5i’

SECTION 3

TRANSMISSION STANDARDS

Recommendation P.48

SPECIFICATION FOR AN INTERMEDIATE REFERENCE SYSTEM

(Geneva, 1976; amended at Geneva, 1980,

Malaga-Torremolinos, 1984, Melbourne, 1988)

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 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.

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.

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.

Figure 1/P.48 p.

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 1/P.35. 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.

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 /0° 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 /0° 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 ± | % 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 /0° 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 /0° 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]).

H.T. [T1.48]

TABLE 1/P.48

Nominal sending sensitivities and receiving sensitivities of the IRS

(These values were adopted provisionally)

Frequency (Hz) *S* *Je* }	*S* *mJ*	{
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

Table 1/P.48 [T1.48], p.

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 Shapes and tolerances on sensitivities of sending and receiving parts

The shape of the sensitivity/frequency characteristics of the sending and receiving parts of the IRS shall lie within the limits of masks formed by Table 2/P.48 and plotted in Figures 2/P.48 and 3/P.48. The sending and receiving loudness ratings shall both be set to 0 ± 0.2 dB when calculated in accordance with the principles laid down in Recommendation P.79.

Note — One excursion above or one excursion below the limits is permitted provided that:

a) the excursion is no greater than 2 dB above the upper or below the lower limit;

b) the width of the excursion as it breaks the appropriate limit is no greater than 1/10th of the frequency at the maximum or minimum of the excursion.

H.T. [T2.48]
TABLE 2/P.48
Coordinates of sending and receiving sensitivity limit curves

Limite curve	Frequency (Hz)	Sending sensitivity (dB with respect to an arbitrary level)	Frequency (Hz)	Receiving sensitivity (dB with respect to an arbitrary level)
Upper limit	100 200 400 3400 3600 6000	—41 —16 —6 +6 +4 —60	100 200 300 500 3400 3600 4500	—24 0 +9 +14 +16 +13 —40
Lower limit	Under 200 200 400 3000 3400 Over 3400	—∞ —21 —11 —1 —4 —∞	Under 200 200 300 500 3200 3400 Over 3400	

} —∞ —20 +4 +9 +10 +4 —∞ }	{				

Tableau [T2.48] p. 3

Figure 2/P.48, p. 4

Figure 3/P.48, p. 5

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.

SECTION 4

OBJECTIVE MEASURING APPARATUS

Recommendation P.50

ARTIFICIAL VOICES

(Melbourne, 1988)

The CCITT,

considering

(a) that it is highly desirable to perform objective telephonometric measurements by means of a mathematically defined signal reproducing the characteristics of human speech;

(b) that the standardization of such a signal is a subject for general study by the CCITT,

recommends

the use of the artificial voice described in this Recommendation.

Note 1 — For objective loudness rating measurements, less sophisticated signals such as pink noise or spectrum-shaped Gaussian noise can be used instead of the artificial voice.

Note 2 — The artificial voice here recommended has not yet been exhaustively tested in all possible applications; further studies being carried out within Question 14/XII.

1 Introduction

The signal here described reproduces the characteristics of human speech for the purposes of characterizing linear and non-linear telecommunication systems and devices, which are intended for the transduction or transmission of speech. It is known that for some purposes, such as objective loudness rating measurements , more simple signals can be used as well. Examples of such signals are pink noise or spectrum-shaped Gaussian noise, which nevertheless cannot be referred to as ‘‘artificial voice’’ for the purpose of this Recommendation.

The artificial voice is a signal that is mathematically defined and that reproduces the time and spectral characteristics of speech which significantly affect the performances of telecommunication systems [1]. Two kinds of artificial voice are defined, reproducing respectively the spectral characteristics of female and male speech.

The following time and spectral characteristics of real speech are reproduced by the artificial voice:

a) long-term average spectrum,

The specifications given here are subject to future enhancement and therefore should be regarded as provisional.

b) short-term spectrum,

c) instantaneous amplitude distribution,

d) voiced and unvoiced structure of speech waveform,

e) syllabic envelope.

2 Scope, purpose and definition

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. carbon microphones, loudspeaking telephone sets, nonlinear coders, echo controlling devices, syllabic compandors, nonlinear systems in general.

The use of the artificial voice instead of real speech has the advantage of both being more easily generated and having a smaller variability than samples of real voice.

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.

2.2 *Definition*

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

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.50.

Figure 1/P.50, p.

3.1 *electrical artificial voice*

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

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

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.

4 Characteristics

4.1 *Long-term average spectrum*

The third octave filtered long-term average spectrum of the artificial voice is given in Figure 2/P.50 and Table 1/P.50, normalized for a wideband sound pressure level of —4.7 dBPa. The table is calculated from the theoretical equation reported in [2].

Note — The values of the long-term spectrum of the artificial voice at the MRP can be derived from the equation:

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

where $S(f)$ is the spectrum density in dB relative to 1 pW/m^2 sound intensity per Hertz at the frequency f. The definition frequency range is from 100 Hz to 8 kHz.

The curve of the spectrum is shown in Figure 2/P.50. The values of $S(f)$ at 1/3 octave ISO frequencies are given in the fourth column of Table 1/P.50. The tolerances are given in the fifth column of Table 1/P.50. The tolerances below 200 Hz apply onto to the male artificial voice.

The total sound pressure level of the spectrum defined in Equation (1-1) is —4.7 dBPa. However, this spectrum is also applicable for the levels from —19.7 to +10.3 dPBa. In other words, the first term of Equation (1-1) may range from —391.44 to —361.44.

Figure 2/P.50, p.

H.T. [T1.50]
TABLE 1/P.50
Long-term spectrum of the artificial voice

1/3 octave center frequency (Hz) (1)	Bandwidth correction factor 10 log 10 Δf (dB) (2)	Sound pressure level (third octave) (dBPa) (3)	Spectrum density (dB) (3) — (2)	Tolerance (dB)
100	13.6	—23.1	—36.7	—
125	14.6	—19.2	—33.8	+3, —6 | ua)
160	15.6	—16.4	—32,7	+3, —6 | ua)
200	16.6	—14.4	—31,7	+3, —6
250	17.6	—13.4	—31,7	±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	—37,7	±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	—

a) The given tolerances apply to the long-term spectrum of male speech and must also be complied with by speech shaped noises. However, they do not apply to the female speech spectrum, whose energy content in this frequency range is virtually negligible.

Table 1/P.50 [T1.50], p.

