INTERNATIONAL TELECOMMUNICATION UNION

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

P.50
(09/99)

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

Objective measuring apparatus

Artificial voices

ITU-T Recommendation P.50

(Previously CCITT Recommendation)

ITU-T P-SERIES RECOMMENDATIONS

TELEPHONE TRANSMISSION QUALITY, TELEPHONE INSTALLATIONS, LOCAL LINE NETWORKS

Vocabulary and effects of transmission parameters on customer opinion of transmission quality	Series P.10
Subscribers' lines and sets	Series P.30
	P.300
Transmission standards	Series P.40
Objective measuring apparatus	**Series P.50**
	P.500
Objective electro-acoustical measurements	Series P.60
Measurements related to speech loudness	Series P.70
Methods for objective and subjective assessment of quality	Series P.80
	P.800
Audiovisual quality in multimedia services	Series P.900

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

ITU-T RECOMMENDATION P.50

ARTIFICIAL VOICES

Summary

The "artificial voice" described in this Recommendation reproduces the characteristics of human speech for the purpose of characterizing linear and non-linear telecommunication systems and devices, which are intended for the transduction or transmission of speech. 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. Two kinds of artificial voice are defined, reproducing respectively the characteristics of female and male speech.

FOREWORD

ITU (International Telecommunication Union) is the United Nations Specialized Agency in the field of telecommunications. The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of the ITU. The ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Conference (WTSC), which meets every four years, establishes the topics for study by the ITU-T Study Groups which, in their turn, produce Recommendations on these topics.

The approval of Recommendations by the Members of the ITU-T is covered by the procedure laid down in WTSC Resolution No. 1.

In some areas of information technology which fall within ITU-T's purview, the necessary standards are prepared on a collaborative basis with ISO and IEC.

NOTE

In this Recommendation the term *recognized operating agency (ROA)* includes any individual, company, corporation or governmental organization that operates a public correspondence service. The terms *Administration, ROA* and *public correspondence* are defined in the *Constitution of the ITU (Geneva, 1992)*.

INTELLECTUAL PROPERTY RIGHTS

The ITU draws attention to the possibility that the practice or implementation of this Recommendation may involve the use of a claimed Intellectual Property Right. The ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU members or others outside of the Recommendation development process.

As of the date of approval of this Recommendation, the ITU had not received notice of intellectual property, protected by patents, which may be required to implement this Recommendation. However, implementors are cautioned that this may not represent the latest information and are therefore strongly urged to consult the TSB patent database.

CONTENTS

Recommendation P.50

ARTIFICIAL VOICES

(Melbourne, 1988; amended at Helsinki, 1993, Geneva, 1999)

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. 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;

b)   short-term spectrum;

c)   instantaneous amplitude distribution;

d)   voiced and unvoiced structure of speech waveform;

e)   syllabic envelope.

Appendix I/P.50 includes a CD-ROM containing useful test signals. The signals on this CD-ROM include the signal described in Recommendation P.50 as well as other signals that have been found useful by some Administrations. Additionally, the full speech database that was used to develop Recommendation P.50 is also on this CD-ROM. Appendix I/P.50 is published separately.

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 artificial voice described in this Recommendation is mainly used for objective evaluation of speech processing systems and devices, in which a single-channel signal with continuous activity (i.e. without pauses) is sufficient for measuring characteristics. An example is evaluation of speech codecs. For objective evaluation that needs two signals with pauses (e.g. evaluation of devices with speech detectors), the artificial conversational speech signal described in Recommendation P.59 should be used.

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.

1   Electrical artificial voice
2   Artificial mouth excitation signal
3   Acoustic artificial voice
MRP   Mouth Reference Point
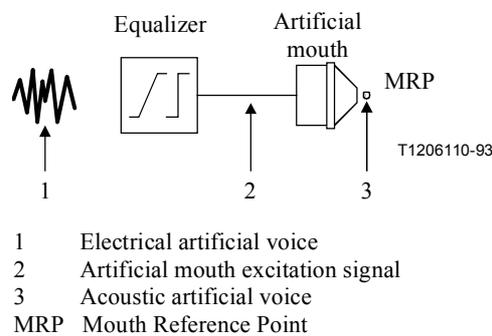
Figure 1/P.50

3.1 Electrical artificial voice

The artificial voice produced as an electrical signal 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 – 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

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.

