat) response, we should in principle simulate these conditions also at one-ear telephone listening.

Then, at the earphone listening, we should have a frequency response of the earphone as in Figure 1 to simulate the diffraction effect we have at free-field two-ear listening [1]. However, most Administrations seem to prefer a flat response and to put the corresponding correction at the sending end. It may also be easier to construct a receiver with high efficiency if the goal is a flat response. Reference [2] has suggested a response as in Figure 2 optimized for a mean local line. Where mains noise may cause problems, a response with greater loss at lower frequencies, e.g. at 200 Hz and lower frequencies, may be appropriate.

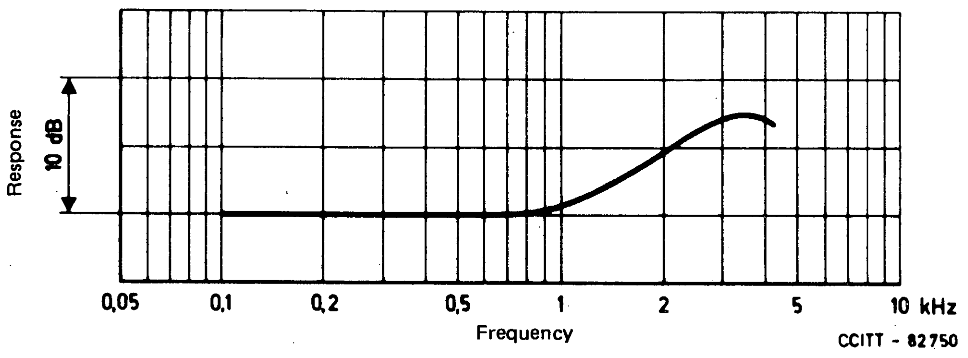


FIGURE 1

The diffraction effect around the head at 1 m distance in free field [1]

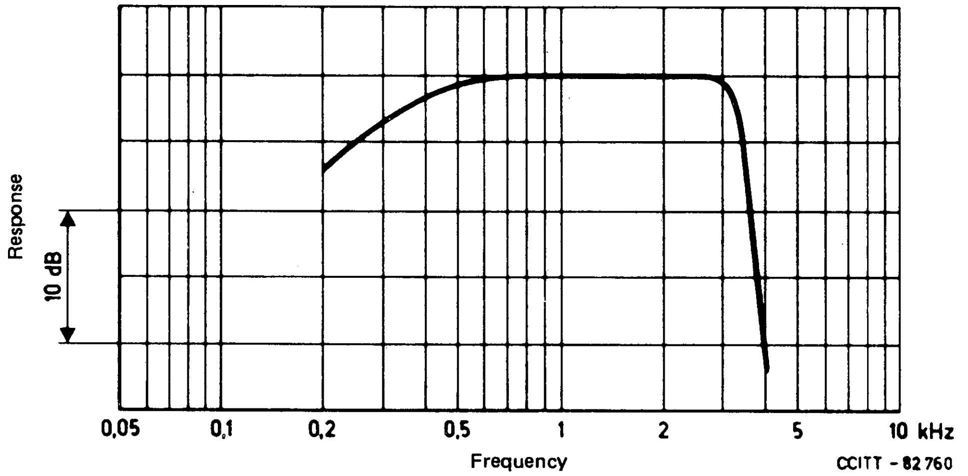


FIGURE 2

Receiving frequency response according to [2]

3 Receiving sensitivity

Receiving sensitivity today often is represented by values between CRE = 0 to CRE = −8 dB (RLR of −4 dB and −12 dB respectively).

A further increase of the sensitivity by the use of amplifiers might technically be possible. However, the probability for the audibility of crosstalk will increase with increased sensitivity. Therefore, the information gathered in Recommendation P.16 must be considered and it is doubtful if it can be recommended to increase the sensitivity further beyond CRE = −8 dB (RLR of −12 dB).

Increasing the receiving sensitivity also decreases the margins against the effects of speech-off noise on the connection, e.g. unwanted modulation products from PCM systems. The stability against singing will also be affected.

4 Sending frequency response

Having chosen the receiving response to be flat, the sending frequency response can be optimized to give the proper overall characteristic. Reference [3] suggests an optimization achieved by asking the listeners for the "preferred" response. The result is shown in Figure 3. Reference [4] suggests a 2 to 3 dB increase per octave with increasing frequency. This result was obtained in tests regarding "naturalness". Reference [2] suggests a steeper curve (Figure 4) as a result of an optimization where maximum loudness, minimum listening effort and lowest output level are combined. The degree of freedom used by [2] is of course less than in [3] and [4]. Here we may have a difference in opinion concerning which assumptions we must include in the optimization. If the signal-to-noise ratio is a problem, some decibels could be gained (without overloading) in the way shown by [2]. If there are no signal-to-noise ratio problems, an optimization for best naturalness as in [3] and [4] can be used. Thus, the result will depend on the assumptions.

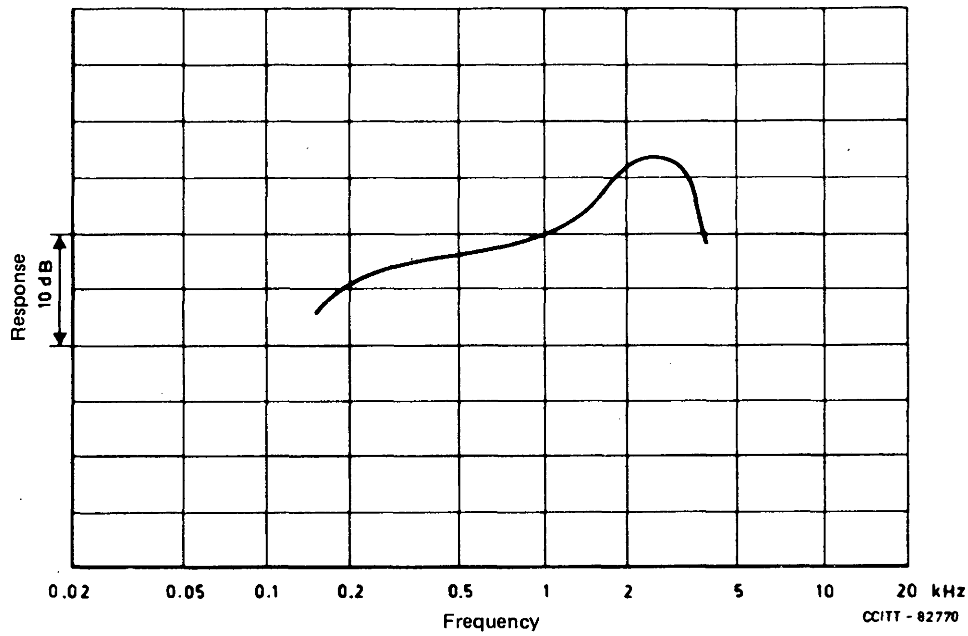


FIGURE 3

Sending frequency response according to [3]

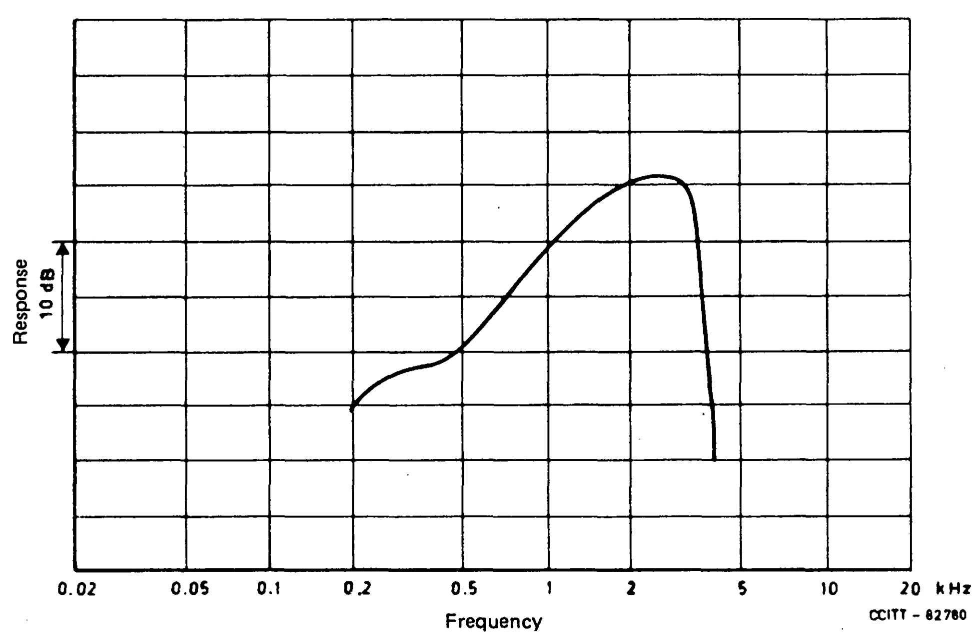


FIGURE 4

Sending frequency response according to [2]

Different opinions may also exist about the local cable length for which the frequency response should be optimized and if the high frequency loss at long lines should be compensated. Reference [2] suggests optimization of the mean local line which will be optimum to the highest number of subscribers (because of the statistical distribution of cable lengths).

The curves according to Figure 4 and [4] give with a flat receiving frequency response an overall characteristic close to what is obtained by the diffraction effect at free-field listening. However, this is probably not the whole explanation to the preferred curves. Even if the receiving responses were flat during sealed measuring conditions, hardly anyone keeps the earphone tight to the ear during conversation. Therefore, the actual responses during conversation probably give some additional low frequency cut-off that certainly has an influence on the results.

5 Sending sensitivity

When we want to choose the sending sensitivity we have one degree of freedom less that at the receiving end. We must consider both the probability of crosstalk and the probability of overloading other parts of the telephone system. Actual output levels from the telephone must be considered. As shown in [5] different output levels for the same SRE-value have been found in different countries. However, the different results show one important feature in common: output levels during normal conversation are generally lower than during reference equivalent measurements. Hopefully we will get better agreement on this point in the future if we use the measuring distance defined in Recommendation P.76, Annex A for loudness rating measurements.

Recommendation G.121, § 1 permits the corrected sending reference equivalent of a local telephone circuit on a limiting line to be about 11 dB higher than the corresponding corrected receiving reference equivalent (assuming no differential gain in the circuits of the national extension).

6 Regulation

A possibility to increase the sending sensitivity on long lines exists if we use sending regulation dependent on line length. The probability for overloading and the probability for far end crosstalk will not increase if the mean power is kept to the same value as today. See also [2]. The probability of near end crosstalk in the local cable will of course increase and has to be considered.

If regulation is introduced both at sending and receiving, more subscribers can have overall reference equivalents or Loudness ratings close to a preferred optimum, i.e. less calls will be rated poor and unsatisfactory. Another reason to introduce regulation is to obtain a better sidetone performance on short and long lines at the same time.

7 Handset dimensions

The shape and the dimension of the handset have an important influence on both send and receive levels. As stated in [2] "The earpiece must be capable of forming a good seal to the ear and the handgrip of the handset must be such that it will encourage the user to hold it to the head in the optimum position".

If we manage to get a design which gives a good seal during normal conversation it may influence the preferred frequency response (see § 4).

Reference [7] is an ergonomic study which suggests certain dimensions (see Figure 5) which will cater for 95% of the distribution of the relevant finger and head dimensions.

A later head dimension investigation carried out in the People's Republic of China [11] gives an ellipse of about the same size and the same distance to the ear as in Figure 5, but with the angle $\alpha = 21°$ instead of 17.7°. A further investigation, also from the People's Republic of China [12], shows that a handset with a mouthpiece tangential to the ellipse found in reference [11], although it only excludes 5% of lip positions, is only acceptable to 65% of the users.