4.2 *Short-term spectrum*

The short-term spectrum characteristics of the male and female artificial voices are described in Annex A.

4.3 *Instantaneous amplitude distribution*

The probability density distribution of the instantaneous amplitude of the artificial voice is shown in Figure 3/P.50 [3].

Figure 3/P.50, p.

4.4 *Segmental power level distribution*

The segmental power level distribution of the artificial voice, measured on time windows of 16 ms, is shown in Figure 4/P.50. 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) [4], [5].

Figure 4/P.50, p.

4.5 *Spectrum of the modulation envelope*

The spectrum of the modulation envelope waveform is shown in Figure 5/P.50 and should be reproduced with a tolerance of ± | dB on the whole frequency range.

Figure 5/P.50, p.

4.6 *Time convergence*

The 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 s.

5 Generation method

Figure 6/P.50 shows a block diagram of the generation process of the artificial voice signals, a glottal excitation signal and a random noise, to a time-variant spectrum shaping filter. The artificial voice generated by the glottal excitation signal and by the random noise corresponds respectively to voiced and unvoiced sounds. The frequency response of the spectrum shaping filter simulates the transmission characteristics of the vocal tract.

Figure 6/P.50, p.

5.1 *Excitation source signal*

The artifical voice is obtained by randomly alternating four basic unit elements, each containing voiced and unvoiced segments. While one unit element starts with an unvoiced sound, followed by a voiced one, the other three elements start with a voiced sound, followed by an unvoiced one and end with a voiced sound again (see also Figure 9/P.50). The ratio of the unvoiced sound duration T_{uv} to the total duration of voiced segments T_v for each unit element is 0.25. The duration $T = T_{uv} + T_v$ of unit elements varies according to the following equation:

$$T = —3.486 (\log_{10} r)$$

where r | denotes a uniformly distributed random number (0.371 r 0.609).

The time lengths of the voiced and unvoiced sounds of the four unit elements are as follows:

Element a: Unvoiced (T_{uv}) ; Voiced (T_v)

Element b: Voiced (T_v/4) + Unvoiced (T_{uv}) + Voiced (3T_v/4)

Element c: Voiced (T_v/2) + Unvoiced (T_{uv}) ; Voiced (T_v/2)

Element d: Voiced (3T_v/4) + Unvoiced (T_{uv}) + Voiced (T_v/4)

Unit elements shall be randomly iterated for at least 10 s in order to comply with the artificial voice characteristics as specified in § 4.

5.2 *Glottal excitation*

The glottal excitation signal is a periodic waveform as shown in Figure 7/P.50. The pitch frequency (1/T_0 in Figure 7/P.50) varies according to the variation pattern shown in Figure 8/P.50 during the period T_v. The starting value of the pitch frequency (F_s in Figure 8/P.50) is determined according to the following relationships:

$F_s = F_c — 31.82\ T_v + 39.4\ R$ | for the male artificial voice

$F_s = F_c — 51.85\ T_v + 64.2\ R$ | for the female artificial voice

where F_c and R respectively denote the center frequency and a uniformly distributed random variable ($—1 < R < 1$). F_c is 128 Hz for the male artificial voice and 215 Hz for the female artificial voice. In the trapezoid of the pitch frequency variation pattern, the area of the trapezoid above F_c should be equal to that below F_c (shaded in Figure 8/P.50). For the elements b), c) and d) in Figure 7/P.50 the pitch frequency variation pattern applies to the combination of the two voiced parts, irrespectively of where the unvoiced segment is inserted.

Figure 7/P.50, p.

Figure 8/P.50, p.

5.3 *Unvoiced sounds*

The transfer function of the low-pass filter located after the random noise generator (low emphasis) is $1/(1 - z^{-1})$, where z^{-1} denotes the unit delay.

5.4 *Power envelope*

The power envelope of each unit element of the excitation source signal is so controlled that the short-term segmental power (evaluated over 2 ms intervals) of the artificial voice varies according to the patterns shown in a) to d) of Figure 9/P.50. This is obtained by utilizing the following relationship providing input and output signals of the spectrum shaping filter:

where:

P_{in} is the input power to the spectrum shaping filter

P_{out} is the output power from the spectrum shaping filter

k_i is the *i* th coefficient of the spectrum shaping filter.

The rising, stationary and decay times of each trapezoid of a) to d) of Figure 9/P.50 shall be mutually related by the same proportionality coefficients (2 | | | |) of the pitch frequency variation pattern shown in Figure 8/P.50. For each unit element, the average power of unvoiced sounds (P_{uv}) shall be 17.5 dB less than the average power of voiced sounds (P_v).

5.5 *Spectrum shaping filter*

The spectrum shaping filter has a 12th order lattice structure as shown in Figure 10/P.50. Sixteen groups, each of 12 filtering coefficients (k_1 — k_{12}), are defined; thirteen groups shall be used for generating the voiced part, while three groups shall be used for generating the unvoiced part. These coefficients are listed in Table 2/P.50 both for male and female artificial voices.

The twelve filter coefficients shall be updated every 60 ms while generating the signal. More precisely, during each 60 ms period the actual filtering coefficients must be adjourned every 2 ms, by linearly interpolating between the two sets of values adopted for subsequent 60 ms intervals. In the voiced sound part, each of 13 groups of coefficients shall be chosen at random once every 780 ms (= 60 ms × 13), and in the unvoiced sound part each of 3 groups of coefficients shall be chosen at random once every 180 ms (= 60 ms × 3).

Note — The described implementation of the shaping filter should be considered as an example and is not an integral part of this Recommendation. Any other implementation providing the same transfer function can be alternatively used.