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 and Table 1, normalized for a wideband sound pressure level of −4.7 dBPa.

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

$$S(f) = -376.44 + 465.439(\log_{10}f) - 157.745(\log_{10}f)^2 + 16.7124(\log_{10}f)^3 \qquad (4\text{-}1)$$

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

The total sound pressure level of the spectrum defined in Equation (4-1) 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 (4-1) may range from −391.44 to −361.44.
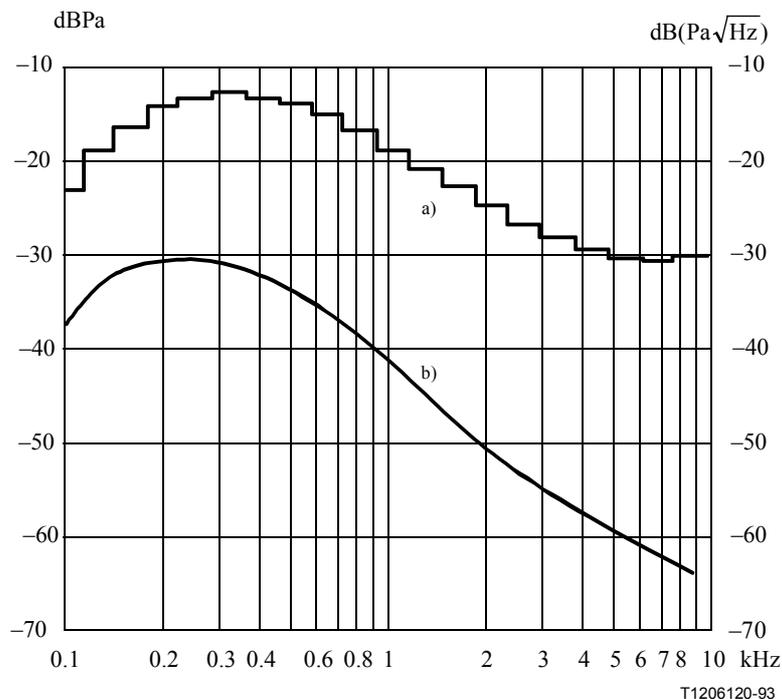
a) Third octave spectrum [Column (3), Table 1].

b) Spectrum density [Column (3)-(2), Table 1].

Figure 2/P.50 – Long-term spectrum of artificial voice

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.

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. 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).

Table 1/P.50 – 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) (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[a]
160	15.6	−16.4	−32.7	+3, −6[a]
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.

4.5 Spectrum of the modulation envelope

The spectrum of the modulation envelope waveform is shown in Figure 5 and should be reproduced with a tolerance of ±5 dB on the whole frequency range.
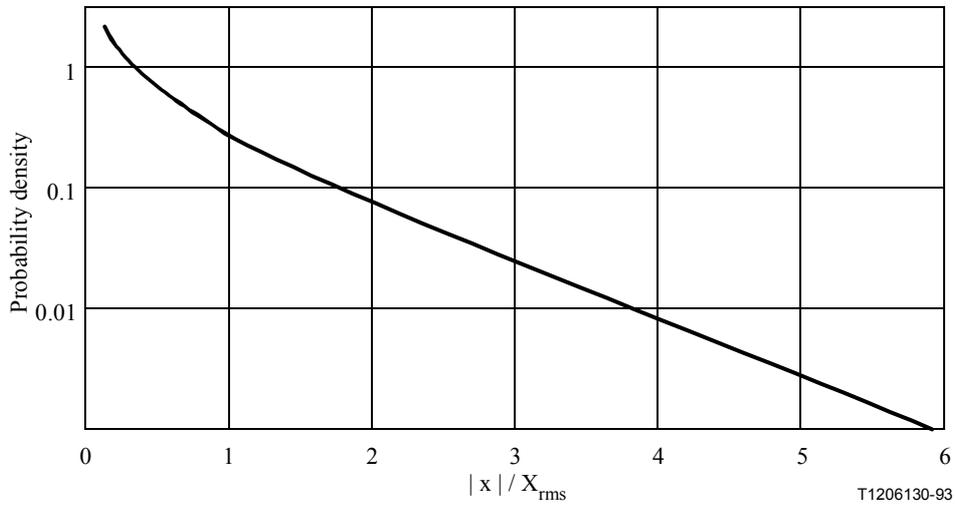
$|x|$ Absolute value of the instantaneous amplitude

