3k. Speech and Hearing

EDWIN B. NEWMAN[1]

Harvard University

The data concerning hearing are, without exception, empirical in derivation. Consequently, the values reported always represent some parameter of a population, most often a mean, and the reader is warned to bear constantly in mind the many sources of variability that attach to any particular measurement.

3k-1. Physical Dimensions of the Ear

TABLE 3k-1. PHYSICAL DIMENSIONS OF THE EAR[*]

- Pinna:
  - Mean length, young men, 65.0 mm
  - Range, 52–79 mm
- Auditory meatus:
  - Cross section, 0.3–0.5 cm^2
  - Diameter, 0.7 cm
  - Length, 2.7 cm
  - Volume, 1.0 cc
- Tympanic membrane:
  - Area, 0.5–0.9 cm^2 (roughly circular)
  - Thickness, about 0.1 mm
  - Volume elasticity for 10 Hz, equivalent to about 8 cc air
  - Displacement amplitude for 1,000-Hz tone (at threshold), 10^{-9} cm
  - Displacement amplitude for low-frequency tones (threshold of feeling), about 10^{-2} cm
- Middle ear:
  - Total volume, about 2 cc
  - Malleus:
    - Weight, 23 mg
    - Length, 5.5–6.0 mm
  - Incus: weight, 27 mg
  - Stapes:
    - Weight, 215 mg
    - Length of footplate, 3.2 mm
    - Width of footplate, 1.4 mm
    - Area of footplate, 3.2 mm^2
    - Width of elastic ligament, 0.015–0.1 mm
- Cochlea:
  - Length of cochlear channels, 35 mm
  - Height of scala vestibuli or scala tympani, about 1 mm (great variability)
  - Round window: area, 2 mm^2
  - Basilar membrane:
    - Width at stapes, 0.04 mm
    - Width at helicotrema, 0.5 mm
  - Helicotrema: area of opening, 0.25–0.4 mm^2

[*] S. S. Stevens, ed., "Handbook of Experimental Psychology," John Wiley & Sons, Inc., New York, 1951.

3k-2. Acoustic Impedance of the Ear. Reasonable agreement on measurements below 1,000 Hz has been obtained. The reference point for measurements is just

---

[1] This section benefited from the advice and assistance of Dr. S. S. Stevens and Mrs. Nancy C. Waugh.

within the external meatus. The values in Table 3k-2 are representative but are subject to wide variations among individuals.[1]

3k-3. Minimum Audible Sound. Table 3k-3 lists the minimum audible (threshold) sound pressures of pure tones measured at the entrance to the external meatus. The pressure measurements were made when the subject heard the tone one-half the time it was presented via an earphone applied to his ear with a standard static force. Observations were made on young persons, eighteen to twenty-five years of age, with no record of hearing impairment. Sound pressures were determined with a probe-tube microphone and are given in decibels relative to 2×10^{-4} dyne/cm^2.

The results of such measurements made in various laboratories show a considerable amount of variation. The pressures in Table 3k-3 are based on measurements made in two independent laboratories; see the first footnote for details.

The variance in the threshold sound pressures measured at the entrance to the meatus has been so great that such pressures cannot serve usefully as standards for audiometry. Experience has shown that the most accurate method for storing audiometric standard threshold information is as follows. Measurements of threshold voltages on an earphone applied to a number of young persons at the various audiometric frequencies are the primary data. The sound pressures which are produced by these voltages when the earphone is applied to an artificial ear (coupler) then serve as the standard thresholds for that particular earphone-coupler combination. This method of measuring and storing standard threshold sound pressures is now in use in several countries. A comparison of the standard thresholds was completed under the auspices of Technical Committee 43 on Acoustics of the International Organization for Standardization (ISO). An internationally agreed-upon standard threshold has been issued by ISO in its Recommendation R389, Standard Reference Zero for the Calibration of Pure Tone Audiometers. The standard data in it are sound pressures corresponding to the threshold of hearing for five earphone-coupler combinations now in use in several countries.[2]

3k-4. Threshold of Feeling or Discomfort. The upper limit for a tolerable intensity of sound rises substantially with increasing habituation. Moreover, a variety of subjective effects are reported, such as discomfort, tickle, pressure, and pain, each at a slightly different level. As a simple engineering estimate it can be said that naïve listeners reach a limit at about 125 dB SPL and experienced listeners at 135 to 140 dB. These are overall measures of sound falling within the audible range and are roughly independent of frequency.

3k-5. Differential Thresholds for Pure Tones and Noise. A differential threshold represents a careful determination by laboratory methods of the ability of a subject to just detect, and report, a difference in any specific property of a sound, all other factors presumably being held constant.