Figure 9/P.50, p. 15

Figure 10/P.50, p. 16

H.T. [T2.50]

TABLE 2/P.50

Coefficients

k

i

a) k

k 1	*k* 2	*k* 3	*k* 4	*k* 5	*k* 6	*k* 7	*k* 8	*k* 9	| fIk 1 0	| fIk 1 1	| fIk 1 2		
Unvoiced	1 2 3	-0.471 -0.284 -0.025	-0.108 -0.468 -0.496	0.024 0.030 -0.176	-0.048 0.090 0.162	0.140 0.124 0.236	0.036 -0.020 -0.012	0.054 0.087 0.068	0.004 0.067 0.001	0.123 0.131 0.096	0.044 0.011 0.029	0.099 0.076 0.086	-0.003 -0.024 -0.018
	1	0.974	0.219	0.025	-0.123	-0.132	-0.203	-0.103	-0.174	-0.079	-0.153	-0.010	-0.061
	2	0,629	-0.152	-0.138	-0.142	-0.118	-0.135	0.147	0.019	0.077	-0.040	0.029	-0.007
	3	0.599	-0.119	0.067	0.051	0.103	0.023	0.106	0.036	-0.006	-0.133	-0.052	-0.094
	4	0.164	-0.364	-0.248	-0.076	0.168	0.072	0.103	0.045	0.112	0.010	0.048	-0.034
	5	0.842	0.022	0.171	0.173	0.067	-0.057	0.089	-0.045	-0.039	-0.134	-0.034	-0.122
	6	0.933	-0.537	-0.137	-0.161	-0.216	-0.139	0.115	-0.042	0.027	-0.163	0.102	-0.107
Voiced	7	0.937	-0.413	0.132	-0.059	-0.103	-0.134	0.047	-0.115	-0.105	-0.097	0.039	-0.108
	8	0.965	-0.034	0.032	0.001	-0.107	-0.189	-0.057	-0.175	-0.109	-0.163	-0.003	-0.055
	9	0.870	-0.476	-0.016	-0.136	-0.125	-0.107	0.091	-0.008	0.021	-0.128	0.042	-0.069
	10	0.686	-0.030	0.178	0.197	0.155	-0.026	0.078	0.004	-0.001	-0.128	-0.004	-0.102
	11	0.963	-0.232	0.086	-0.018	-0.147	-0.192	-0.040	-0.179	-0.144	-0.133	0.042	-0.042
	12	0.930	-0.461	0.071	-0.144	-0.122	-0.096	0.034	-0.066	-0.021	-0.171	0.067	-0.091
	13	0.949	-0.334	0.143	-0.040	-0.112	-0.161	0.010	-0.156	-0.123	-0.119	0.049	-0.070

Tableau 2/P.50 [T2.50], p. 17

Blanc

ANNEX A

(to Recommendation P.50)

Short-term spectrum characteristics of the artificial voice

The artificial voice is generated by randomly selecting each of sixteen short-term spectrum patterns once ever 960 ms (= 60 ms × 16 patterns). The spectrum density of each pattern is provided by Equation (A-1) and Table A-1/P.50, and the short-term spectrum of the signal during the 60 ms interval occurring between any two subsequent pattern selections varies smoothly from one pattern to the next.

Note — The spectrum patterns in Equation (A-10) and Table A-1/P.50 are expressed in power normalized form.

Blanc

H.T. [T3.50]

TABLE A-1/P.50
{
Coefficients
A
i
j
}
{
a) A
i
j
for male artificial voice
}

j i	0	1	2	3	4	5	6	7	8	9	| 0	| 1	| 2
1	2.09230	—1.33222	1.32175	—1.14200	0.99352	—0.94634	0.72684	—0.63263	0.41196	—0.42858	0.22070	—0.19746	0.10900
2	9.34810	—8.55934	7.35732	—6.35320	5.33999	—4.47238	3.62417	—2.85246	2.12260	—1.49424	0.93988	—0.44998	0.12400
3	11.69068	—10.91138	9.46588	—8.11729	6.94160	—5.90977	4.95137	—3.89587	2.88750	—1.97671	1.14892	—0.50255	0.12100
4	12.56830	—11.81209	10.36030	—8.82879	7.37947	—6.01017	4.66740	—3.46913	2.42182	—1.60880	0.91652	—0.39648	0.12000
5	6.83438	—6.18275	5.59089	—4.71866	4.06004	—3.44767	2.65380	—2.12140	1.50334	—1.07904	0.64553	—0.31816	0.11500
6	12.37251	—11.52358	9.89962	—8.31774	6.99062	—5.86272	4.69809	—3.56806	2.53340	—1.70522	0.99232	—0.45403	0.13400
7	21.07637	—19.62125	16.56781	—13.67518	11.41379	—9.61940	7.93529	—6.32841	4.92443	—3.53539	2.09095	—0.86543	0.18100
8	30.77371	—29.17365	25.52254	—21.51978	17.80583	—14.30488	10.87190	—7.71572	5.14643	—3.20113	1.72149	—0.68054	0.14400
9	4.18618	—3.36611	3.36793	—2.92133	2.38452	—2.06047	1.57550	—1.34240	0.84994	—0.70462	0.38685	—0.21857	0.12100
10	14.12359	—13.14611	11.25804	—9.47510	7.97588	—6.70717	5.44803	—4.23843	3.10807	—2.12879	1.25096	—0.53230	0.12600
11	26.36971	—24.95984	21.80496	—18.41045	15.30642	—12.49415	9.84879	—7.40287	5.29262	—3.43906	1.84980	—0.71546	0.14800
12	11.50808	—10.74609	9.34328	—7.91953	6.66959	—5.54500	4.34328	—3.27036	2.33714	—1.61333	0.96597	—0.44666	0.13500
13	5.32020	—4.61998	4.29145	—3.62118	3.01310	—2.67071	2.13992	—1.72147	1.22163	—0.93163	0.53317	—0.28989	0.11900
14	20.61945	—19.39682	16.80034	—14.14817	11.84307	—9.78712	7.73534	—5.77921	4.06200	—2.66324	1.49831	—0.59887	0.12600
15	30.02641	—28.42244	24.75314	—20.70178	16.98199	—13.72247	10.81050	—8.20966	5.94148	—3.90501	2.11507	—0.81306	0.16400
16	27.62370	—26.17896	22.93678	—19.42253	16.18997	—13.17171	10.19859	—7.42299	5.07437	—3.21481	1.73980	—0.67818	0.14000

Tableau A-1/P.50 [T3.50], A L'ITALIENNE, p. 18

References

[1] CCITT — Contribution COM XII-No. 76, Study Period 1981-1984

[2] CCITT — Contribution COM XII-No. 108, Study Period 1981-1984

[3] CCITT — Contribution COM XII-No. 11, Study Period 1981-1984

[4] CCITT — Contribution COM XII-No. 150, Study Period 1981-1984

[5] CCITT — Contribution COM XII-No. 132, Study Period 1981-1984

Recommendation P.51

ARTIFICIAL EAR AND ARTIFICIAL MOUTH

(amended at Mar del Plata, 1968, Geneva, 1972, 1976,

1980, Malaga-Torremolinos, 1984 and Melbourne, 1988)

The CCITT,

considering

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

(b) that the standardization of the artificial ear and mouth used in the construction of such apparatus is a subject for general study by the CCITT,

recommends

(1) the use of the artificial ears described in § 1 of this Recommendation;

(2) the use of the artificial mouth described in § 2 of this Recommendation.