X_{rms} Root mean square of the signal

Figure 3/P.50 – Instantaneous amplitude distribution
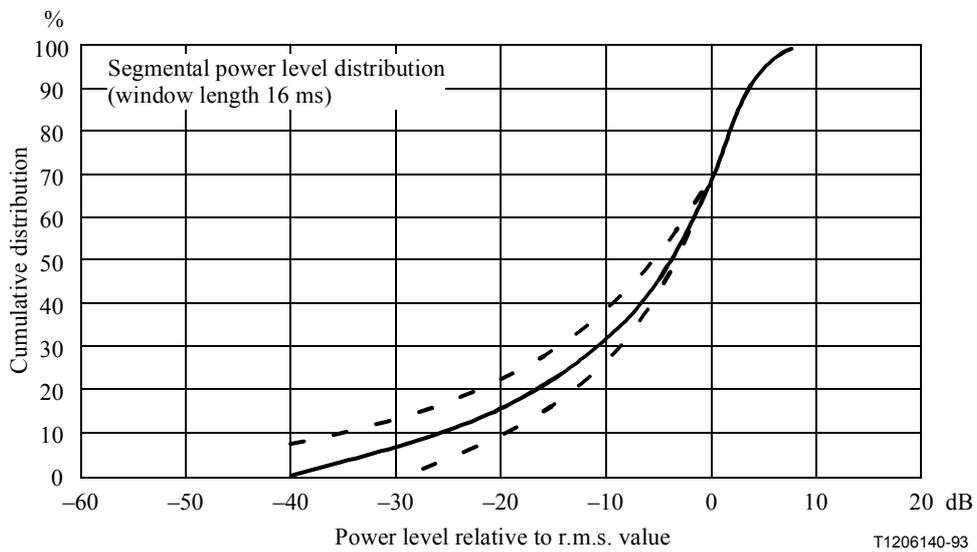
Figure 4/P.50 – Segmental power level distribution
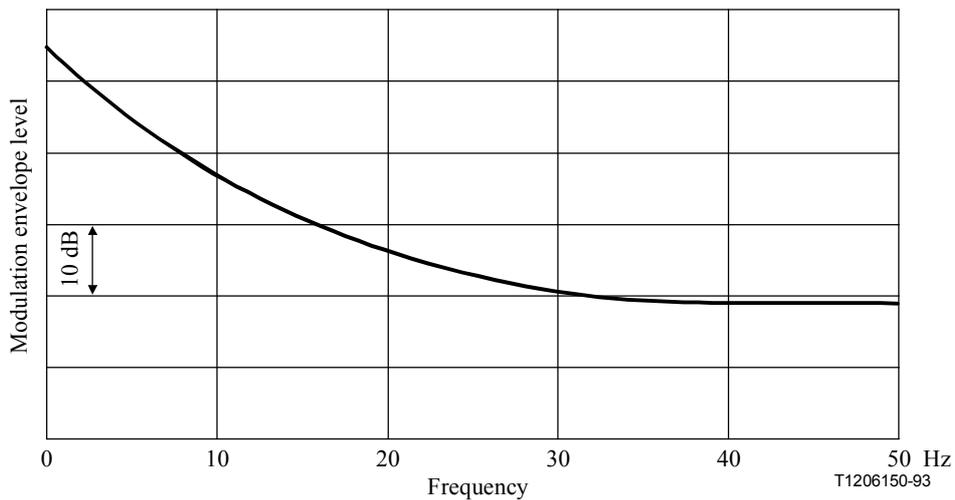
Figure 5/P.50 – Spectrum of modulation envelope