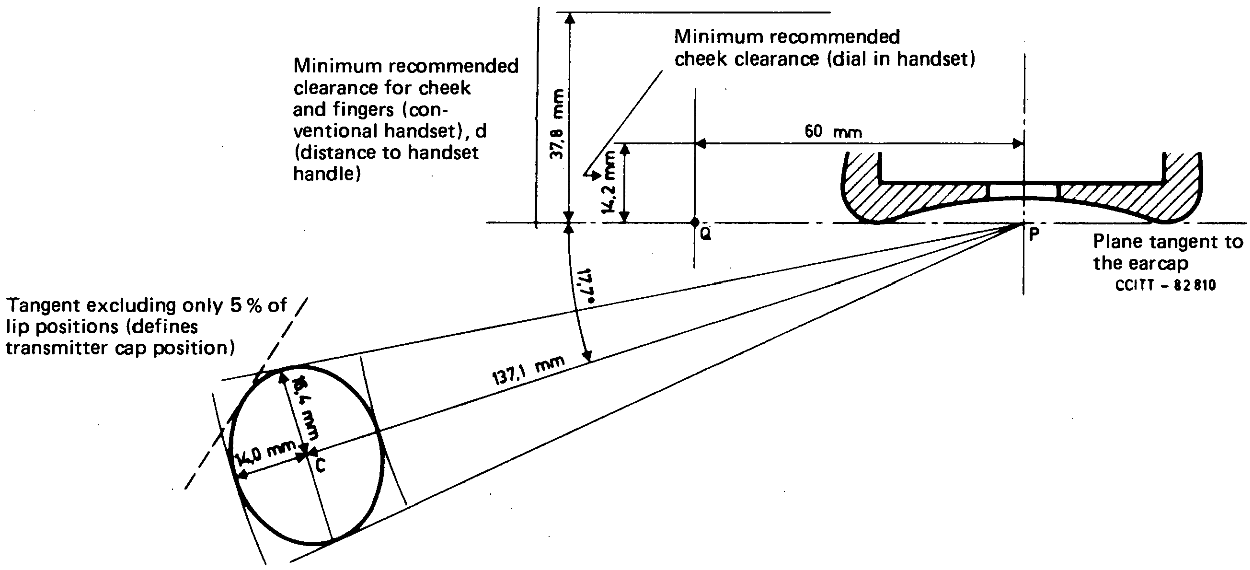


FIGURE 5

Telephone handset design criteria referred to in [7]

The investigation shows that for convenience the mouthpiece of the handset should be 8 mm or more outside the ellipse (the distance depending on the angle) to be acceptable to 95% of the people. However, when a longer mouth-microphone distance is chosen, the signal-to-noise ratio will be worse. Therefore both signal-to-noise ratio and distance for convenience must be considered and probably a compromise must be chosen. It seems interesting that a handset designed according to reference [12] is slightly shorter than one designed according to reference [7].

However, it is noteworthy that in respect of dimension d in Figure 5, few handset designs exhibit the value 37.8 mm. A value of 30 mm is more representative of modern designs (including the type used for the CCITT IRS). The weight balance of the handset also depends on this distance and for given weights of sending and receiving insets, is worse if $d = 37.8$ mm.

8 Impedance presented to the line

Some considerations concerning this topic are as follows:

- a conjugate match with the line maximizes the power transferred but creates sidetone problems on short lines and also stability/echo problems on long-distance calls;
- an image match to the line reduces the range of impedance presented to the exchange and eases the sidetone problem except for short subscriber-lines connected to resistive junction plant (e.g. PCM circuits);
- an impedance approximating the reference resistance (e.g. 600 ohms) eases standardization problems particularly in respect of alternative uses of the local line for non-speech services, but the optimum in respect of sidetone cannot be attained over the whole range of local line lengths.

References [2], [8] and [14] touch upon this subject.

9 Sidetone balance impedance

The degree of sidetone suppression is governed by the following parameters:

- microphone sensitivity;
- earphone sensitivity;
- sidetone balancing arrangement within the telephone instrument circuit;
- the impedance of the line to which the telephone is connected.

The microphone and earphone sensitivities and the instrument circuit are in part controlled by the required sending and receiving sensitivities. The impedance of the line to which the telephone is connected is not usually within the control of the telephone instrument designer. The only parameter freely available to the telephone designer to control the sidetone level is Z_{SO}, the sidetone balance impedance [8], [9], the impedance which when connected to the telephone completely suppresses sidetone [1)]. If a transformer hybrid is used in the telephone then the internal balance network impedance is equal to the sidetone-balance impedance Z_{SO} modified by the turns ratio of the transformer. However, the concept Z_{SO} is not affected if the circuit uses any other form of balancing arrangement instead of a transformer.

10 Interworking with the existing network

The design of new handset telephones to be introduced into the telephone network must take account of the need to give satisfactory transmission on connections to existing local telephone circuits either directly or via the long-distance network. Reference [8] contains information touching upon this aspect.

Reference [13] is an example of a specification used in North America. Guidance for desirable sending and receiving levels are given as well as characteristics to be minimally acceptable for connection to the public switched network. It should be noted that this specification uses IEEE terminology, which is different from that found in CCITT Recommendations.

References

[1] CCITT Recommendation *Description of the ARAEN*, Green Book, Vol. V, Rec. P.41, Fig. 4, ITU, Geneva, 1972.

[2] CCITT Contribution COM XII-No. 32 (U.K. Post Office), Study Period 1973-1976.

[3] CCITT Contribution COM XII-No. 22 (Australia), Study Period 1973-1976.

[4] GLEISS (N.): Sound transmission quality, Tele. No. 1, 1972, pp. 44-53.

[5] CCITT Recommendation *Subjective effects of direct crosstalk; thresholds of audibility and intelligibility*, Yellow Book, Vol. V, Rec. P.16, ITU, Geneva, 1981.

[6] Supplement No. 3 and Supplement No. 4 to Series P Recommendations, Yellow Book, Vol. V, ITU, Geneva, 1981.

[7] CCITT Contribution COM XII-No. 49 (ITT), Study Period 1973-1976.

[8] CCITT *Manual Transmission planning of switched telephone networks*, Chapter V, Annex 1, ITU, Geneva, 1976.

[9] RICHARDS (D. L.): Telecommunications by speech, Chapter 5, *Butterworths*, London, 1973.

[10] CCITT Contribution COM XII-No. 105 (LME), Study Period 1973-1976.

[11] CCITT Contribution COM XII-No. 21 (People's Republic of China), Study Period 1977-1980.

[12] CCITT Contribution COM XII-No. 122 (People's Republic of China), Study Period 1977-1980.

[13] EIA Specification RS 470.

[14] CCITT Contribution COM XII-No. 144 (British Telecom), Study Period 1981-1984.

[15] CCITT manual *Telephonometry* (to be published in 1985).

1) See also § 3.3.9 of the manual *Telephonometry* [15].

SOME EFFECTS OF SIDETONE ON THE SUBSCRIBER WHEN TALKING

(Malaga-Torremolinos, 1984)

(referred to in Recommendations P.11 and P.79)

Over a number of years sidetone has been studied in CCITT Study Group XII under Question 9/XII. Some important provisional conclusions have been reached from the point of view of the subscriber in his role as a talker. These provisional conclusions relate to the effect on the talker as regards the subjective effect of listening to his own voice, and the effect of sidetone on his talking level. These effects are summarized in the Figures 1 and 2.

Figure 1 shows that there is a preferred range for sidetone when the subscriber is talking under quiet conditions, and that the difference between the sidetone being objectionable or too quiet is of the order of 20 dB. (These results were obtained from talking-only tests and need to be confirmed by conversation tests.) The preferred range lies between 7 and 10 dB, STMR (sidetone masking rating – Recommendatin P.76) [1], [5]. It is not possible at this stage to suggest an acceptable range for the wide range of telephone connection and room noise conditions that can be encountered in practice. However, under good conditions of low room noise a "probable acceptable range" is indicated. In general, the higher the loss in a telephone connection the higher the value of STMR required to maintain good conversational conditions.

Figure 2 shows the way in which the talking level changes with sidetone level [1], [2], [3], [4]. These results were obtained by means of conversation tests [6], for a connection close to the preferred overall loss. The speech voltage will also be a function of room noise for the same connection conditions.

The effects of sidetone when the subscriber is listening in conditions of moderate to high level room noise is under study in Question 9/XII.

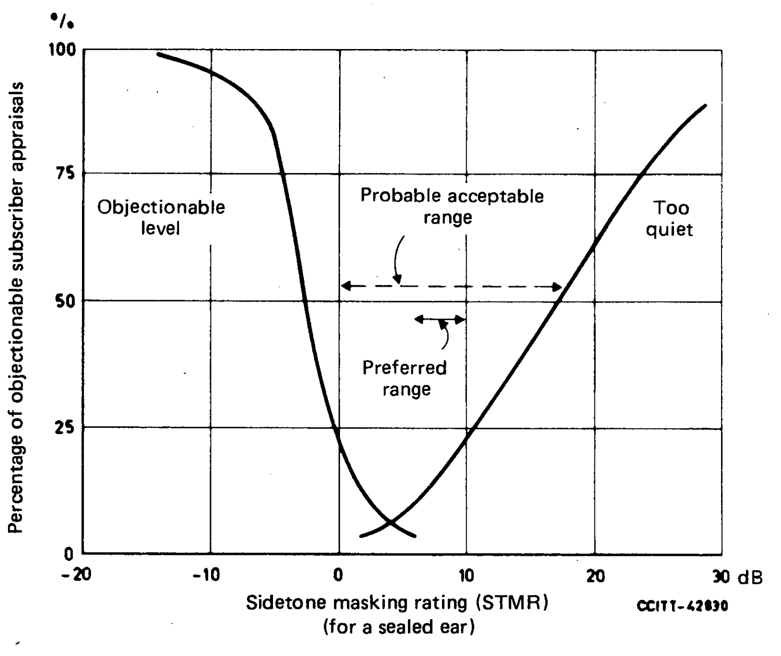


FIGURE 1

Curves showing sidetone levels that are objectionable and too quiet, together with the preferred range

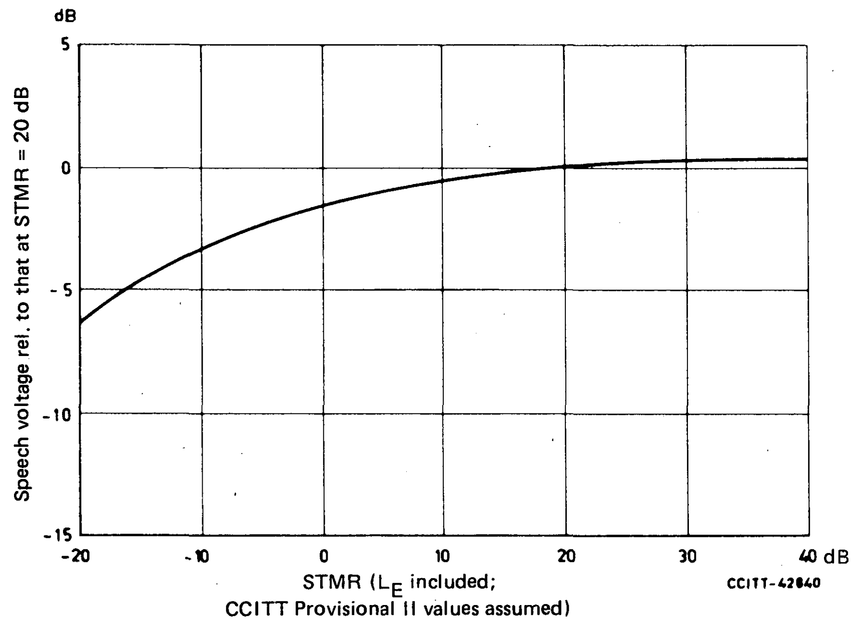


FIGURE 2

Speech voltage as a function of STMR

References

[1] CCITT Contribution COM XII-No. 50, Study Period 1977-1980.

[2] CCITT Contribution COM XII-No. 171, Study Period 1977-1980.

[3] CCITT Contribution COM XII-No. 199, Study Period 1977-1980.

[4] CCITT Contribution COM XII-No. 116, Study Period 1977-1980.

[5] CCITT Contribution COM XII-No. 152, Study Period 1981-1984.

[6] Results of conversation tests sent directly to Special Rapporteur for Question 9/XII, British Telecom, 1978.

Supplement No. 12

COMPOSITION AND ALIGNMENT OF THE NOSFER SYSTEM

(Malaga-Torremolinos, 1984)

(quoted in Recommendation P.42)

1 Introduction

This Supplement describes the composition and alignment procedures for the two NOSFER systems currently in use in the CCITT Laboratory, these are:

a) the original NOSFER based on the ARAEN and introduced around 1960;

b) the NOSFER-1984 which has been used experimentally since 1977.