Note — Administrations may, if they wish, use 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.

1 Artificial ears

Three types of artificial ears are defined:

1) a wideband type for audiometricand telephonometric measurements,

2) a special type for measuring insert earphones,

3) a type which faithfully reproduces the characteristics of the average human ear, for use in the laboratory.

Type 1 is covered by IEC Recommendation 318 [1], the second IEC Recommendation 711 [2] and the third is the object of further study in the IEC.

It is recommended that the artificial ear conforming to IEC 318 [1] should be used for measurements on supra-aural earphones, e.g. handsets, and that the insert ear simulator conforming to IEC 711 [2] should be used for measurements on insert earphones, e.g. some headsets.

Note 1 — 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.

Note 2 — The sound pressure measured by the IEC 711 artificial ear is referred to the eardrum. The correction function given in Table 1/P.51 shall be used for converting data to the ear reference point (ERP), where loudness rating algorithms (Recommendation P.79) are based. The corrections apply to free field open-ear conditions and to partially or totally occluded conditions as well.

H.T. [T1.51]
TABLE 1/P.51

Frequency (Hz)	{ S_{DE} (dB) }
100	0.0
125	0.0
160	0.0
200	0.0
250	0.0
315	—0.2
400	—0.5
500	—1.1
630	—1.0
800	—1.8
1000	—2.0
1250	—2.5
1600	—4.1
2000	—7.2
2500	—10.6
3150	—10.4
4000	—6.0
5000	—2.1

S_{DE} is the transfer function eardrum to ERP: $S_{DE} = 10 \log \frac{f\,IP_E}{f\,IP_D}$ (dB), where

P sound pressure at the ERP

P sound pressure at the eardrum.

Table 1/P.51 [T1.51], p.

2 Artificial mouth

2.1 *Introduction*

The artificial mouth is a device that accurately reproduces the acoustic field generated by the human mouth in the near field. It is used for measuring objectively the sending characteristics of handset-equipped telephone sets as specified in Recommendation P.64. It may also be used for measuring the sending characteristics of loudspeaking telephones at distances up to 0.5 m from the lip plane, but the accuracy with which it reproduces the sound field of the human mouth is slightly reduced.

2.2 *Definitions*

2.2.1 **lip ring**

Circular ring of thin rigid rod, having a diameter of 25 mm and less than 2 mm thick. It shall be constructed of non-magnetic material and be solidly fixed to the case of the artificial mouth. The lip ring defines both the reference axis of the mouth and the mouth reference point.

Note — The provision of the lip ring for locating the lip planes and the reference axis is not mandatory. However, when not provided, adequate markings or other suitable geometric reference shall be alternatively available.

2.2.2 **lip plane**

Outer plane of the lip ring.

2.2.3 **reference axis**

The line perpendicular to the lip plane containing the center of the lip ring.

2.2.4 **vertical plane**

A plane containing the reference axis that divides the mouth into symmetrical halves. It shall be vertically oriented in order to reproduce the acoustic field generated by a person in the upright position.

2.2.5 **horizontal plane**

The plane containing the reference axis, perpendicular to the vertical plane. It shall be horizontally oriented in order to reproduce the acoustic field generated by a person in the upright position.

2.2.6 **mouth reference point (MRP)**

The point on the reference axis, 25 mm in front of the lip plane.

2.2.7 **normalized free-field response (at a given point)**

Difference between the third-octave spectrum level of the signal delivered by the artificial mouth at a given point in the free field and the third-octave spectrum level of the signal delivered simultaneously at the MRP. The characteristic is measured by feeding the artificial voice (see Recommendtion P.50) a speech-shaped random noise or a pink noise.

2.2.8. **reference obstacle**

Disc constructed of hard, stable and on-megnetic material, such as brass, having a diameter of 63 mm and 5 mm thick. In order to measure the normalized obstacle diffraction, it shall be fitted with a ¼" pressure microphone, mounted at the centre with the diaphragm flush on the disc surface.

2.2.9 **normalized obstacle diffraction**

Difference between the third-octave spectrum level of the acoustic pressure delivered by the artificial mouth at the surface of the reference obstacle and the third-octave spectrum level of the pressure simultaneously delivered at the point on the reference axis, 500 mm in front of the lip plane. The characteristic is defined for positions of the reference obstacle in front of the artificial mouth, with the disc axis coinciding with the reference axis, and is measured by feeding the artificial mouth with a complex signal such as the artificial voice, a speech shaped random noise or a pink noise.

2.3 *Acoustic characteristics of the artificial mouth*

2.3.1 *Normalized free-field response*

The normalized free-field response is specified at seventeen points: ten in the near field and seven in the far field. Near-field points are listed in Table 2/P.51, while far-field points are listed in Table 3/P.51.

Table 4/P.51 provides the normalized free-field response of the artificial mouth, together with tolerances, for the bandwidth between 100 Hz and 8 kHz. The requirements at each point not lying in the vertical plan shall also be met by the corresponding point in the symmetrical half-space.

The characteristic shall be checked by using appropriate microphones, as specified in Table 5/P.51. Pressure microphones shall be oriented with their axes perpendicular to the sound direction, while free-field microphones shall be oriented with their axes parallel to the direction of sound.

Note — If a compressor microphone is used with the mouth, it (or an equivalent dummy) shall be left in place while checking the normalized free-field response.

H.T. [T2.51]
TABLE 2/P.51
Coordinates of points in the near field

Measurement point On-axis displacement from the lip plane (mm) } Off-axis, perpendicular displacement (mm) }	{ {	
1	12.5	0
2	50 |	0
3	100 |	0
4	140 |	0
5	0 |	20 horizontal
6	0 |	40 horizontal
7	25 |	20 horizontal
8	25 |	40 horizontal
9	25 |	20 vertical (downwards)
10	25 |	40 vertical

Tableau 2/P.51 [T2.51], p. 20

H.T. [T3.51]
TABLE 3/P.51
Coordinates of points in the far field

Measurement point Distance from the lip plane (mm) } Azimuth angle (horizontal) (degree) } Elevation angle (vertical) (degree) }	{ { {		
11	500	0	0
12	500	0	+15
13	500	0	+30
14	500	0	—15
15	500	0	—30
16	500	15	0
17	500	30	0

Tableau 3/P.51 [T3.51], p. 21

Blanc

H.T. [T4.51]