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 shows a block diagram of the generation process of the artificial voice. It is generated by applying two different types of excitation source 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.
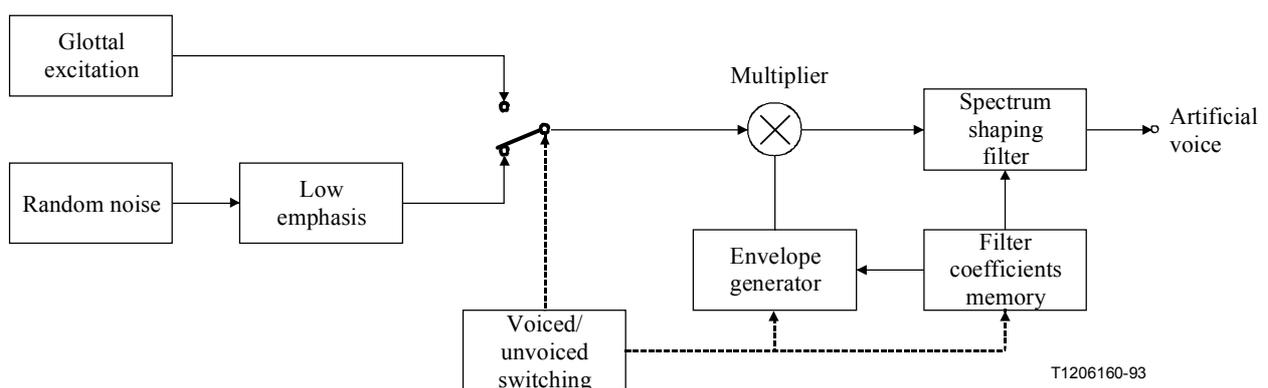
Figure 6/P.50 – Artificial voice generation process

5.1 Excitation source signal

The artificial 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). 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:

$$T = -3.486 \ln(r)$$

where r denotes a uniformly distributed random number ($0.371 \leq r \leq 0.609$).

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 ($3\ T_v/4$)

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

Element d: Voiced ($3\ T_v/4$) + 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 clause 4.

5.2 Glottal excitation

The glottal excitation signal is a periodic waveform as shown in Figure 7. The pitch frequency ($1/T_0$ in Figure 7) varies according to the variation pattern shown in Figure 8 during the period T_v. The starting value of the pitch frequency (F_s in Figure 8) 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 ($0 < R < 1$). 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). For the elements b), c) and d) in Figure 9 the pitch frequency variation pattern applies to the combination of the two voiced parts, irrespectively of where the unvoiced segment is inserted.

5.3 Unvoiced sounds

The transfer function of the low-pass filter located after the random noise generator (low emphasis) is $1/(1 - (1 - \varepsilon)\ z^{-1})$, $\varepsilon = 0.001$ where z^{-1} denotes the unit delay. The random noise generator generates Gaussian white noise with a discrete zero-mean random process whose samples are uncorrelated.

a = 0.375 ms
b = 0.5 ms
c = e = 0.3125 ms
d = 0.1875 ms
f = 0.625 ms
DC = 1.176/(16000 T_0 – 23) for male
       1.176/(16000 T_0 – 11.5) for female
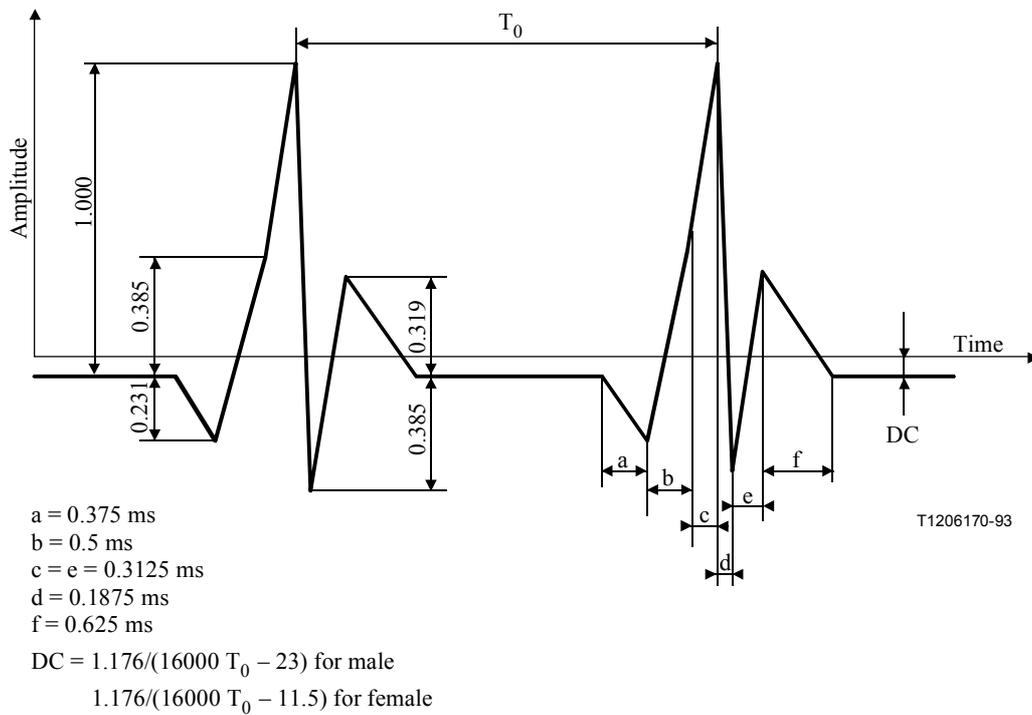
Figure 7/P.50 – Glottal excitation signal