Both systems are designed to fulfil the basic specification as detailed in Recommendation P.42. Similarly, the basic representation of a complete NOSFER as shown in Figure 1/P.42 is applicable to both models. The main working details related to composition and alignment for use are given or referred to below.

2 The original NOSFER system as based on ARAEN

The composition of this system is shown in Figure 4/P.42. It was based on the components of the ARAEN as given in Recommendation P.41, *Yellow Book* version. Modifications were made to the ARAEN as shown in the figure, in order to achieve sending and receiving sensitivity/frequency characteristics similar to those of SFERT [1] and thus ensure the validity of previously obtained reference equivalents made by comparisons with SFERT.

2.1 *Composition of the original NOSFER*

The original NOSFER was based on a linear wideband stable microphone (moving coil type 4021E) calibrated under free-field conditions so that the results were applicable to the real voice use of NOSFER at least in the frequency range 80-8000 Hz.

Details of the absolute calibration method previously used for the microphone are given in Volume V, CCITT *Red Book*, 1962, Annex 5. This includes information on the fundamental (Rayleigh Disk) method used for calibrating probe tube microphones as secondary standards and using these in a closed coupler device to check the calibration of the NOSFER microphone type 4021E. For this purpose, the relationship between the free-field calibration and the coupler test had been previously established using a number of microphones type 4021E.

Similarly, a stable wideband NOSFER receiver (type 4026A or type DR 701) with an attached rubber cushion of specified shape and properties has been used. The latter was provided to minimize acoustic leakage and noise ingress over a fairly wide frequency range but primarily to ensure a reasonable degree of agreement between real ear and artificial ear calibrations by giving a more repeatable seal between the human ear and the receiver.

Details of the NOSFER receiver are given in Annex D to Recommendation P.42 and the method of its calibration in Annex C of that Recommendation and in [2]. Apart from the transducer requirements detailed above, other components are as shown in Figure 4/P.42. Details of these may be obtained from Recommendation P.41 and former Recommendation P.42 (*Yellow Book* version). (Although Recommendation P.41 relates to the ARAEN it contains appreciable detail of the original NOSFER equipment which is common to both systems.)

2.2 *Alignment of the original NOSFER*

The alignment (normal adjustment) of the original NOSFER necessitated:

a) periodic tests of all frequency-dependent items, e.g. transducers and equalizers, together with adequate tests of fixed gain/loss items such as amplifiers and attenuators;

b) daily simple checks, prior to use on voice-ear tests, to ensure that the send, receive and overall system performance was being maintained.

Descriptions of these have been given in earlier CCITT Books with references as given in the former Recommendation P.42. However, for convenience, Table 1 has been assembled here for those elements of NOSFER which are frequency-dependent and re-expressed to align with current practice, i.e. using frequencies corresponding to ISO mid-frequencies of 1/3 octave bands and, where appropriate, with sensitivities related to the mouth and ear reference points so that the data may be applied to both types of NOSFER and to loudness calculations.

Microphone sensitivities (columns 1 and 2) are in terms of open circuit emf. For the free-field test the flat circular acoustic grid of the microphone lies in the same horizontal plane as the direction of the incident sound wave (as in normal use). Calibration methods as used previously are detailed in [2] and [3].

NOSFER send end sensitivity (column 6) which is used for loudness calculation is also in terms of free-field pressure, but referred to the mouth reference point at 25 mm in front of the equivalent lip position. The resulting voltage is measured across the input of the 600 ohm junction (J_s).

For the receiver (column 3) and NOSFER receive end (column 7) sound pressures are measured with the rubber cushion of the receiver seated on a flat plate of the ARAEN or Model IEC 318 artificial ear as shown in Annex C to Recommendation P.42.

In column 7 the voltage is applied across the 600 ohm input of the NOSFER receive end (J_r) and the sound pressure measured as in column 3.

Insertion losses (columns 4 and 5) are measured between 600 ohm terminations.

In the periodic check, a necessary stage of alignment is to ensure that the shapes of the nicrophone and receiver sensitivities lie within the masks formed by the sensitivity/frequency characteristics of columns 1 or 2, and column 3 respectively with the appropriate tolerances.

For daily alignment check of the complete NOSFER system, the practice is to use a 1000 Hz injected signal, as it is shown in a level diagram in Figure 7/P.42 (*Yellow Book* version).

TABLE 1

Sensitivities and insertion losses of the original NOSFER components

Frequency (Hz)	Sensitivity			Insertion loss NOSFER equalizer		Sensitivity NOSFER system	
	Microphone type 4021E		Receiver type 4026A	Send (dB)	Receive (dB)	Sending end	Receiving end
	Free field (dB V/Pa) 1	Closed coupler (dB V/Pa) 2	(dB Pa/V) 3	4	5	(dB V/Pa) 6	(dB Pa/V) 7
100	−65.70	−67.50	26.00	12.60	28.70	−6.60	1.30
125	−64.50	−66.00	26.40	12.50	28.30	−5.50	2.10
160	−63.30	−64.60	26.30	12.40	28.00	−5.00	2.40
200	−62.50	−63.70	26.50	12.30	27.30	−4.80	3.30
250	−62.00	−63.10	26.20	12.30	26.60	−5.20	4.00
315	−61.50	−62.70	25.60	12.00	25.30	−5.60	4.80
400	−61.70	−62.60	25.30	11.50	23.80	−5.40	5.70
500	−61.70	−62.60	24.70	11.00	21.40	−4.60	7.30
630	−61.30	−62.30	23.60	10.40	21.10	−3.40	6.70
800	−62.30	−62.70	22.50	9.60	20.20	−2.40	6.50
1000	−63.40	−63.10	21.30	8.70	19.00	−1.40	6.60
1250	−64.60	−63.50	20.50	8.00	18.20	−1.10	6.40
1600	−66.10	−64.50	20.30	7.40	17.80	−1.30	6.70
2000	−66.70	−65.30	21.30	6.80	18.00	−0.90	7.50
2500	−67.40	−65.60	23.30	6.50	19.80	−0.90	7.70
3150	−66.20	−65.40	26.80	6.30	24.90	0.50	6.10
4000	−65.90	−64.80	27.90	7.60	27.00	−0.10	4.10
5000	−65.40	−64.10	25.60	12.20	20.20	−4.20	7.40
6300	−64.60	−58.40	29.00	20.80	17.20	−11.80	12.50
8000	−65.60	−59.70	20.10	23.80	22.20	−15.90	2.00

Note 1 – Tolerances on columns 1, 2, 3, 6 and 7 are ± 1 dB (0.1-4.0 kHz), ± 1.5 dB (4.0-6.3 kHz) and ± 2.0 dB (6.3-8.0 kHz).

Note 2 – The closed coupler used for column 2 was specific to the 4021E microphone and used for day by day stability checks. It was not intended to produce sensitivities equal to the free-field response.

3 NOSFER-1984

In the currently used NOSFER-1984 (which has been in provisional use since 1977) a 1 inch condenser microphone (Bruel & Kjaer type 4144) with its associated preamplifier replaces the original moving coil microphone. As with the condenser microphone the sensitivity is independent of frequency over the range 80-8000 Hz. Consequent advantages arise from greatly reduced calibration processes which can now be readily reproduced in most laboratories. In addition there is now no need for a microphone equalizer.

A replacement receiver (IWATSU, Japan, type DR 701), has also been substituted for the original model (STC, UK, type 4026A). Both receivers, when fitted with the specified rubber cushion, have the same impedance and substantially similiar performance.

These transducers are associated with modernized integrated circuits and active equalizers, thus achieving an appreciable degree of miniaturization.

As with NOSFER based on ARAEN, variable elements are provided in both the send and receive amplifiers to give a certain degree of fine adjustment (± 2.5 dB).

The composition of the NOSFER-1984 remains the same as described in § 2.2.

The physical representation of the NOSFER-1984 is shown in Figure 5/P.42. Specification details and performance requirements of constituent parts are contained in working documents held by the CCITT Laboratory.

The dynamic range of NOSFER-1984 permits to accept up to +13 dB/Pa of sound pressure on the diaphragm of the NOSFER microphone, i.e. +25 dB/V read on the speech voltmeter (SV) without visible increase of harmonic distortion of the NOSFER system.

3.1 *Alignment of NOSFER-1984 sending end*

The NOSFER condenser microphone should be positioned in an anechoic chamber with the diaphragm in a horizontal plane and its centre at 140 mm from the plane of a vertical guard-ring of an artificial mouth. In this position the perpendicular axis of the circular guard-ring lies in the plane of the microphone diaphragm.

A standard microphone of sufficiently small dimensions (e.g. ½ inch or 1 inch manufactured by B&K), to avoid disturbance of the sound field in the range 100-8000 Hz, is first used to set up the sound pressure in the absence of the NOSFER microphone at the ISO 1/3 octave frequencies given in Table 1.

With the conditions for establishing the sound field known, the standard microphone is replaced by the NOSFER microphone. The voltage at the output of the NOSFER sending end, i.e. across the input of the 600 ohm junction, is then measured for each of the pure tone frequencies concerned.

The shape of the sensitivity/frequency characteristic so obtained, in terms of dBV/Pa across the input to the 600 ohm junction (J_s) with free-field sound pressure at the microphone, is then checked to see that it falls within the tolerance mask based on column 6, Table 1. If it is outside these tolerances, individual items in the sending end will need to be checked against their particular specification.

When within the tolerance for shape, a fine gain adjustment may still be required to adjust the sending sensitivity to its specified value.

The method for this is given in Annex B to Recommendation P.42. In principle it is based on the calculation of a send loudness rating (SLR) of the calibrated system against the specified system with sensitivities as in column 6, Table 1. Ideally this should yield SLR = 0; any small departure should be taken up by an appropriate change in the send amplifier gain to obtain an SLR within ± 0.2 dB of zero.

3.2 *Alignment of NOSFER-1984 receiving end*

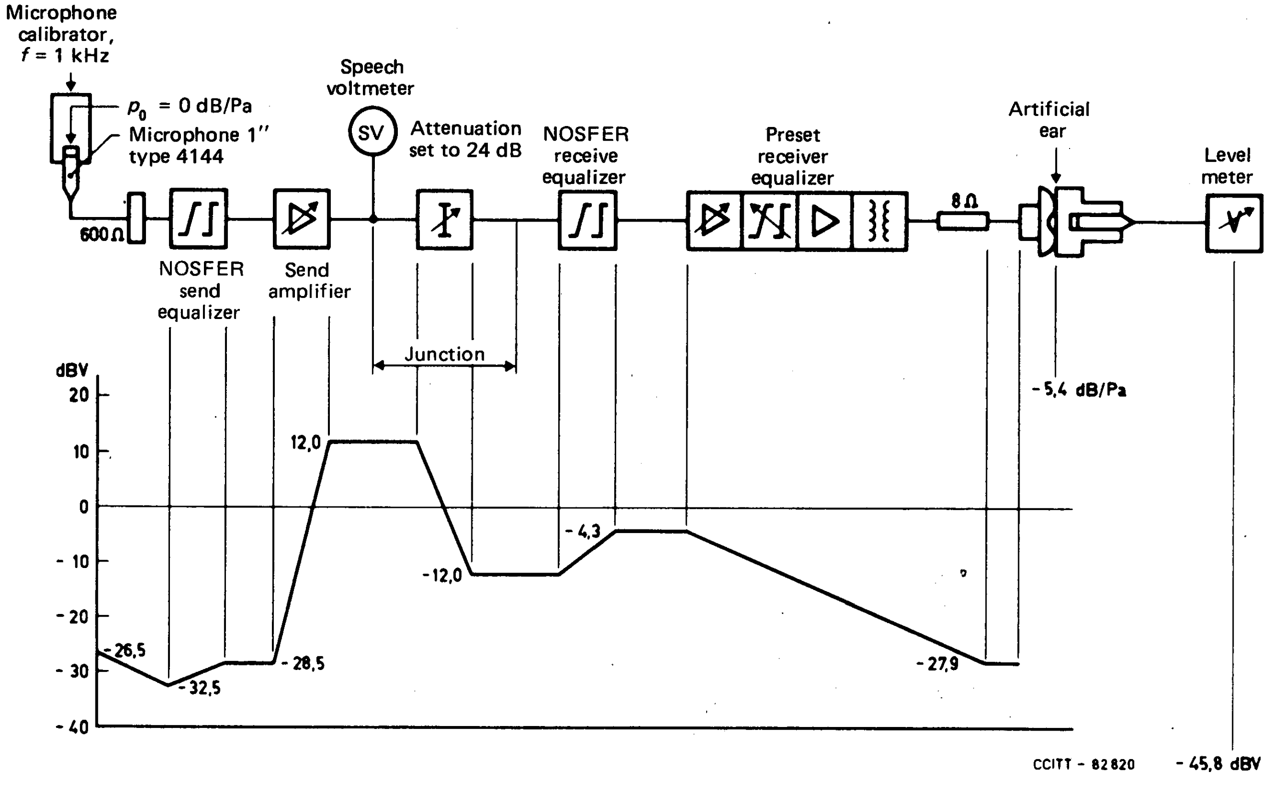
In this case the ISO 1/3 octave frequencies of Table 1 are injected into the input of the NOSFER receiving system from a 600 ohm source simulating the NOSFER junction and the corresponding sound pressures from the receiver are measured, so that the sensitivity may be expressed in dB relative to 1 Pa/V in the manner of column 7, Table 1.