TABLE 4a/P.51

Normalized free field response at points on axis in the near field

Frequency (Hz)	Measurement point 1 (dB)	2 (dB)	3 (dB)	4 (dB)	Tolerance (dB)
100	4.2	—5.0	—11.0	—13.6	±1.5
125	4.2	—5.0	—10.9	—13.6	±1.5
160	4.2	—5.0	—10.7	—13.6	±1.5
200	4.0	—5.0	—10.7	—13.3	±1.5
250	4.0	—5.0	—10.6	—13.2	±1.5
315	4.0	—5.0	—10.6	—13.2	±1.0
400	4.0	—5.0	—10.6	—13.2	±1.0
500	4.1	—5.0	—10.6	—13.2	±1.0
630	4.2	—4.9	—10.5	—13.4	±1.0
800	4.2	—4.8	—10.5	—13.4	±1.0
1000	4.1	—4.8	—10.4	—12.9	±1.0
1250	3.9	—4.8	—10.2	—12.7	±1.0
1600	3.8	—4.8	—10.0	—12.7	±1.0
2000	3.6	—4.7	—10.0	—12.7	±1.0
2500	3.5	—4.6	—9.4	—12.3	±1.0
3150	3.6	—4.6	—9.4	—12.0	±1.0
4000	3.7	—4.6	—9.7	—12.3	±1.5
5000	3.7	—4.5	—9.7	—12.6	±1.5
6300	3.8	—4.5	—9.7	—12.6	±1.5
8000	3.8	—4.9	—10.0	—12.7	±1.5

Tableau 4a/P.51 [T4.51], p. 22

H.T. [T5.51]

TABLE 4b/P.51

Normalized free-field response at points on axis in the near field

Frequency (Hz)	Measurement point 5 [a] (dB)	6 (dB)	7 (dB)	8 (dB)	9 (dB)	10 (dB)	Tolerance (dB)
100	5.2	—1.7	—1.4	—4.0	—1.6	—4.2	±1.5
125	5.2	—1.7	—1.3	—3.8	—1.5	—4.2	±1.5
160	5.2	—1.7	—1.2	—3.8	—1.5	—4.2	±1.5
200	5.2	—1.7	—1.2	—3.8	—1.5	—4.2	±1.5
250	5.2	—1.8	—1.3	—3.8	—1.4	—4.2	±1.5
315	5.1	—1.8	—1.3	—3.8	—1.3	—4.2	±1.0
400	5.1	—1.8	—1.3	—3.8	—1.3	—4.0	±1.0
500	5.0	—1.6	—1.3	—3.8	—1.3	—3.9	±1.0
630	5.0	—1.6	—1.3	—3.8	—1.3	—3.9	±1.0
800	5.0	—1.6	—1.3	—3.8	—1.3	—4.0	±1.0
1000	4.8	—1.7	—1.3	—3.9	—1.3	—4.1	±1.0
1250	4.8	—1.8	—1.4	—4.0	—1.3	—4.3	±1.0
1600	4.7	—1.8	—1.4	—3.8	—1.3	—4.0	±1.0
2000	4.7	—1.8	—1.2	—3.7	—1.3	—3.6	±1.0
2500	4.7	—1.9	—1.0	—3.6	—1.1	—3.5	±1.0
3150	4.7	—2.1	—1.1	—3.5	—1.2	—3.4	±1.0
4000	4.5	—2.9	—1.5	—4.1	—1.3	—3.0	±1.5
5000	3.8	—3.6	—1.5	—4.8	—1.3	—3.7	±1.5
6300	3.2	—4.8	—1.8	—5.2	—1.7	—3.7	±1.5
8000	2.5	—5.2	—2.0	—6.1	—2.2	—4.2	±1.5

[a] The measurements on the human mouth at point 5 are quite scattered, so the response at this point is only indicatively provided and no tolerances are specified.

Tableau 4b/P.51 [T5.51], p. 23

H.T. [T6.51]
TABLE 4c/P.51
Normalized free field response in the far field

Measurement point	Response (dB)	Tolerance (dB)
11	—24.0	± | .0
12	—24.0	± | .0
13	—24.0	± | .0
14	—24.0	± | .0
15	—24.0	± | .0
16	—24.0	± | .0
17	—24.0	± | .0

Tableau 4c/P.51 [T6.51], p. 24

H.T. [T7.51]
TABLE 5/P.51
Recommended microphone types for free-field measurements

Measurement point	Microphone size (max.)	Microphone equalization
1, 2, 5, 6, 7, 8, 9, 10	1/4"	Pressure
3, 4	1/2"	Pressure
11, 12, 13, 14, 15, 16, 17	1"	Free-field
MRP	1/4"	Pressure

Tableau 5/P.51 [T7.51], p. 25

2.3.2 *Normalized obstacle diffraction*

The normalized obstacle diffraction of the artificial mouth is defined at three points on the references axis, as specified in Table 6/P.51.

Note — If a compressor microphone is used with the mouth, it (or an equivalent dummy) shall be left in place while checking the normalized obstacle diffraction.

2.3.3 *Maximum deliverable sound pressure level*

The artificial mouth shall be able to deliver steadily the acoustic artificial voice at sound pressure levels up to at least +6 dBPa at the MRP.

2.3.4 *Harmonic distortion*

When delivering sine tones, with amplitudes up to +6 dBPa at the MRP, the harmonic distortion of the acoustic signal shall comply with the limits specified in Table 7/P.51.

H.T. [T8.51]

TABLE 6/P.51

Normalized obstacle diffraction

Frequency	Measurement point			
(Hz)	18 (dB)	19 (dB)	20 (dB)	Tolerance (dB)
100	32.2	27.0	21.7	±2.0
125	32.0	27.0	21.4	±2.0
160	32.0	27.3	21.4	±2.0
200	31.2	26.5	20.6	±2.0
250	31.2	26.5	20.5	±2.0
315	31.9	27.0	21.0	±1.5
400	31.8	27.0	20.9	±1.5
500	31.3	26.4	20.4	±1.5
630	31.0	26.0	20.0	±1.5
800	30.1	25.1	19.4	±1.5
1000	29.3	24.4	18.8	±1.5
1250	29.0	24.3	18.8	±1.5
1600	28.9	24.5	19.6	±1.5
2000	28.6	25.2	20.5	±1.5
2500	29.0	26.3	23.2	±1.5
3150	29.0	26.5	21.8	±1.5
4000	29.6	27.3	22.8	±2.0
5000	31.2	26.9	22.4	±2.0
6300	31.7	26.0	22.5	±2.0
8000	30.0	23.0	18.0	±2.0