T1206180-93
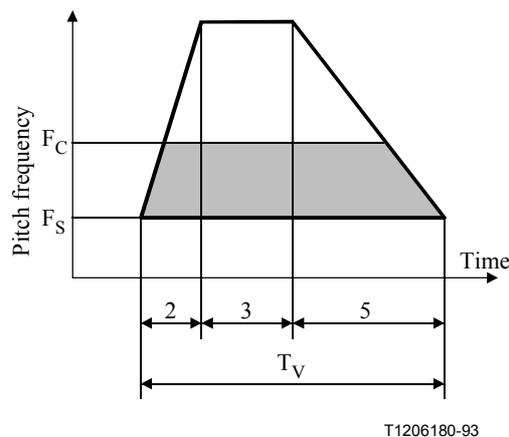
Figure 8/P.50 – Pitch frequency variation pattern

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 Figure 9. This is obtained by utilizing the following relationship providing input and output signals of the spectrum shaping filter:

$$P_{in} = P_{out} \prod_{i=1}^{12} \left(1 - k_i^2\right)$$

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 ith coefficient of the spectrum shaping filter

The rising, stationary and decay times of each trapezoid of Figure 9 shall be mutually related by the same proportionality coefficients (2:3:5) of the pitch frequency variation pattern shown in Figure 8. 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. 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 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. A sampling frequency of 16 000 Hz belongs to Table 2a and Table 2b.

a)

b)

$P_{vb1} = 12\ P_v/13$
$P_{vb2} = 40\ P_v/39$

c)

$P_{vc1} = 16\ P_v/13$
$P_{vc2} = 10\ P_v/13$

d)