For this purpose it is essential to mount the receiver complete with rubber cushion on the flat plate of an Model IEC 318 artificial ear in the manner specified in Annex C of Recommendation P.42. It is also advisable to include the normal receiver connecting cable as an integral part of the receiver, as the resistance of the cable is not generally negligible (about 2 ohms) against the receiver impedance (nominally $|Z|$ = 22 ohms).

The shape of the sensitivity/frequency characteristic so obtained must then be compared with the tolerance mask of the column 7, Table 1 response. If it falls outside the mask then individual circuit elements of the receiving end will need to be checked and corrected where necessary[1]. When within tolerance, the procedure is similar to that for the sending system in that a calculation of RLR is made in the manner of Annex C to Recommendation P.42. Any small departure from the result RLR = 0 is appropriately adjusted by a small adjustment of the receive amplifier gain to obtain an RLR within ± 0.2 dB of zero.

[1] As individual receiver (type 4026A or type DR 701) sensitivity/frequency characteristics may sometimes show departures from the nominal characteristic presented in column 3, Table 1, a receiver (active) (preset corrector) has been introduced in the NOSFER-1984 receive end (see Figure 5/P.42). This allows the alignment of the NOSFER-1984 receiving end to maintain its characteristic within the tolerance mask.

The daily alignment check of the complete NOSFER-1984 system is made using as a source a microphone calibrator producing 1000 Hz, and it is positioned directly on the NOSFER microphone. This yields a level diagram which is presented as shown in Figure 1.



Note – A 1000 Hz signal as the acoustic sound level 0 dB/Pa is supplied to the NOSFER 1 inch condenser microphone by using a microphone calibrator.

FIGURE 1

Level diagram for the NOSFER-1984 used in the CCITT Laboratory

References

[1] CCIF – *Measuring methods and apparatus*, Green Book, Vol. IV, Part 3, pp. 27-43, ITU, Geneva, 1956.

[2] *Absolute calibration of the ARAEN at the CCITT Laboratory*, White Book, Vol. V, Supplement No. 9, ITU, Geneva, 1969.

[3] Research Report No. 13200, U.K. Post Office, April 1950.

NOISE SPECTRA

(Malaga-Torremolinos, 1984)

(quoted in Recommendations P.44 and P.45 and Question 24/XII)

(Contribution from British Telecom)

1 Introduction

This Supplement gives the descriptions of noise spectra used in the evaluation of telephony transmission performance that are recommended by the CCITT or have been employed in studying questions assigned to Study Group XII.

Controlled environmental noise is used in subjective evaluations such as:

a) AEN determinations as described in Recommendations P.44 [1] and P.45 [2];

b) conversation and listening experiments as described, for example, in Supplement No. 2 [3].

Spectra for two different environments are described, one for room noise and two for internal vehicle noise.

2 Room noise

The room noise should have a power density spectrum corresponding to that published by Hoth [4]. Table 1 gives the spectrum density adjusted in level to produce a reading of 50 dBA on a sound level meter conforming to IEC Recommendation Publication 179 [5]. This is reproduced in Figure 1. This spectrum is independent of level, i.e. for 40 dBA the level in each band will be 10 dB less than that shown in Table 1. Additional information on the power in each 1/3rd octave band is also given in Table 1.

3 Internal vehicle noise

Two spectra representing internal vehicle noise [6], [7] have been recommended for use in the study of Question 24/XII [8] for evaluating mobile radio systems. They are adequately represented by simplified curves [9]; one spectrum for moving vehicles and the other for stationary vehicles. Table 2 gives the spectrum densities together with additional information on the power in each 1/3rd octave band. The spectrum density for moving vehicles is shown in Figure 2 *a)* and for stationary vehicles in Figure 2 *b)*. These spectra are independent of level.

Table 3 gives the computed values of the unweighted sound pressure levels for various speeds calculated over the ISO 1/3rd octave frequency bands centred on 63 Hz to 8000 Hz.

TABLE 1

Room noise spectrum

Frequency (Hz)	Spectrum density (dB SPL/Hz)	Bandwidth 10 $\log_{10} \Delta f$ (dB)	Total power in each 1/3rd octave band (dB SPL)	Tolerance (dB)
100	32.4	13.5	45.9	± 3
125	30.9	14.7	45.5	
160	29.1	15.7	44.9	
200	27.6	16.5	44.1	
250	26.0	17.6	43.6	
315	24.4	18.7	43.1	
400	22.7	19.7	42.3	
500	21.1	20.6	41.7	
630	19.5	21.7	41.2	
800	17.8	22.7	40.4	
1000	16.2	23.5	39.7	
1250	14.6	24.7	39.3	
1600	12.9	25.7	38.7	
2000	11.3	26.5	37.8	
2500	9.6	27.6	37.2	
3150	7.8	28.7	36.5	
4000	5.4	29.7	34.8	
5000	2.6	30.6	33.2	
6300	−1.3	31.7	30.4	
8000	−6.6	32.7	26.0	

Note 1 – The electrical input signal, e.g. white noise, shall be band-limited to the 1/3rd octave bands centred on the ISO preferred frequencies (ISO 266) between 100 Hz and 8000 Hz with the band edges conforming to the filters described in IEC 225.

Note 2 – The acoustical room noise is difficult to control at low frequencies, especially in the unspecified region below 100 Hz because of the dimensions of typical test cabinets, poor attenuation of such cabinets and the influence of extraneous noises, e.g. air-conditioning plant. It is therefore desirable to select a test cabinet that keeps these unwanted low frequency sound pressure levels to a minimum.

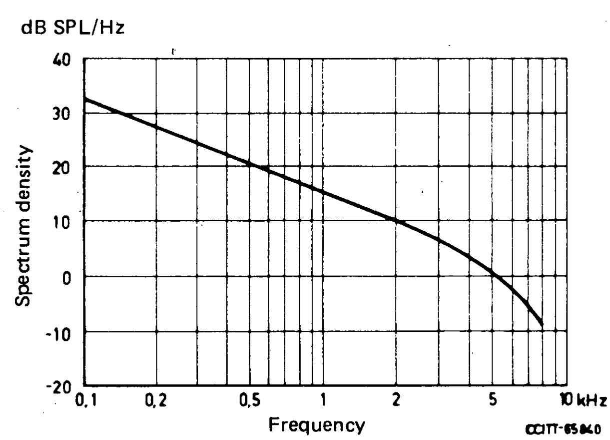
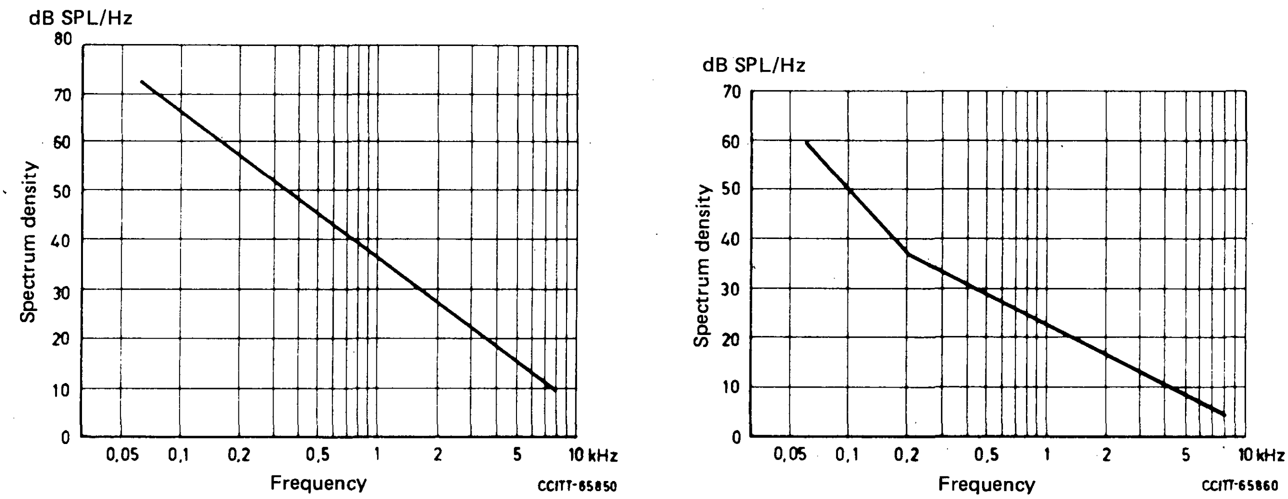


FIGURE 1

Room noise spectrum density



a) Spectrum density for moving vehicles

b) Spectrum density for stationary vehicles

FIGURE 2

TABLE 2

Internal vehicle noise spectra

Frequency (Hz)	Spectrum density (dB SPL/Hz)		Bandwidth $10 \log_{10} \Delta f$ (dB)	Total power in each ⅓ octave band (dB SPL)		Tolerance (dB)
	Moving	Stationary		Moving	Stationary	
63	72.3	58.3	11.7	84.0	70.0	
80	69.3	55.0	12.7	82.0	66.7	
100	66.5	49.8	13.5	80.0	63.3	
125	63.3	45.1	14.7	78.0	60.0	
160	60.3	42.0	15.7	76.0	56.7	
200	57.5	36.8	16.5	74.0	53.3	
250	54.4	34.7	17.6	72.0	52.3	
315	51.3	32.6	18.7	70.0	51.3	
400	48.3	30.6	19.7	68.0	50.3	
500	45.4	28.7	20.6	66.0	49.3	
630	42.3	26.6	21.7	64.0	48.3	
800	39.3	24.6	22.7	62.0	47.3	± 3
1000	36.5	22.8	23.5	60.0	46.3	
1250	33.3	20.6	24.7	58.0	45.3	
1600	30.3	18.6	25.7	56.0	44.3	
2000	27.5	16.8	26.5	54.0	43.3	
2500	24.4	14.7	27.6	52.0	42.3	
3150	21.3	12.6	28.7	50.0	41.3	
4000	18.3	10.6	29.7	48.0	40.3	
5000	15.4	8.7	30.6	46.0	39.3	
6300	12.3	6.6	31.7	44.0	38.3	
8000	9.3	4.6	32.7	42.0	37.3	

TABLE 3

Computed sound pressure levels of spectra

Spectra		Sound pressure level, unweighted (dB SPL)
Moving	30 km/hr	80
	80 km/hr	85
	110 km/hr	90
Stationary		75

Notes to Tables 2 and 3:

Note 1 – These values apply for typical vehicles. Discretion may be used to adjust the levels downwards for luxury vehicles and upwards for noisier vehicles.

Note 2 – Because of the practical difficulty of generating such high sound pressure levels at low frequencies, and because normal speech contains no apparent energy below about 63 Hz in which range of frequencies the ear is also comparatively insensitive it is probably advisable to restrict the recommended noise spectrum to frequencies above 63 Hz. However, it should be borne in mind that low and medium frequency vibrations have important physiological and psychological effects which should be studied in their own right.