Tableau 6/P.51 [T8.P.51], p. 26

H.T. [T9.51]

TABLE 7/P.51

Maximum harmonic distortion of the artificial mouth

	Harmonic distorsion	
	2nd harmonic	3rd harmonic
100 Hz-125 Hz	< | 0%	< | 0%
125 Hz-200 Hz	< | 4%	< | 4%
200 Hz-8 Hz	< | 1%	< | 1%

Tableau 7/P.51 [T9.P.51], p. 27

Blanc

2.3.5 *Linearity*

A positive or negative variation of 6 dB of the feeding electrical signal shall produce corresponding variation of 6 dB ± 0.5 dB at the MRP for outputs in the range —14 dBPa to +6 dBPa. This requirement shall be met both for complex excitations, such as the artificial voice, and for sine tones in the range 100 Hz to 8 kHz.

2.4 *Miscellaneous*

2.4.1 *Delivery conditions*

The artificial mouth shall be delivered by the maufacturer with the mechanical fixtures required to place the ½" calibration microphone at the MRP, as specified in Recommendation P.64. Suitable markings shall be engraved on the device housing for identifying the vertical plane position.

Each artificial mouth shall be delivered with a calibration chart specifying the free-field radiation and obstacle diffraction characteristics as defined in this Recommendation

2.4.2 *Stability*

The device shall be stable and reproducible.

2.4.3 *Stray magnetic field*

Neither the d.c. nor the a.c. magnetic stray fields generated by the artificial mouth shall neither influence the signal transduced by microphones under test.

It is recommended that the a.c. stray field produced at the MRP shall lie below the curve formed by the following coordinates:

Frequency
(Hz)
Magnetic output

(dB A/m/Pa) | 00 —10 1 | 00 —40 10 | 00 —40

It is also recommended that the d.c. stray field at the MRP be lower than 400 A/m.

Note — The recommended d.c. stray field limit of 400 A/m applies specifically to mouths intended for measuring electromagnetic microphones. For measuring other kinds of microphones, a higher limit of 1200 A/m is acceptable.

2.4.4 *Choice of model*

The results of measurements made on the BK 4219 source (no longer produced) and on the newer BK 4227, with its mouthpiece replaced by the UA 0899 conical adaptor, show a satisfactory agreement between the two models and compliance with the present Recommendation. The models actually used in tests shall always be stated, together with the results of measurements.

Note — It should be noted that the BK 4227 artificial mouth generates a d.c. stray magnetic field at the MRP which exceeds 400 A/m. It is then not suitable for measuring electromagnetic microphones.

References

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

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

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 defined in Recommendation P.56. Comparison of results using the active speech level meter and some meters described in this Recommendation can be found in Supplement No. 18.

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

— 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]. Further results can be found in Supplement No. 18 of the present volume.

Blanc

H.T. [T1.52]

TABLE 1/P.52

Principal characteristics of the various instruments used for monitoring the volume or peaks during telephone conversations or sound-programme transmissions

<table>
<tr><th>Type of instrument</th><th>Rectifier characteristic (see Note 3)</th><th>Time to reach 99% of final reading (milliseconds)</th><th>Integration time (milliseconds) (see Note 4)</th><th>Time to return to zero (value and definition)</th></tr>
<tr><td>(1)
‘‘Speech voltmeter’’ United Kingdom Post Office Type 3 (S.V.3) identical to the speech power meter of the l’ARAEN</td><td>2</td><td>230</td><td>100 (approx.)</td><td>equal to the integration time</td></tr>
<tr><td>(2)
VU meter (United States of America)
(see Note 1)</td><td>1.0 to 1.4</td><td>300</td><td>165 (approx.)</td><td>equal to the integration time</td></tr>
<tr><td>(3)
Speech power meter of the ‘‘SFERT volume indicator’’</td><td>2</td><td>around 400 to 650</td><td>200</td><td>equal to the integration time</td></tr>
<tr><td>(4)
Peak indicator for sound-programme transmissions used by the British Broadcasting Corporation (BBC Peak Programme Meter)
(see Note 2)</td><td>1</td><td></td><td>10 (see Note 5)</td><td>3 seconds for the pointer to fall to 26 dB</td></tr>
<tr><td>(5)
Maximum amplitude indicator used by the Federal German Republic
(type U 21)</td><td>1</td><td>around 80</td><td>5 (approx.)</td><td>1 or 2 seconds from 100% to 10% of the reading in the steady state</td></tr>
<tr><td>(6)
OIRT | (em | rogramme level meter:
type A sound meter
type B sound meter</td><td></td><td>for both types:
less than 300 ms for meters with pointer indication
and less than 150 ms for meters with light indication</td><td></td><td></td></tr>
</table>

} for both types: 1.5 to 2 seconds from the 0 dB point which is at 30% of the length of the operational section of the scale }	10 | (+- | 60 | (+- | 0	{			

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)
| | fIV (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.

Tableau 1/P.52 [T1.52], p. 28

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 | fR OF CIRCUIT NOISE)

Refer to Recommendation O.41, CCITT Blue 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.

Recommendation P.55

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.

Recommendation P.56

OBJECTIVE MEASUREMENT OF ACTIVE SPEECH LEVEL

(Melbourne, 1988)

1 Introduction

The CCITT considers it important that there should be a standardized method of objectively measuring speech level, so that measurements made by different Administrations may be directly comparable. Requirements of such a meter are that it should measure active speech level and should be independent of operator interpretation.

In this Recommendation, a meter is a complete unit that includes the input circuitry, filter (if necessary), processor and display. The processor includes the algorithm of the detection method.

In its present form, this meter can safely be used for laboratory experiments or can be used with care on operational circuits. Further study is continuing on:

a) how the meter can be used on 2-wire and 4-wire circuits to determine who is talking and whether it is an echo, and

b) how such an instrument can discriminate between speech and signalling, for example.

The method described herein maintains maximum comparability and continuity with past work, provided suitable monitoring is used, e.g. an operator performing the monitoring function. In particular, the new method yields 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. A method using operator monitoring can be found in Annex A.

This Recommendation describes a method that can be easily implemented using current technology. It also acts as a reference against which other methods can be compared. The purpose of this Recommendation is not to exclude any other method but to ensure that results from different methods give the same result.

Active speech level shall 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 (see § 4);

— *Method B* — measuring a quantity called active speech level, used for other purposes (see § 5).