$P_{vd1} = 12\ P_v/39$
$P_{vd2} = 40\ P_v/13$
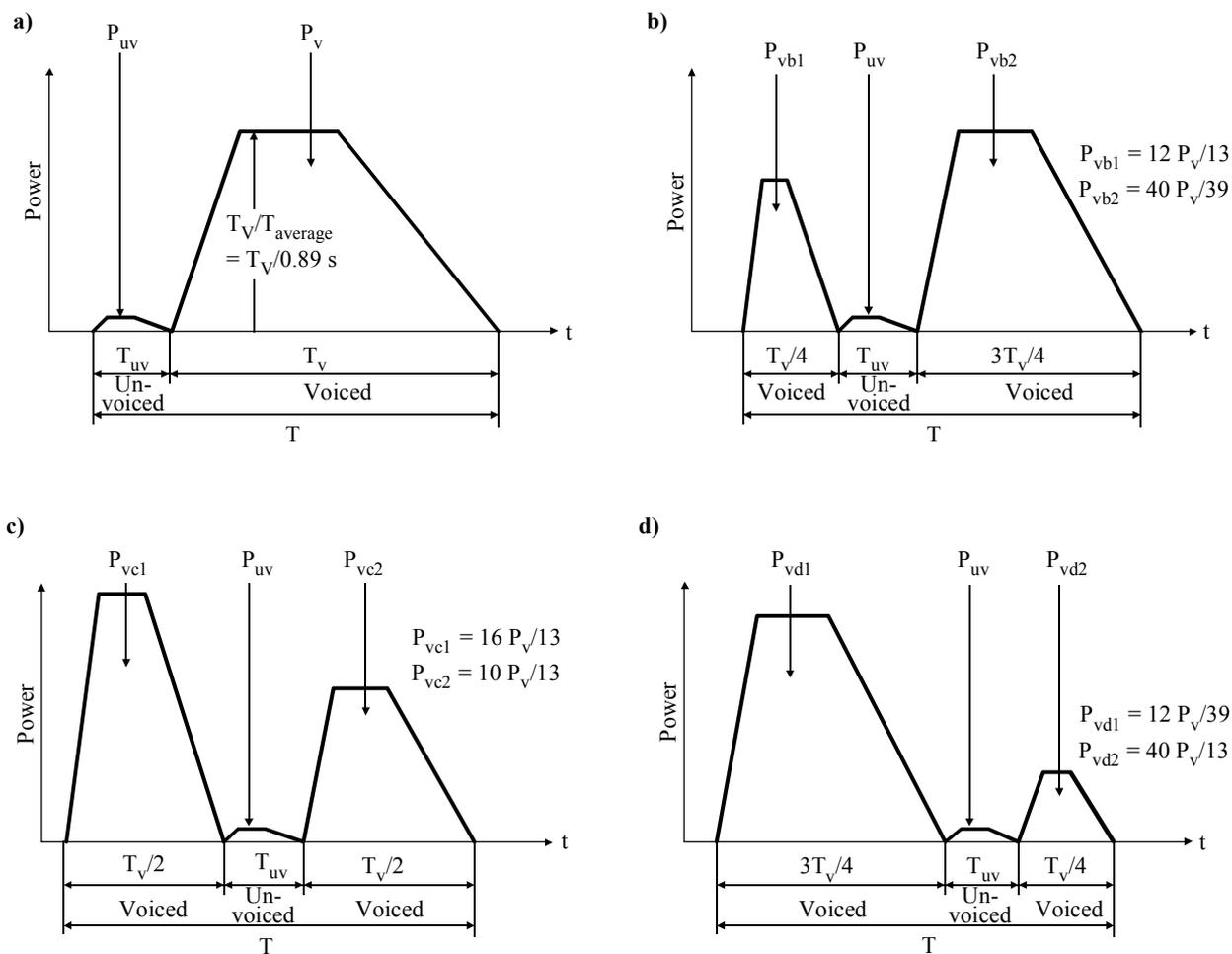
T1206190-93

**Figure 9/P.50 – Short-term power variation patterns of the four unit elements
used to generate the artificial voice**

z^{-1} Unit delay
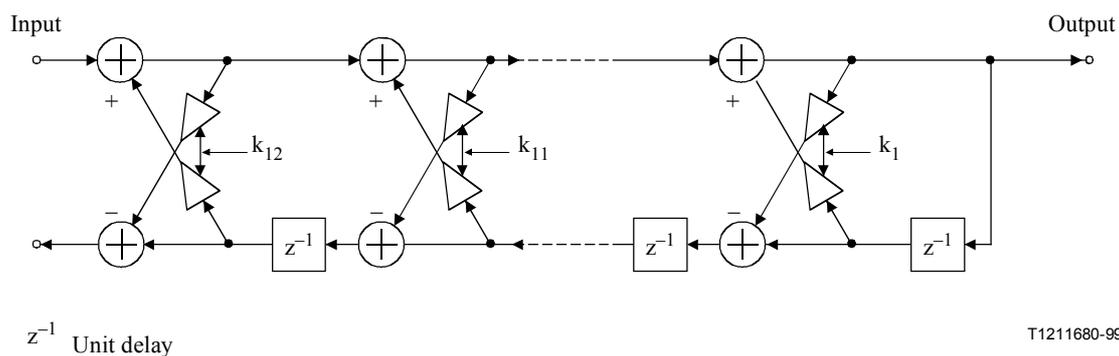
T1211680-99

Figure 10/P.50 – Spectrum shaping filter

NOTE – The output port of the filter at the lower side in Figure 10 is not used but shown to highlight the symmetry of the filter structure.