Note 3 – The electrical input signal, e.g. white noise, shall be band-limited to the 1/3rd octave bands centred on the ISO preferred frequencies (ISO 266) between 63 Hz and 8000 Hz with the band edges conforming to the filters described in IEC 225.

Note 4 – The acoustical room noise is difficult to control at low frequencies especially in the unspecified region below 63 Hz because of the dimensions of typical test cabinets, poor attenuation of such cabinets and the influence of extraneous noises e.g. air-conditioning plant. It is therefore desirable to select a test cabinet that keeps these unwanted low frequency sound pressure levels to a minimum.

References

[1] CCITT Recommendation *Description and adjustment of the reference system for the determination of AEN (SRAEN)*, Yellow Book, Vol. V, Rec. P.44, ITU, Geneva, 1981.

[2] CCITT Recommendation *Measurement of the AEN value of a commercial telephone system (sending and receiving) by comparison with the SRAEN*, Yellow Book, Vol. V, Rec. P.45, ITU, Geneva, 1981.

[3] *Methods used for assessing telephony transmission performance*, Supplement No. 2, Yellow Book, Vol. V, ITU, Geneva, 1981.

[4] HOTH (D. F.): Room noise spectra at subscribers' telephone locations, *J.A.S.A.*, Vol. 12, pp. 499-504, April 1941.

[5] IEC Recommendation Publication 179, *Precision sound level meters*, 1965.

[6] CCITT Question 24/XII, Contribution COM XII-No. 120, (Noise inside light motor vehicles), Study Period 1981-1984.

[7] CCITT Question 24/XII, Contribution COM XII-No. 134, (Internal vehicle noise spectra), Study Period 1981-1984.

[8] CCITT Question 24/XII, Contribution COM XII-No. 1, (Link with mobile stations), Study Period 1981-1984.

[9] CCITT Contribution COM XII-No. 208, (Comparison of the results of vehicle noise submitted by France and BT), Study Period 1981-1984.

SUBJECTIVE PERFORMANCE ASSESSMENT OF DIGITAL PROCESSES USING THE MODULATED NOISE REFERENCE UNIT (MNRU)

(Malaga-Torremolinos, 1984)

(quoted in Recommendation P.70)

1 Introduction

The primary purpose of this Supplement is to define a specific subjective testing method for evaluating digital processes in a manner such that the quantization distortion effects of these processes on transmission performance can be taken into account in the evolving international telephone network. This implies both the ability to uniquely assign a numerical contribution to each digital process and the ability to use this assigned contribution in conjunction with other impairments to estimate telephone connection performance.

Secondary purposes of the Supplement are to suggest ways in which the subjective test results can be treated to arrive at the assigned impairment level of a particular digital process and how this assigned impairment level can be used in transmission performance analysis.

The defined test method, the listening-only Absolute Category Rating (ACR) test method described in Annex A to this Supplement, has been used successfully by a number of organizations as a basis for selecting a 32 kbit/s algorithm for standardization by the CCITT [1]. (The method should be used with caution when comparing within the same test systems with widely differing degradations, e.g. digital codecs, frequency division multiplex systems, vocoders).

2 Background for the ACR Test Method of Annex A

The method is based on a procedure utilized in an experiment conducted by a working group of the IEEE (Institute of Electrical and Electronics Engineers) in which representatives from seven countries participated (Canada, France, Italy, Japan, Norway, the United Kingdom and the United States) [2]. The aim of this experiment was to determine whether comparable results could be obtained when the same test is performed in several countries. Speech samples in the native languages of the participating countries were processed at a central location through 38 communications circuits. The recordings of the processed speech were returned to each country for evaluation on a five-point category rating scale by native listeners.

The communications circuits included 22 references (continuous random noise, MNRU, μ-255 PCM) and 16 adaptive differential PCM (ADPCM) systems. (The type of ADPCM system used was a first order fixed predictor [3].) An important part of the data analysis was estimation of the quality of the 16 ADPCM conditions at one location given measurements of ADPCM quality elsewhere.

Results (mean opinion scores [MOS]) obtained at the different locations differed [2]. Nonetheless, analysis of the results indicated that a reasonably accurate estimate of ADPCM quality in country B is the quality measured in country A adjusted by an additive constant.

Changes in the methodology were discussed at an IEEE working group meeting in May 1982 in Paris. The methodology incorporating these changes was recommended by Study Group XII in June 1982 as a basis for evaluating candidate 32 kbit/s algorithms for CCITT standardization. Subjective tests using the methodology were conducted under the auspices of Study Group XVIII in late 1982 with the results that a codec algorithm was selected and improvements (not related to telephone speech transmission issues) were identified. A second series of subjective tests in late 1983 confirmed that telephone speech transmission performance of the improved algorithm was suitable. (Differences between test results from the different participating organizations were also found in the 1982 and 1983 CCITT tests.)

The preceding discussion should not be taken to indicate that the subjective testing methodology is completely satisfactory: the reasons for differences found between countries [2] and [4] are thus far not explained. Nevertheless the testing methodology has the important feature of having been used by several countries.

3 Impairment reference scale for digital processes

Two reference scales that have been used for performance assessment of digital processes are continuous random noise and random noise with amplitude proportional to the instantaneous signal amplitude. Random noise with amplitude proportional to the instantaneous signal amplitude in terms of the Q ratio, according to the MNRU as specified in Recommendation P.70, should be used.

The reasons for this proposal are:

1) the signal processed through the MNRU is perceptually very similar in character to the signal processed through various digital processes, thus resulting, in principle, in easier assessment by test subjects, and

2) considerable experience and information has been accumulated with the MNRU.

Note – It has not been documented that Q represents a more suitable reference scale than continuous random noise.

4 Comparison method

A variety of methods are possible to characterize the performance of digital processes in terms of Q. These comprise listening-only tests involving:

1) opinion (category) ratings of test conditions generated using the MNRU and using digital processes being studied,

2) pair or multiple comparisons of MNRU and digital process conditions, and

3) articulation tests of MNRU and digital process conditions.

Listening-only tests involving ACRs (Absolute Category Ratings) using the methodology of Annex A to this Supplement are recommended. The methodology according to Annex A provides a basis for digital process performance evaluation.

A modification of the ACR test method, the Degradation Category Rating (DCR) test method, is described in Annex B to this Supplement. Based on results from one Administration, the DCR test method provides greater discrimination between conditions than does the ACR test method [5].

Future experience may result in changes in the methodology of Annex A. Furthermore, there is the need for supplementing listening-only tests with conversation tests. These considerations need to be kept in mind in future studies.

5 Analysis of test results

The purpose in conducting test of digital processes is to determine their suitability for use in telephone networks. A procedure which has been used is to assign Q values, determined using the reference system of Recommendation P.70, to processes of interest. Various methods of data analysis are possible, but it appears desirable to define a single method to be used in order to assure expressing results in common terms. The proposed provisional method is based on the use of MOS (mean opinion score) values obtained using the procedures of Annex A.

Hypothetical results obtained from a subjective test conducted according to the methodology of Annex A to this Supplement are shown in Figures 1 to 4. (Straight lines are used simply to connect data points.) Generally such results will display a saturation effect at and near the very good conditions (high MOS) and the very bad conditions (low MOS). (For high MOS, the saturation is caused by the 5-point scale and possibly by the idle circuit noise of the subjective test system without added impairments, e.g. idle circuit noise, Q, codec quantization distortion. For low MOS, the saturation is caused by the 5-point scale.) Experience [2] has shown that due to this saturation effect, acceptable accuracy for the determination of Q is obtained for the range of about 5 dB to 25 dB.

Analysis of test results should include statistical analysis to establish that MOS values obtained are due to the test conditions and not to other factors. Student's test may be suitable, but there is some indication that analysis of variance is more appropriate.

The principles of a method of analysis used by one organization are outlined in Annex C of this Supplement. The method uses analytic values, called fit means, calculated from subjective test results; these analytic values are similar to MOS values calculated from test results.

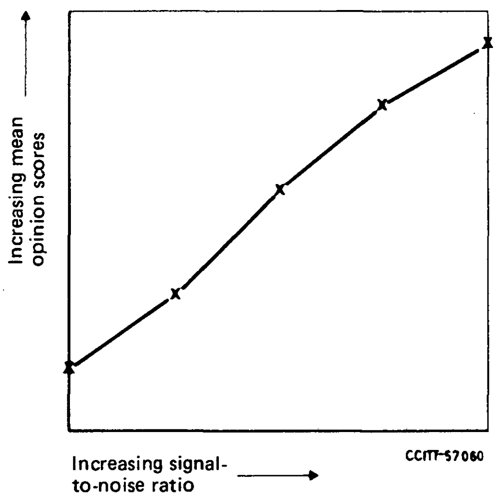


FIGURE 1

Mean opinion scores for SNR; speech power in dBm, idle circuit noise power in dBmp

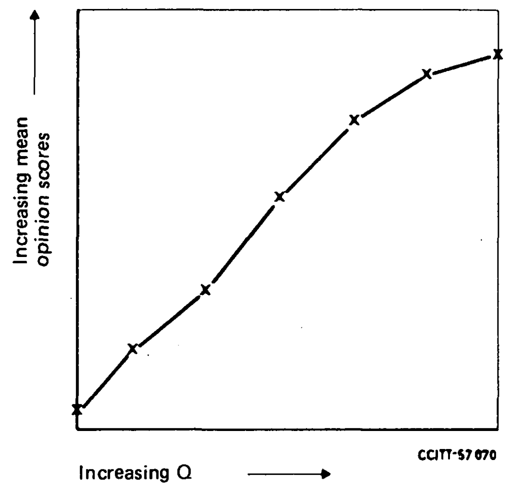


FIGURE 2

Mean opinion scores for Q; Q is the setting of the MNRU

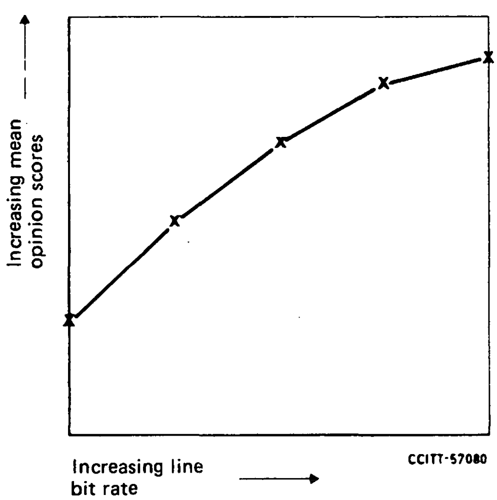


FIGURE 3

Mean opinion scores for a single codec at different line bit rates

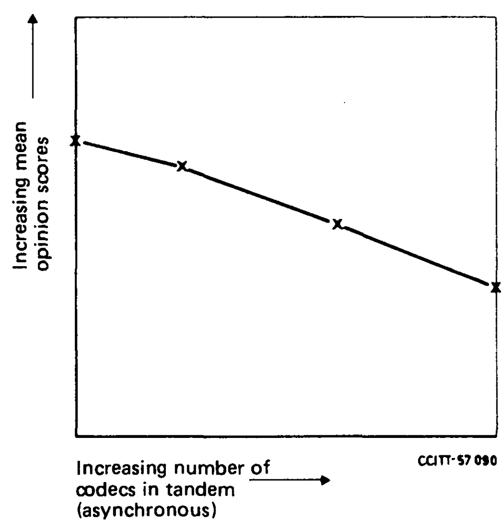


FIGURE 4

Mean opinion scores for tandem codecs

ANNEX A

(to Supplement No. 14)

Absolute category rating (ACR)
method for subjective testing of digital processes

A.1 *Introduction*

The listening-only test method consists in principle of three parts: preparation of source tapes; processing of the sources tapes to obtain stimulus tapes containing the test conditions of interest; conduct of subjective tests using the stimulus tapes. Certain steps may be combined if interchange of source/stimulus tapes between locations is not involved.

The methodology is based on the notion of simulating a connection comprising a sending system, a receiving system and an interconnection system which provides for inserting the impairment of interest (idle channel noise and quantization distortion from the MNRU and from digital processes).