Comparison of readings given by meters of methods A and B can be found in Supplement No. 18.

Note — This meter cannot be used to determine peak levels but sufficient information exists [1] giving the instantaneous peak/r.m.s. ratio, provided the signal has not been restricted or modified in any way, e.g. peak clipping.

2 Terminology

The recommended terminology is as follows:

speech volume until now used interchangeably with *speech level* , should in future 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 these terms [2], and other related terms such as those for the meters themselves [3], should be adjusted accordingly.

3 General

3.1 *Electrical, acoustic and other levels*

This Recommendation 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 (e.g., acoustical). 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).

3.2 *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).

3.3 *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 grounds that the distribution of mean active speech levels (dB values) is approximately Gaussian.

4 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/P.56.

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.

H.T. [T1.56]

TABLE 1/P.56

Application	Meter	Quantity observed
{ Control of vocal level in live-speech loudness balances }	ARAEN volume meter (SV3)	Level exceeded in 3 s
Avoidance of peak limiting	Peak programme meter	Highest reading
{ Maintenance of optimum level in making magnetic tape recording }	VU meter	{
Average of peaks (excluding most extreme) }		

Table 1/P.56 [T1.56], p.

5 Method B: active speech level for other applications than those mentioned in method A

5.1 *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 . However, 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 [4]. 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 the 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).

This Recommendation describes the detection method that meets the requirements. 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; despite that, by accumulating sufficient information during the process, it is possible to apply the correct threshold retrospectively, 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.

5.2 *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 the 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 where the probability of exceeding that value is 0.001.

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 the sample value; that is to say, 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 $i = 1, 2, 3, \ldots$ 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

$g = \exp(-t/T)$ coefficient of smoothing

H hangover time in seconds

$I = H/t$ rounded up to next integer

M margin in 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:

$$p_i = g \cdot p_{i-1} + (1-g) \cdot |x_i|$$

$$q_i = g \cdot q_{i-1} + (1-g) \cdot p_i$$

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 j th 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.

Initially, all the a_j values are set equal to zero, and the h_j values set equal to I .

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 a significant d.c. offset, this may be estimated as s | (mu | fIv /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 is equal to sq | (mu | fIv ^{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 (sq | (mu | fIv ^{2}/n)—20 log r

Active-level estimate, A_j = 10 log (sq | (mu | fIv ^{2}/a_j) —20 log r

Threshold, C_j = 20 log (c_j | (mu | fIv)—20 log 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 fulfils 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 as fulfilling the definition.

5.3 *Values of the parameters*

The values of the parameters given in Table 2/P.56 should be used. They have been found suitable for the purpose and have stood the test of many years of application by various organizations [4].

H.T. [T2.56]

TABLE 2/P.56

Parameter	Value	Tolerance
f	694 samples/second	not less than 600
T	0.03 seconds	± | % |
H	0.2 seconds	± | % |
M	15.9 dB	{
± | .5		
}		

Note — The value M = 15 dB might appear to be implied in [4], but the threshold level there described 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 (voltage/mean absolute voltage) for a sine wave.

Table 2/P.56 [T2.56], p.

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 Recommendation.

Note — 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 a 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.

6 Approximate equivalents of method B

Other methods under development use a broadly similar principle of measurement but depart in detail from the algorithm given above.

It is not the intention to exclude any such method, 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 s (active time)

Bandwidth 300 to 3400 Hz and 100 to 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.

This verification procedure is valid until a suitable speech-like signal has been recommended and found suitable to perform this function (see Questions 12/XII and 13/XII).

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 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* .

Certain measurement systems with fixed thresholds (instead of the retrospectively selected threshold as described in § 5.3), may still give an active speech level according to the definition in cases where the margin turns out to be within the specified limits.

7 Specification

A speech voltmeter normally consists of three parts, namely:

i) input circuitry,

ii) filter, and

iii) processor and display.

Figure 1/P.56 shows a typical layout of such a meter.

Whether all or part of the components that make up i) and ii) are used will depend on where the meter is to be used. However, it is recommended that a meter for general usage should conform to this specification.

Figure 1/P.56, p.

7.1 *Signal input*

7.1.1 *Input impedance*

The meter is normally used as a bridging instrument and, if so, its impedance must be high so as not to influence the results. An impedance of 100 kohm is recommended.

7.1.2 *Circuit protection*

It is recommended that the meter should withstand voltages far in excess of those in the measurement range as accidental usage may occur and the circuit under test may have higher voltages than anticipated. Examples of this are mains 110/240 V or 50 V exchange voltages.

7.1.3 *Connection*

It is recommended that the connection should be independent of polarity. The meter should have the facility of connection in both balanced and unbalanced modes.

7.2 *Filter*

When measuring the speech levels of circuits in the conventional telephony speech bandwidth (300-3400 Hz), it is often practical to use a filter that will reject unwanted hum, tape noise, etc. yet pass the frequencies of greatest interest without affecting the speech level measurement. The set of coordinates in Table 3/P.56 meet these requirements. Figure 2/P.56 gives an example of such a filter.

The following noise requirements should also be met:

Output noise level:

wideband (20-20 | 00 Hz) <—75 dBm

telephone weighted <—90 dBmp.

H.T. [T3.56]

TABLE 3/P.56

Frequency (Hz)	(dB)
	{
Upper limit response relative to 1 kHz	
}	
| 16	—49.75
| 60	+0.25
7 | 00	+0.25
70 | 00	—49.75
	{
Lower limit response relative to 1 kHz	
}	
Under 200	—∞
200	—0.25
5500	—0.25
Over 5500	—∞

Tableau 3/P.56 [T3.56], p. 32

Figure 2/P.56, p. 33

7.3 *Speech level measurements*

7.3.1 *Working range for speech*

The recommended working range for speech refers to the active level and should be at least 0 to —30 dBV.

Note 1 — The dynamic range of the instrument will depend on the analogue-to-digital converter (ADC). If the ADC is set to a 10 volt maximum input level (i.e. the all 1 code) and 12-bit arithmetic is used, based on the most significant bits from the ADC, then 1 sign bit +11 bits magnitude provides a 66 dB range. The measurable range sill be some 35 dB less when allowance is made for the peak/mean ratio of 18 dB (peaks of speech will only exceed the maximum input level for less than 0.1% of the time [1]) and margin *M* of 15.9 dB; the largest speech signal is therefore around +2 dBV with a smallest speech signal of —30 dBV. However, the practical working range has been found to be +5 dBV to —35 dBV.