Table 2a/P.50 – Coefficients k_i for male artificial voice

		k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8	k_9	k_{10}	k_{11}	k_{12}
Unvoiced	1	−0.471	−0.108	0.024	−0.048	0.140	0.036	0.054	0.004	0.123	0.044	0.099	−0.003
	2	−0.284	−0.468	0.030	0.090	0.124	−0.020	0.087	0.067	0.131	0.011	0.076	−0.024
	3	−0.025	−0.496	−0.176	0.162	0.236	−0.012	0.068	0.001	0.096	0.029	0.086	−0.018
Voiced	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
	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

Table 2b/P.50 – Coefficients k_i for female artificial voice

		k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8	k_9	k_{10}	k_{11}	k_{12}
Unvoiced	1	0.488	−0.388	0.145	0.053	0.122	0.027	0.135	0.035	0.080	0.017	0.068	0.028
	2	−0.093	−0.444	−0.102	0.121	0.154	0.009	0.102	−0.031	0.084	0.019	0.101	−0.020
	3	−0.709	−0.179	0.134	0.007	0.142	0.027	0.099	0.000	0.115	0.007	0.075	−0.037
Voiced	1	0.355	−0.247	−0.092	−0.043	0.032	0.046	0.113	−0.023	0.071	−0.030	−0.000	−0.116
	2	0.976	0.150	−0.062	−0.187	−0.172	−0.200	−0.122	−0.207	−0.054	−0.127	0.012	−0.111
	3	0.737	−0.324	−0.175	−0.197	−0.153	0.023	0.110	−0.018	0.040	−0.062	0.034	−0.091
	4	0.598	0.234	0.126	0.011	−0.005	−0.026	0.131	0.032	0.073	−0.063	0.011	−0.088
	5	0.808	0.118	0.262	0.139	0.063	−0.024	0.001	−0.184	−0.056	−0.100	0.014	−0.115
	6	0.914	−0.500	−0.051	−0.115	−0.211	−0.012	−0.077	−0.179	0.064	−0.102	0.037	−0.092
	7	0.933	−0.359	0.089	−0.107	−0.178	−0.050	−0.137	−0.206	0.046	−0.088	−0.004	−0.074
	8	0.966	−0.023	0.044	−0.105	−0.178	−0.195	−0.150	−0.233	−0.045	−0.092	0.029	−0.097
	9	0.870	−0.469	−0.244	−0.107	−0.140	−0.037	0.084	−0.131	0.021	−0.066	−0.003	−0.091
	10	0.673	−0.292	0.392	0.158	0.143	0.160	0.019	−0.281	−0.105	−0.195	−0.156	−0.185
	11	0.962	−0.191	0.030	−0.089	−0.207	−0.133	−0.141	−0.263	0.007	−0.054	0.014	−0.074
	12	0.879	−0.340	0.046	−0.049	−0.071	−0.024	−0.039	−0.188	0.017	−0.078	−0.014	−0.117
	13	0.941	−0.258	0.122	−0.073	−0.163	−0.089	−0.151	−0.250	0.025	−0.062	−0.006	−0.093

6   Bibliography

[1]   CCITT Contribution COM XII-No. 76 (1982), An artificial voice (Italy).

[2]   CCITT Contribution COM XII-No. 108 (1982), Draft partial text of section 3, Recommendation P.51 (Special Rapporteur).

[3]   CCITT Contribution COM XII-No. 11 (1981), Measurement of carbon microphone using real voice and artificial test signals (Canada).

[4]   CCITT Contribution COM XII-No. 150 (1983), Improvement of artificial voice by inserting fricative sounds (Italy).

[5]   CCITT Contribution COM XII-No. 132 (1983), Evaluation of non-linear distortion using simulated speech signals (Canada).

ANNEX A

Short-term spectrum characteristics of the artificial voice

The artificial voice is generated by randomly selecting each of sixteen short-term spectrum patterns once every 960 ms (= 60 ms × 16 patterns). The spectrum density of each pattern is provided by Equation (A-1) and Table A.1, 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-1) and Table A.1 are expressed in power normalized form.

$$\text{Spectrum density } S_i(f) = \frac{1}{A_{io} + 2\sum_{j=1}^{12} A_{ij}\left[\cos(2\pi ijf)\right]}, i = 1, 2, \ldots 16 \qquad \text{(A-1)}$$

Table A.1a/P.50 – Coefficients A_{ij} for male artificial voice

i	j												
	0	1	2	3	4	5	6	7	8	9	10	11	12
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