Listener responses in the subjective tests are influenced by a number of sources of variation, e.g. speech material, talker voice characteristics, presentation orders, time effects, etc. Unless controlled in some way, these variables may bias the outcome of the experiment. It is therefore recommended that appropriate experimental design be applied to take this into account. Principles for experimental design may be found in textbooks on statistics.

A.2 *Preparation of source tape(s)*

The recording system consists of a tape recorder, means for injecting calibration tones and a suitably defined sending system.

A.2.1 *Tape recorder*

The tape recorder should be a high (studio) quality two-track machine. The type of equalization should be stated, but IEC is preferred. One of the tracks is used for recording the speech samples; the second channel is available for other purposes, e.g. cueuing tones to allow computer start/stop control of the tape recorder. The tape recorder should be operated at 19 cm/sec.

Low print-through, low-noise tape should be used and the tape should be stored "tail-out" so that it is necessary to rewind the tape before it is played.

Note – The use of an A/D converter and a television cassette recorder should be considered as a means for recording and storing high quality source and test tapes.

A.2.2 *Calibration tones*

It is recommended that calibration tones be recorded on the source tape(s) to enable checking the sensitivity/frequency characteristics of the connection simulation from input to the source tape recorder to output from the stimulus presentation tape recorder. Tones should be recorded in sequence at 250, 500, 1000, 2000, 3000 and 4000 Hz of 5 seconds duration each, with a level 6 dB below the maximum r.m.s. input level of the tape recorder. These tones should be followed by a 15 second recording of a 1 kHz test tone at maximum r.m.s. input level to enable calibration of the interconnection and listening systems. This should be followed by several metres of leader tape.

A.2.3 *Sending system*

The sensitivity/frequency characteristics of sending systems of different countries are likely to differ and, thus, results of different countries may differ because of attenuation distortion. Furthermore, the performance of complex digital codec algorithms may be dependent on the shape of the sending system sensitivity/frequency characteristics. Therefore, it is desirable that at least for some of the conditions in a test the sending system characteristic be as given in Table A-1 (simulates the IRS send part without filter).

TABLE A-1

IRS characteristics before adding SRAEN filter

Frequency (Hz)	S_{MJ} (dB V/Pa)	S_{Je} (dB Pa/V)
100	−22.00	−21.00
125	−18.00	−17.00
160	−14.00	−13.00
200	−10.00	− 9.00
250	− 6.80	− 5.70
315	− 4.60	− 2.90
400	− 3.30	− 1.30
500	− 2.60	− 0.60
630	− 2.20	− 0.10
800	− 1.20	+ 0.00
1000	+ 0.00	+ 0.00
1250	+ 1.20	+ 0.20
1600	+ 2.80	+ 0.40
2000	+ 3.20	+ 0.40
2500	+ 4.00	− 0.30
3150	+ 4.30	− 0.50
4000	+ 0.00	−11.00
5000	− 6.00	−23.00
6300	−12.00	−35.00
8000	−18.00	−53.00

It may be desirable to include conditions for which the sending system represents a typical (average) local system according to the testing organization's (country's) network and/or needs. This system comprises a handset telephone set, a simulated physical cable pair, a feeding bridge and a resistive termination (e.g. 600 ohms, 900 ohms) to which the source tape recorder is connected. The telephone set can utilize a linear telephone microphone with a real voice sensitivity/frequency characteristic such that the acoustic-to-electric response of the sending system represents the organization's average local system. It may also be desirable to include conditions obtained with a carbon telephone microphone representative of the type(s) used in the organization's (country's) network. (See Recommendation P.64.) The characteristics (and feeding current) should be reported. It may also be desirable to report the characteristic measured using an artificial sound source. (See Recommendations P.51 and P.64.)

A.2.4 *Recording environment*

The recording environment should be that of a quiet living room or office. The ambient room noise level should be 25-30 dB(A). The noise spectrum should, if possible, have the shape of the Hoth spectrum of Supplement No. 13. Special tests may be required using other noise levels and/or spectral characteristics (e.g. typewriter noise, etc.).

The room noise characteristic should be reported in as complete a form as is possible [e.g. dB(A), long-term spectrum, amplitude/time distribution, etc.].

A 30 second recording of the room noise through the local system should follow the calibration tones. This should be accomplished with a talker holding the telphone handset in a normal use manner. (Special precautions may be necessary in order to avoid breath sounds if desired.)

A.2.5 *Speech samples*

A source tape is made of 4 × C samples (4 talkers, samples consisting of training, reference and test conditions). Each sample should comprise 2 or 3 sentences separated by at least 1 second.

All samples should be different to avoid repetition of sentences during a test. When reporting test results, it may be desirable to provide a list of the sentences used (i.e. 8 × C or 12 × C sentences).

Each sample is expected to be 6-10 seconds in length. The samples should be separated by 5 seconds of silence to allow for control (e.g. turning the tape recorder on and off) and of the amount of time needed for subjects to vote.

The r.m.s. level of the speech samples (speech power while active) should be 12 dB below the r.m.s. level of the 1 kHz calibration tone in order to avoid peak clipping of the speech samples by the tape recorder and to measure in an easy way the actual r.m.s. level of the speech.

A.2.6 *Talkers*

At least 4 different talkers (2 female, 2 male) with different voice characteristics should be used. Selection of the talkers will depend on the judgement of the experimenter.

A.3 *Preparation of stimulus tape(s)*

The interconnection system will consist of the source tape recorder (resistive, 600 or 900 ohms), an input filter, a means for inserting test conditions, an output filter, and the stimulus tape recorder (resistive, 600 or 900 ohms). The characteristics of the filters should be provided.

A.3.1 *Test conditions*

The test conditions comprise the digital codec(s) of interest. The codec(s) should be defined as simply and completely as possible (e.g. A-law/μ-law, ADPCM with first order fixed predictor, etc.). This is to enable unique performance specification for codecs of the same type.

Because codecs may have different performances at different speech input levels, they should preferably be tested not only at a nominal fully-loaded condition, but also at levels below and above this level, say ± 10 dB. These changes in input level to codecs should be "off-set" by corresponding adjustments of their outputs to maintain an approximately constant output level for the test. (Listening level may also affect relative performance of different digital processes. See also § A.4.4.)

The codec(s) should be tested singly (one encoding/decoding pair) and with 2.4 and (possibly) 8 codecs connected in tandem asynchronously. (It may also be desirable to include conditions in tandem synchronously.) The codecs may be hardware or software implemented; if the latter, injected circuit noise expected for practical codecs should be included.

For the single codec(s) conditions, the line bit rate should be the design value and, if possible, line bit rates both above (to ensure subjective saturation) and below (to ensure degraded performance). These conditions may be useful in assigning a performance level(s) to the codec(s). (For example, a nominal 32 kbit/s ADPCM algorithm might also be tested at 16, 24, 40 and 48 kbit/s.)

The tandem conditions should utilize the codec(s) at the design line bit rate(s).

Codec conditions with line errors should be included. Bit error rates covering the range 10^{-3} to 10^{-6} should be used.

A.3.2 *Reference conditions*

Reference conditions which should be included are Q values within the range 5 dB to 25 dB with a minimum of 4 steps. (It may also be desirable to include Q values of 0 dB and 30 dB.)

It is desirable to include injected circuit noise values to provide SNRs within the range 5 dB to 45 dB with a minimum of 4 steps. (SNR is the dB ratio of speech power in milliwatts while active to injected circuit noise in milliwatts; the circuit noise conditions should be band-limited by filters having the same characteristics as the filter of the MNRU.) Note that the 45 dB ratio could be dependent on the inherent system noise, e.g. noice from the source tape preparation process, noise from the source tape recorder, etc.

Source conditions should also be included. (These are obtained by removing the injected idle circuit noise.)

The purpose of including the injected circuit noise conditions is to enable the relating of test results to results available on the effects of loss and circuit noise (Question 4/XII) and to allow use of the test results in subjective opinion model studies (Question 7/XII).

Other reference conditions can be included at the discretion of the testing organization. For example, particular organizations may have available information from previous tests of A-law/μ-law companded PCM, and it may be desirable to include some PCM conditions to allow comparison with previous results.

A.3.3 *Calibration*

The insertion loss of the interconnection circuit should be 0 dB at 1 kHz between the resistive source/termination. This should apply for the better conditions e.g. Q = 25 dB, SNR = 45 dB and the test codec(s) operated at design line bit rate(s).

The r.m.s. level of the 1 kHz calibration tone at the input to the inter-connection circuit should be 3 dB below the codec(s) overload level (which should be quoted). This will ensure that r.m.s. level of the speech samples will be 15 dB below the codec(s) r.m.s. sinewave overload level.

With the above calibration, the injected circuit noise levels in dBm across the output resistive termination should be adjusted to an appropriate level relative to the output 1 kHz calibration tone level in dBm. Note that in particular the circuit noise impairment should be present during the speech sample idle periods but not before and after the speech sample.

The stimulus tape recorder calibration should be the same as that for the source tape recorder.

A.3.4 *Stimulus tape(s)*

Stimulus tapes should begin with the 1 kHz calibration tone recorded (without introduced impairments), 12 practice conditions and then the test and reference conditions.

The practice conditions should be selected to introduce the test subjects to the test format and range of speech quality. These conditions should consist of each of the four talkers with 3 practice conditions.

The basic test and reference conditions will be 4 (i.e. number of talkers) times the number of nominal conditions. These conditions should appear in random order. There should be at least 2 stimulus tapes with different random orders. (These could be used in different tests with different subject groups.)

It may also be desirable to include replication of at least some of the test/reference conditions. However, this may not be possible for a practical subjective test size.

The timing of conditions in the stimulus tapes is the same as that for the source tapes, e.g. approximately 6-10 seconds (2 or 3 sentences) with each condition separated by 5 seconds of silence.

The calibration tones on the source tape need not appear on the stimulus tape (except for 1 kHz calibration tone as noted above). However, the calibration tone levels should be measured at the interconnection system output resistive termination so that the system sensitivity/frequency characteristics can be measured and reported for all condition types.

A.4 *Testing procedure*

A.4.1 *Listeners*

The preferred number of listeners is 32, assigned equally to each tape. At least 12 test subjects should be used. It is desirable that the subjects be selected to represent the typical customer population (e.g. half of the group females and half males, ages approximating the population distribution of ages, normal hearing, etc.).

A.4.2 *Listening system*

For reasons given in the first paragraph of § A.2.3, the receiving system characteristic should be as given in Table A-1 (simulates the IRS receive part without filter).

It would be desirable if the listening system simulated the organization's typical (e.g. average) local system representing the central office source impedance, feeding bridge, physical cable pair and the handset telephone set. The electric-to-acoustic sensitivity/frequency characteristic of the listening system should be determined (see Recommendation P.64). Sidetone in the listening system should be suppressed.

A.4.3 *Listening environment*

The listening handset(s) should be located in a room with an ambient room noise level $\leqslant$ 40 dB(A), preferably 25-30 dB(A) (simulating a quiet office or living room). The noise spectrum should, if possible, have the shape of the Hoth spectrum of Supplement No. 13. The actual ambient room noise level and spectrum, if different from the above, should be reported.

A.4.4 *Speech level*

The 1 kHz calibration tone on the stimulus tape when played through the listening system should be adjusted such that reproduction occurs at a level of -3 dB Pa as measured with the artificial ear recommended by the CCITT. (See Recommendation P.51.) This will result in a speech level of about -15 dB Pa which is close to the preferred level. It may also be desirable to include conditions with a 10 dB lower level and 10 dB higher level since the listening level may affect the relative performance of different digital processes. [1)]

A.4.5 *Test instructions*

Test subjects will be provided with a written set of instructions which will also be read to them (either by the test administrator or by means of a tape recording). The instructions should be given before the practice conditions. Subjects should not be instructed that the practice conditions represent the full range of quality to be encountered in the test. After the practice conditions, there should be sufficient time allowed for answering possible questions by the subjects.

The subjects should be asked whether the speech they have just heard comes from a connection that was (in English) "Excellent, Good, Fair, Poor and Bad (or very Poor)". In countries for which English is not the native language, the appropriate terms in the native language should be used.