Note 2 — To cater for a wider range of speech levels, an attenuator or low noise amplifier may be inserted in the input circuitry. Care must be exercised to maintain the input requirements of § 7.1.1.

7.3.2 *Linearity*

The linearity of the meter is specified for r.m.s. sine wave measurements since for speech the algorithm is correct by definition, and only the precision or repeatability of measurements need to be considered; this is specified in § 7.3.4.

Assuming that:

a) the measurement is for a minimum period of 5 s,

b) the sine wave is present for the whole of the measurement period, the linearity specified is:

Frequency (Hz)	*Input range* (dBV)	*Accuracy* (dB)
100 to 4000	+16 to —45	± 0.1
4000 to 8000	+13 to —45	± 0.3

Note — The maximum input for the frequency range 4000 to 8000 Hz should ideally be the same as for 100 to 4000 Hz, but practical limitations in commercially available ADCs (due to the limited ‘‘ slewing rate ´as the power in the 8000 Hz band for speech is 30 dB down on the level at 500 Hz it is likely that any error will be extremely small.

7.3.3 *Frequency response*

The frequency response of the meter without filter when measured in the frequency range 100 to 8000 Hz should be flat within the specified tolerances:

Frequency (Hz)	*Input range* (dBV)	*Tolerance* (dB)
100 to 4000	+16 to —45	± 0.2
4000 to 8000	+13 to —45	± 0.4

Note 1 — Tolerances are referred to 1000 Hz.

Note 2 — The note of 7.3.2 applies.

7.3.4 *Repeatability*

When a given speech signal, having its active level within the recommended working range and its duration not less than 5 s active time, is repeatedly measured on the same meter, the active-level readings shall have a standard deviation of less than 0.1 dB.

8 Routine calibration of method-B meter

The following routine calibration procedures, using non-speech-like signals, will ensure that the meter is performing satisfactorily. The calibration can only be made using speech.

A suitable circuit arrangement is shown in Figure 3/P.56. Wherever suitable, measurements should be made with two settings of the attenuator, 0 and 20 dB. All source signals are from a 600 ohm source and the meter is terminated in 600 ohm.

Figure 3/P.56, p.

8.1 *No input signal*

With no input applied the meter should display the following results:

Activity factor 0 + 0.5% Active-level < —60 dBV Long-term level < —60 dBV 8.2 *Continuous tone*

With a 1000 Hz sine wave calibrated to be 0 dBV, the meter should display the following results for the two settings of the attenuator when applied for 12 + 0.2 s:

Attenuator = 0 dB Attenuator = 20 dB Activity factor 100 to 0.5% 100 to 0.5% Active-level 0 ± 0.1 dBV —20 ± 0.1 dBV Long-term level 0 ± 0.1 dBV —20 ± 0.1 dBV

8.3 *White noise*

8.3.1 *Without filter*

With the meter having no filter in circuit and the white noise source calibrated to be 0 dBV, the meter should display the following results for the two settings of the attenuator when applied for 12 + 0.2 s:

Attenuator = 0 dB Attenuator = 20 dB Activity factor 100 to 0.5% 100 to 0.5% Active-level 0 ± 0.5 dBV —20 ± 0.5 dBV Long-term level 0 ± 0.5 dBV —20 ± 0.5 dBV

8.3.2 *With filter*

With the meter having the filter in circuit and the white noise source calibrated to be 0 dBV, the meter should display the following results for the two settings of the attenuator when applied for 12 + 0.2 s:

Attenuator = 0 dB Attenuator = 20 dB Activity factor 100 to 0.5% 100 to 0.5% Active-level —6.9 ± 0.5 dBV —26.9 ± 0.5 dBV Long-term level —6.9 ± 0.5 dBV —26.9 ± 0.5 dBV

8.3.3 *Pulsed noise*

With the meter having no filter in circuit and the white noise source pulsed at 3 s ‘‘ON’’ and 3 s ‘‘OFF’’ and calibrated to be 0 dBV when ‘‘ON’’, the meter should display the following results for the two settings of the attenuator when applied for 12 + 0.2 s:

Attenuator = 0 dB Attenuator = 20 dB Factor activity 55 ± 1.5% 55 ± 1.5% Active-level 0 ± 1 dBV —20 ± 1 dBV Long-term level —2.7 ± 1 dBV —22.7 ± 1 dBV

Note — It is possible that § 8 could be revised to calibrate both method B and B-equivalent meters when a speech-like signal has been found suitable to perform this function.

ANNEX A

(to Recommendation P.56)

A method using a speech voltmeter complying

with method B in network conditions

A speech voltmeter complying with method B is not suitable in its present form for speech measurements (see, for example, Recommendation G.223) on real connections since the meter is unable to distinquish between speech coming from one or the other end of the connection.

However, if the meter is connected to a 4-wire point in a connection of the type 2-4-2 wire, then measurements may be made using an operator monitoring the beginning and the end of the conversation. The operator can perform this function using earphones (provided the subscriber’s permission has been obtained) or by an auxiliary meter (for example conforming to P.52). The circuit arrangement is shown in Figure A-1/P.56.

The operator monitors the conversation, using the auxiliary meter or earphones, and then by means of a start/stop button can measure the beginning and end of the relevant conversation.

Figure A-1/P.56, p.

References

[1] RICHARDS (D. | .): Telecommunication by speech, § 2.1.3.2, pp. 56-69, *Butterworks* , London, 1973.

[2] ITU — *List of Definitions of Essential Telecommunication Terms* , Definition 14.16, Second impression, Geneva, 1961.

[3] ITU — *List of Definitions of Essential Telecommunication Terms* , Definitions 12.34, 12.35, 12.36, Second impression, Geneva, 1961.

[4] BERRY (R. | .): Speech-volume measurements on telephone circuits, *Proc. IEE* , Vol. 118, No. 2, pp. 335-338, February 1971.

Bibliography

BRADY (P. | .): Equivalent Peak Level: a thre shold-independent speech level measure, *Journal of the Acoustical Society of America* , Vol. 44, pp. 695-699, 1968.

CARSON (R.): A digital Speech Voltmeter — the S V6, *British Telecommunications Engineering* , Vol. 3, Part 1, pp. 23-30, April 1984.

CCITT — Contribution COM XII-No. 43 *A method for sp eech-level measurements using IEC-interface bus and calculation* (Norway), Geneva, 1982.

Blanc

MONTAGE: PAGE 122 = BLANCHE