Table A.1b/P.50 – Coefficients A_{ij} for female artificial voice

i	j												
	0	1	2	3	4	5	6	7	8	9	10	11	12
1	8.92953	–8.28905	7.23150	–6.06571	5.06663	–4.16883	3.34820	–2.64174	1.91152	–1.27122	0.74358	–0.35347	0.13100
2	9.11050	–8.29868	7.05018	–6.03862	5.02156	–4.15784	3.37442	–2.70084	2.04257	–1.41928	0.90339	–0.47240	0.14100
3	13.69058	–12.75539	10.87390	–8.98976	7.28838	–5.87662	4.78110	–3.72852	2.71831	–1.81828	1.06757	–0.48907	0.12900
4	6.95118	–6.29134	5.35757	–4.64910	3.87250	–3.23132	2.69856	–2.12367	1.58326	–1.13885	0.70609	–0.37778	0.12800
5	2.73454	–1.80664	1.95283	–1.79464	1.43897	–1.31656	0.93268	–0.87398	0.53694	–0.47562	0.24159	–0.15438	0.11800
6	10.82358	–10.07808	8.78565	–7.43643	6.14765	–5.10041	4.10027	–3.23241	2.34220	–1.54676	0.91918	–0.45059	0.14200
7	21.58481	–19.92676	16.21532	–12.43566	9.61057	–7.93982	6.97363	–6.00866	4.68271	–3.12797	1.70804	–0.67499	0.14200
8	23.73912	–22.20897	18.74416	–15.03715	11.99248	–9.85513	8.27112	–6.72826	4.94335	–3.10450	1.60004	–0.61090	0.12900
9	4.97162	–4.27705	4.01380	–3.38500	2.78457	–2.45010	1.98057	–1.63020	1.18104	–0.80108	0.51528	–0.29138	0.12500
10	13.37598	–12.45509	10.72295	–8.97928	7.35893	–6.05438	4.88819	–3.86108	2.85164	–1.88876	1.11490	–0.52260	0.13800
11	16.48817	–15.22287	12.62125	–10.23900	8.46966	–7.23692	6.24648	–5.09752	3.77465	–2.46950	1.37404	–0.57453	0.13200
12	18.22041	–17.17540	15.09489	–12.55171	10.24976	–8.45903	6.71874	–5.19063	3.52021	–2.10167	1.08066	–0.41880	0.14300
13	1.32602	–0.31718	0.44277	–0.47070	0.32935	–0.41555	0.25775	–0.32079	0.13791	–0.23640	0.10103	–0.10136	0.11500
14	16.90640	–15.73723	13.30151	–10.82887	8.78690	–7.34521	6.21516	–5.11100	3.80281	–2.43990	1.33506	–0.55971	0.12900
15	21.73895	–20.42432	17.51117	14.44152	11.79131	–9.66735	7.90433	–6.19508	4.41275	–2.75545	1.46525	–0.59916	0.14000
16	21.04832	–19.72714	16.81197	–13.70183	11.07189	–9.12707	7.57941	–6.08064	4.40471	–2.74320	1.43897	–0.58079	0.13300

ITU-T RECOMMENDATIONS SERIES

Series A   Organization of the work of the ITU-T

Series B   Means of expression: definitions, symbols, classification

Series C   General telecommunication statistics

Series D   General tariff principles

Series E   Overall network operation, telephone service, service operation and human factors

Series F   Non-telephone telecommunication services

Series G   Transmission systems and media, digital systems and networks

Series H   Audiovisual and multimedia systems

Series I   Integrated services digital network

Series J   Transmission of television, sound programme and other multimedia signals

Series K   Protection against interference

Series L   Construction, installation and protection of cables and other elements of outside plant

Series M   TMN and network maintenance: international transmission systems, telephone circuits, telegraphy, facsimile and leased circuits

Series N   Maintenance: international sound programme and television transmission circuits

Series O   Specifications of measuring equipment

Series P   Telephone transmission quality, telephone installations, local line networks

Series Q   Switching and signalling

Series R   Telegraph transmission

Series S   Telegraph services terminal equipment

Series T   Terminals for telematic services

Series U   Telegraph switching

Series V   Data communication over the telephone network

Series X   Data networks and open system communications

Series Y   Global information infrastructure and Internet protocol aspects

Series Z   Languages and general software aspects for telecommunication systems