A.4.6 *Data collection*

Subjects' responses can be recorded by computer, on paper or by such other means as are appropriate. If paper and pencil are used, the response to each condition should be recorded on a separate card so that the subject is not looking at a previous opinion while making a new judgement.

A.5 *Results reports*

Reporting all of the raw data may be desirable but results in excessive documentation. Therefore, it may be appropriate to combine data across talkers and report the number of ratings in each of the 5 categories for each condition type, e.g. Q = 15 dB, SNR = 25 dB, etc. (Conclusions resulting from an analysis of the study of possible talker effects should be reported.) In addition, mean opinion scores (MOSs), standard deviation, 95 percent confidence intervals and other statistics computed by the organizations in analyzing the data should be reported.

Other items which should be reported are as follows:

a) microphone type;

b) sensitivity/frequency characteristic of the sending system (Recommendation P.64);

c) description of recording room and ambient noise levels;

---

1) There is some indication that a speech level of -5 dB Pa (1 kHz tone level of $+7$ dB Pa) would be more suitable than -15 dB Pa for discrimination between coder conditions.

d) measurement and adjustment procedure for speech levels;

e) sensitivity/frequency characteristics of the interconnection system for all test/reference condition types;

f) sensitivity/frequency characteristic of the listening system (Recommendation P.64);

g) description of the listening room and ambient noise level;

h) method of recording test subject opinions;

i) description of subject group including age, sex, population, prior experience and, if possible, audiometric threshold;

j) handset dimensions.

Bibliography for Annex A

KIRK (R. E.): Experimental design procedures for the behavioral sciences, *Brooks/Cole Publishing Company*, Belmont California, 1968.

CCITT Recommendation P.64.

CCITT Recommendation P.74.

ANNEX B

(to Supplement No. 14)

Subjective perfomance assessment of digital encoders using the degradation category rating procedure (DCR)

(Contribution of the French Administration)

B.1 *Introduction*

A listening-only test method has been drafted by CCITT SG XII to assess the subjective quality of digital encoders (see Annex A). This procedure, Absolute Category Rating test (ACR), leads to a low sensitivity in distinguishing among good telephone quality coders (within the range of quality of 6-8 bit PCM coders). If higher sensitivity is needed we propose to use a modified version of that procedure, which can be defined as a Degradation Category Rating test (DCR). For image testing CCIR [6] recommends two alternative methods, absolute category ratings and degradation category ratings. The DCR procedure, which in particular uses an annoyance scale and a high quality reference before each judgement, seems to be suitable for evaluating good quality images. Therefore this method has been adapted to evaluate speech quality.

This Supplement first describes the adaptation of the DCR procedure to speech. Then the sensitivity of the method is compared with that of the ACR procedure on the same circuits. Only the differences between ACR and DCR procedure are presented here. One can refer to Annex A for common points which are not covered in this Annex.

B.2 *Degradation category rating procedure (DCR)*

B.2.1 *Speech samples*

Each configuration is evaluated by means of judgements upon four talkers reading two different samples. Each sample should comprise two sentences separated by at least one second. These two samples (S1, S2), hence four different sentences, should be selected from a wider corpus composed of phonetically balanced sentences so that the mean score obtained in evaluating MNRU circuits for these four sentences is about the same as that obtained for the wider corpus. Therefore the corpus consists of eight samples defined as follows:

talker T1 reading samples S1, S2

talker T2 reading samples S1, S2

talker T3 reading samples S1, S2

talker T4 reading samples S1, S2.

This results in a repetition of the two samples during the test. But we feel that this is not so critical for the procedure where a degradation is evaluated with regard to a reference, especially for good telephone quality where the intelligibility of speech is nearly perfect. The use of different samples for each configuration as is done in ACR experiments could be one of the reasons for this procedure's lack of sensitivity.

B.2.2 *Reference conditions*

Reference conditions should include multiplicative noise with Q values within the range of 10 to 30 dB with a minimum of four steps. (It may also be desirable to include Q values of 5 dB and 35 dB).

A high quality reference should be chosen to be inserted before each judgement. Usually source conditions are used, i.e. samples with no more degradation than those introduced by sending systems and limitations of frequency bandwidth. Four "null pairs" (A-A) are included to check the quality of anchoring of the listeners' judgements.

B.2.3 *Stimulus presentation*

The stimuli are presented to listeners by pairs (A-B) or repeated pairs (A-B-A-B) where A is the high quality reference sample and B the same sample processed by a codec. The purpose of the reference sample is to anchor each judgement of the listeners. Using a reference and subjective judgements with respect to that reference is quite a common procedure in psychoacoustics. It tends to result in a good sensitivity for the overall evaluation by listeners. Samples A and B should be separated by 0.75 s and in a repeated pair procedure (A-B-A-B) the separation between the two pairs should be 2 s.

It seems that the classical order effect observed in a one-sample listening test (ACR for example) is not observed with the DCR procedure. Thus, only one random order of presentation can be used. Therefore the basic test and reference conditions will be eight times (four talkers × two samples) the number of nominal conditions.

The timing for the response of listeners is the same as for the ACR test, i.e. 5 s between each presentation (pair or repeated pairs).

B.2.4 *Test instructions*

The subjects should be instructed to rate the conditions according to the five point degradation category scale as follows:

5 – Degradation is inaudible

4 – Degradation is audible but not annoying

3 – Degradation is slightly annoying

2 – Degradation is annoying

1 – Degradation is very annoying.

B.3 *Comparison between the sensitivity of an ACR and a DCR procedure for the same coder configurations*

Tables B-1 and B-2 summarize the results obtained with ACR test and DCR test respectively for the evaluation of three 32 kbit/s ADPCM algorithms.

Figures B-1, B-2 and B-3 show the mean opinion score (MOS) and degradation mean opinion store (DMOS) obtained by the same conditions with the two procedures (ACR and DCR respectively).

From these figures one can note:

– a good agreement between the results obtained with the two procedures;

– a larger spread of the DMOS obtained for MNRU circuits with Q values ranging from 10 dB to 35 dB, and a good anchoring of the judgements of listeners ("null pairs" have obtained a score of 4.98);

– a higher sensitivity of the DCR procedure in the range of good telephone quality ($20 < Q < 35$ dB).

These sensitivities can be quantified by means of a statistical multiple comparison test. When an *a posteriori* comparison of codecs is needed a Tuckey [7] honestly significant difference (HSD) test can be applied effectively. The HSD test is designed to make all pairwise comparisons among the means and to determine the significance of the differences in the mean values. Under identical conditions ($\alpha = 0.01$, $k = 2$, $N = 225$, fixed mode) the HSD limit value ($q_{\alpha,k,N}$) is 3.70 and since the residual errors for ACR and DCR procedures are about the same (0.42), two means can be declared as significantly different if:

$$\Delta = | \overline{X_i} - \overline{X_j} | > 0.21$$

This difference, expressed in Q value, corresponds to:

Range in Q (dB)	ACR test Δ	Range in Q (dB)	DCR test Δ
15 – 20	1.48	15 – 20	1.07
20 – 25	1.87	20 – 25	1.14
25 – 30	3.00	25 – 30	1.36

This means that the resolution of the DCR test may be twice that of the ACR test in terms of Q value in the range of good telephone quality.

B.4 *Conclusion*

A good agreement between the results obtained with the two procedures (ACR and DCR) has been found. The presence of a reference before each judgement for the DCR procedure ensures a good anchoring of the listener's rating and consequently a larger spread of the degradation mean opinion score (DMOS) obtained by the coders. The evaluation of the coders based on the same speech samples leads to a better precision for the DCR procedure at a price, of course, of a decrease of the importance of the effort made to comprehend the samples in the overall quality judgement. Therefore the degradation category rating procedure seems well adapted to evaluate good telephone quality coders.

TABLE B-1

Mean opinion scores (MOS) and 95% confidence interval (INT) for ACR test

Test conditions	X		Y		Z	
	MOS	INT	MOS	INT	MOS	INT
PCM	3.81	0.45	3.89	0.13	4.16	0.13
PCM 2A	3.99	0.13	4.10	0.13	3.90	0.14
PCM 4A	3.35	0.12	4.02	0.14	3.70	0.14
PCM 8A	3.39	0.14	3.48	0.14	3.46	0.12
PCM 10^{-4}	3.31	0.15	3.55	0.14	3.15	0.16
PCM 10^{-3}	1.90	0.15	1.78	0.13	2.10	0.17
PCM +10 dB	3.94	0.15	4.02	0.12	4.14	0.11
PCM −15 dB	3.49	0.16	3.60	0.14	3.41	0.16
ADPCM	3.60	0.15	3.41	0.13	3.65	0.12
ADPCM 2A	3.72	0.13	3.30	0.12	3.38	0.13
ADPCM 4A	3.14	0.13	2.85	0.13	2.63	0.13
ADPCM 8A	2.51	0.14	2.09	0.14	2.23	0.15
ADPCM 2T	3.77	0.12	3.33	0.13	3.42	0.13
ADPCM 4T	3.86	0.14	3.01	0.14	3.80	0.13
ADPCM 10^{-4}	3.54	0.11	3.28	0.12	2.81	0.15
ADPCM 10^{-3}	2.88	0.16	2.55	0.15	1.93	0.13
ADPCM +10 dB	3.80	0.14	3.55	0.14	3.61	0.13
ADPCM −15 dB	3.20	0.15	3.02	0.15	2.92	0.14
ADPCM, C 2A	2.44	0.16	2.62	0.16	2.23	0.14
ADPCM, C 4A	2.13	0.15	2.14	0.13	1.90	0.13
ADPCM, C 8A	1.98	0.14	1.84	0.13	1.59	0.12

S/N 40	3.52	0.15
S/N 35	3.18	0.17
S/N 25	2.04	0.15
S/N 15	1.23	0.09
Q 10	1.41	0.10
Q 15	2.34	0.11
Q 20	3.04	0.10
Q 25	3.61	0.09
Q 30	3.96	0.09

Note 1 – Votes combined across four speakers and two sentences.

Note 2 – Number of votes = 128 except for Q where N = 256.

TABLE B-2

Degradation mean opinion scores (DMOS) and 95% confidence intervals (INT) for DCR test

Test conditions	X		Y		Z	
	DMOS	INT	DMOS	INT	DMOS	INT
PCM	4.35	0.10			4.41	0.11
PCM 8A	3.48	0.16			3.33	0.15
PCM 10^{-3}	2.21	0.11			2.25	0.14
ADPCM	4.33	0.11	4.22	0.11	4.05	0.12
ADPCM 8A	2.63	0.14	2.35	0.14	2.38	0.17
ADPCM 10^{-3}	3.14	0.16	2.83	0.14	1.85	0.14
ADPCM 4T	4.29	0.10	3.69	0.14	4.09	0.13

Q 15	1.99	0.15
Q 20	2.97	0.17
Q 25	3.89	0.18
Q 30	4.66	0.10
Q 35	4.81	0.09
Origin	4.98	0.03

Note 1 – Votes combined across four speakers and two sentences.

Note 2 – Number of votes = 128.

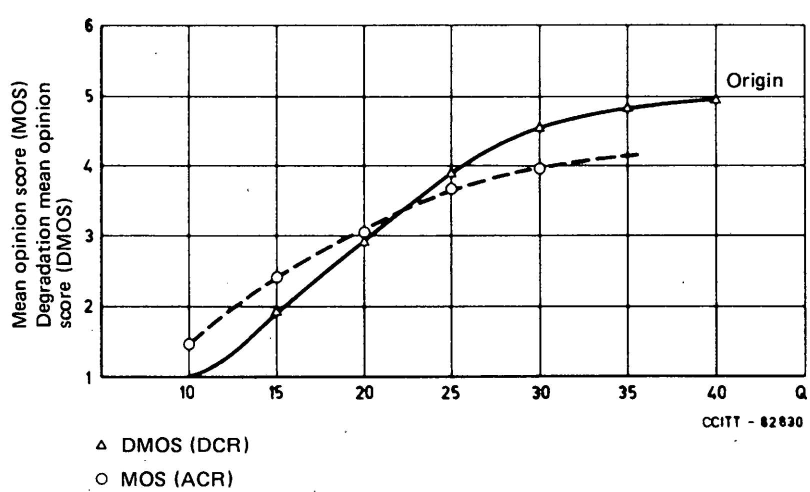
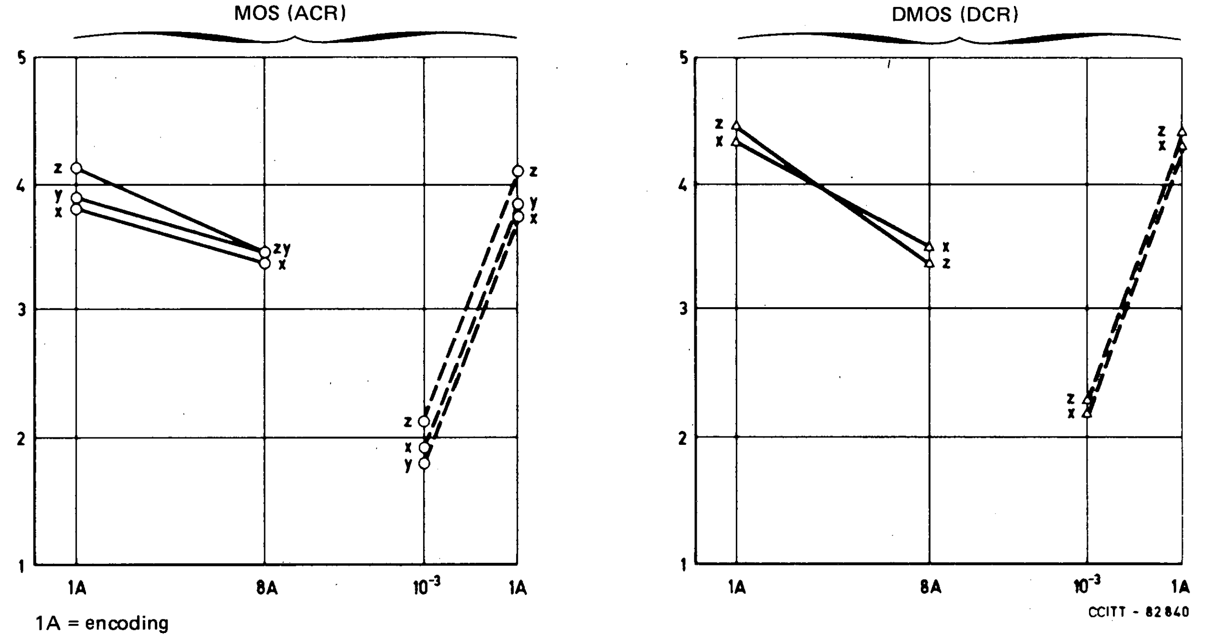


FIGURE B-1

MNRU reference conditions

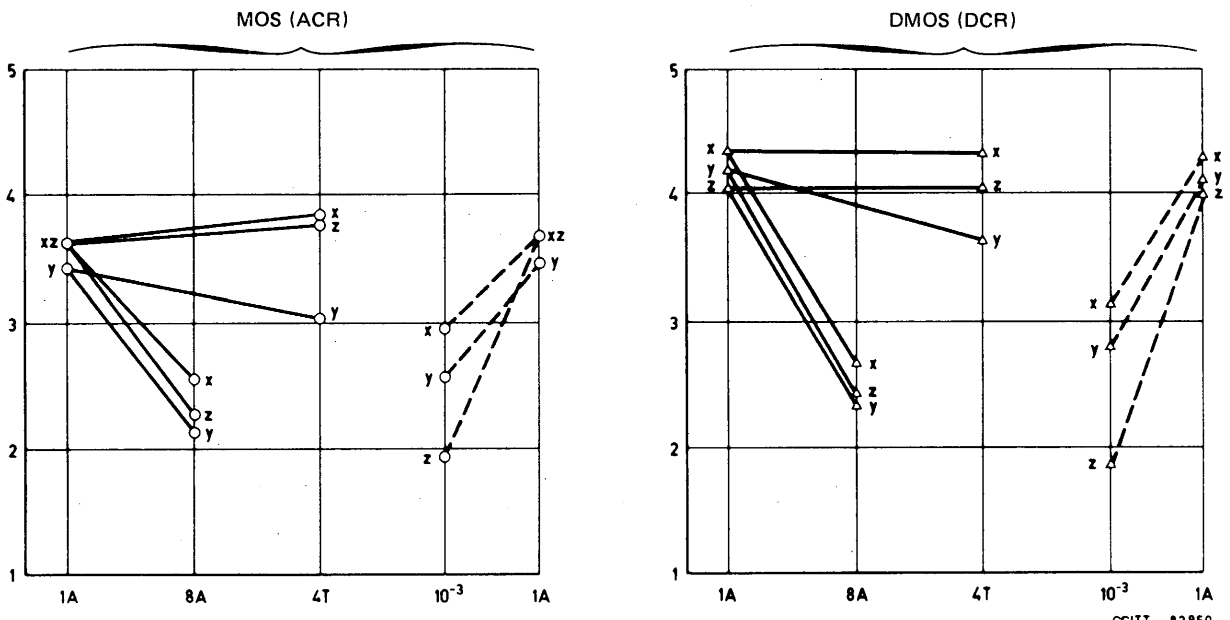


1A = encoding
8A = 8 asynchronous encodings

10^{-3} = line errors injected at a rate of 10^{-3}

FIGURE B-2

PCM reference conditions



1A = 1 encoding
8A = 8 asynchronous encodings

4T = 4 transcodings (synchronous)
10^{-3} = line errors injected at a rate of 10^{-3}

FIGURE B-3

ADPCM coders conditions

ANNEX C

(to Supplement No. 14)

Principles of a method used by one organization in analyzing digital codec performance (Bell Communications Research, Inc.)

Hypothetical mean opinion score (MOS) results obtained from a subjective test conducted according to the methodology of Annex A are shown in Figures 1 to 4 of this Supplement. (Straight lines are used simply to connect data points.) Generally such results will display a saturation effect at and near the very good conditions (high MOS) and the very bad conditions (low MOS). (For high MOS, the saturation is caused by the 5-point scale and possibly by the idle circuit noise of the subjective test system without added impairments, e.g. idle circuit noise, Q, codec quantization distortion. For low MOS, the saturation is caused by the 5-point scale.)

An analytic method of data analysis used by Bell Communications Research, Inc. provides a value called "fit mean" for each test condition [8]. The fit means are then used in analysis of the data. (Fit means and MOSs are nearly equal over the mid-range of MOS values; however, fit means are not numerically constrained for extremely good and extremely bad conditions as are MOSs.) Plots of test results in terms of fit means will be similar to those of Figures 1 to 4, but will numerically exhibit greater spread.

The objective of the analysis is to determine a function:

$$Q_s = f(L) \qquad \text{(C-1)}$$

where

Q_s = Q value for the codec quantization distortion,

L = Line bit rate (e.g. in kbit/s).

(A simple linear relation may be possible in some cases while other cases may require a more complex function.)

The codec Q can then be estimated from Equation (C-1).

Determination of Equation (C-1) needs to take into account in an appropriate manner the saturation effects discussed earlier. For example, the codec design line bit rate may correspond to the middle data point of Figure 3 for which there appears to be a modest saturation effect.

Similarly, the equivalence function (SNR *vs*, Q) internal to the test may need to be considered. [This function is determined from appropriate functions fitted to the curves (in terms of fit means) similar to the curves of Figures 1 and 2.]

An important consideration in the analysis method is obtaining predicted performance values (fit means, Q) approximating as closely as possible the actual performance values (fit means similar in form to the curves of Figures 3 and 4 or Q values obtained by converting the fit mean values to Q values using an appropriate function fitted to fit mean data similar in form to the data of Figure 2). (For present purposes, it is assumed that for asynchronously tandemed codecs the combining law is 15 $\log_{10} n$, where n is the number of tandemed identical codecs. It may also be desirable to include the determination of the combining law in the analysis.)

The Q values obtained for a digital process according to the procedure described above can be used in various ways to assess the effect of quantization distortion on telephone connection performance. The subjective opinion model of Supplement No. 3 is in terms of corrected reference equivalent and idle circuit noise level. This model requires that the Q for an overall connection be converted to an equivalent SNR which can then be converted to an idle circuit noise level based on knowledge of the speech levels for connections of interest.

Equivalence functions used for this purpose have been found to vary (see [9], [10], [11], [12] and [13]). It is not clear if there exists a unique equivalence function which can be agreed on, and what that equivalence function should be. (Perhaps a basic equivalence function should be based on conversational test results.)

The Q values obtained for digital processes can also be used as a basis for specifying codec network application rules in terms of the number of asychronously tandemed 8-bit, μ-255 codecs. For the model of Supplement No. 3, the relation between the number of such codecs and Q (based on a 15 $\log_{10} n$ law) is as follows:

Number 2)	1	2	3	4	5	6	8	10	12	14
Q(dB)	37	32.5	30	28	26.5	25.5	23.5	22	21	20

According to the model of Supplement No. 3, a 7-bit, μ-255 PCM codec would correspond to about 2.5 asynchronously tandemed 8-bit, μ-255 codecs. (Note that the value of Q = 37 dB is 3-4 dB greater than the minimum S/D values of Recommendation G.712 [15]; it is assumed that average 8-bit systems perform at the higher value.)

References

[1] CCITT Recommendation *32 kbit/s adaptive differential pulse code modulation (ADPCM)*, Red Book, Vol. III, Rec. G.721, ITU, Geneva, 1985.

[2] GOODMAN (D. J.), NASH (R. D.): Subjective quality of the same speech transmission conditions in seven different countries, *Proc. ICASSP 82* (International Conference on Acoustics, Speech and Signal Processing), Vol. 2, Paris, May 1982.

[3] DAUMER (W. R.), CAVANAUGH (J. R.): A subjective comparison of selected digital codecs for speech, *Bell System Technical Journal*, Vol. 57, No. 9, November 1978.

[4] RICHARDS (D. L.), BARNES (G. J.): Pay-off between quantizing distortion and injected circuit noise, *Proc. ICASSP 82* (International Conference on Acoustics, Speech and Signal Processing), Vol. 2, Paris, May 1982.

[5] COMBESCURE (P.) *et al.* Quality evaluation of speech coded at 32 kbit/s by means of degradation category ratings, *Proc. ICASSP 82* (International Conference on Acoustics, Speech and Signal Processing), Vol. 2, Paris, May 1982.

[6] CCIR Doc. 11/17 *Subjective assessment of the quality of television pictures* (EBU), Study Period 1978-1982.

[7] TUCKEY: The problem of multiple comparisons, *Ditton*, Princeton University, Ed. 1953.

[8] CAVANAUGH (J. R.), HATCH (R. W.), SULLIVAN (J. L.) Models for the subjective effects of loss, noise and echo on telephone connections, *Bell System Technical Journal*, Vol. 55, November 1976.

[9] CCITT Contribution COM XII-No. 24 (Study on determination of subjectively equivalent noise (NTT)), Study Period 1981-1984.

[10] CCITT Contribution COM XII-No. 61 (Subjectively equivalent noise for linear and carbon microphone originated speech signals – Bell Northern Research, Canada), Study Period 1981-1984.

[11] CCITT Contribution COM XII-No. 124 (Application of information index to quantizing noise in PCM – France), Study Period 1981-1984.

[12] CCITT Contribution COM XII-No. 130 (Subjective equivalence functions for consideration in a quantization distortion opinion model – Bell Northern Research, Canada), Study Period 1981-1984.

[13] CCITT Contribution COM XII-No. 162 (Transmission performance of digital systems – COMSAT), Study Period 1981-1984.

[14] CCITT Recommendation *Transmission impairments*, Yellow Book, Vol. III.1, Rec. G.113, ITU, Geneva, 1981.

[15] CCITT Recommendation *Performance characteristics of PCM channels audio frequencies*, Yellow Book, Vol. III.3, Rec. G.712, ITU, Geneva, 1981.

2) This should not be interpreted as the number of *qdus* [14] to be used for international planning; the relation between the number of codecs and Q applies for the model of Supplement No. 3 which has been used in planning studies in the United States.

Printed in Switzerland — ISBN 92-61-02131-X