The method for determining the differential threshold for intensity of pure tones employed one tone beating with a second tone at 3 beats per second.[3] Much evidence is available to support what should be kept always in mind, that thresholds determined by other methods are a function of numerous psychological parameters and will differ systematically from the values in Table 3k-4. A more conventional method was used to determine the thresholds for white noise, with the results given in the last column.[4]

---

[1] E. Waetzmann and L. Keibs, Hörschwellenbestimmungen mit dem Thermophon und Messungen am Trommelfell, *Ann. Physik* **26**, 141–144 (1936); O. Metz, The Acoustic Impedance Measured on Normal and Pathological Ears, *Acta Oto-Laryngol., Suppl.* 63, 1–254 (1946); A. H. Inglis, C. H. G. Gray, and R. T. Jenkins, A Voice and Ear for Telephone Measurements, *Bell System Tech. J.* **11**, 293–317 (1932).

[2] P. G. Weissler, International Standard Reference Zero for Audiometers, *J. Acoust. Soc. Am.* **44**, 264–275 (1968).

[3] R. R. Reisz, Differential Intensity Sensitivity of the Ear for Pure Tones, *Phys. Rev.* **31**, 867–875 (1928).

[4] G. A. Miller, Sensitivity to Changes in the Intensity of White Noise and Its Relation to Masking and Loudness, *J. Acoust. Soc. Am.* **19**, 609–619 (1947).

The ability to distinguish pitch is subject to a greater range of individual variability than other functions reported here. The data given are for three trained listeners and have been smoothed in both directions. Untrained listeners usually require a greater

TABLE 3k-2. ACOUSTIC IMPEDANCE OF THE EAR IN ACOUSTIC OHMS, MEASURED JUST WITHIN THE MEATUS

Frequency	Total impedance	Resistive component	Reactive component
250	200	50	−190
350	150	40	−145
500	125	35	−115
700	70	25	−65
1,000	55	25	−50

Above 1,000 Hz measurements depend increasingly on the method of measurement.

TABLE 3k-3. MINIMUM AUDIBLE (THRESHOLD) PRESSURE AT ENTRANCE TO EXTERNAL EAR CANAL (MAC)*
(In dB re 2×10^{-4} dyne/cm^2)

	Frequency, Hz										
	125	250	500	1,000	1,500	2,000	3,000	4,000	6,000	8,000	10,000
MAC	35	22	14	8	9	9	10	9	14	17	16

The following quantities are to be added in order to obtain threshold pressures for other conditions:

a. MAC to Threshold Pressure at Tympanic Membrane†

	Frequency, Hz							
	125	250	500	1,000	2,000	4,000	6,000	8,000
Add...........	0.0	0.0	−0.5	−1.0	−4.5	−10.5	−4.0	−2.5

b. MAC to Free Field (MAF) (plane wave, 0° azimuth in absence of head)‡

	Frequency, Hz								
	125	250	500	1,000	2,000	4,000	6,000	8,000	10,000
Add....	+1.0	+0.5	−2.0	−4.0	−11.0	−12.5	−7.0	−3.0	−3.0

---

* J. P. Albrite, R. E. Shutts, M. B. Whitlock, R. K. Cook, E. L. R. Corliss, and M. D. Burkhard, Research in Normal Threshold of Hearing, *AMA Arch. Otolaryngol.* **68**, 194–198 (1958).

† F. M. Wiener and D. A. Ross, The Pressure Distribution in the Auditory Canal in a Progressive Sound Field, *J. Acoust. Soc. Am.* **18**, 401–408 (1946).

‡ L. J. Sivian and S. D. White, On Minimum Audible Sound Fields, *J. Acoust. Soc. Am.* **4**, 288–321 (1933).

TABLE 3k-3. MINIMUM AUDIBLE (THRESHOLD) PRESSURE AT ENTRANCE TO EXTERNAL EAR CANAL (MAC) (*Continued*)

c. Mean Monaural to Mean Binaural Listening§

	Frequency, Hz				
	125–2,000	4,000	6,000	8,000	10,000
Add..................	−2.0	−3.0	−4.0	−5.0	−6.0

d. Reference Age Group (18–25) to Older Age Groups¶

	Frequency, Hz					
	125–1,000	2,000	4,000	6,000	8,000	10,000
Add for:						
Men 30–39..........	+1.0	+2.0	+5.0	+6.0	+6.0	+7.0
Men 40–49..........	+2.0	+5.0	+13.0	+13.0	+11.0	+13.0
Men 50–59..........	+5.0	+13.0	+27.0	+32.0	+35.0	+35.0
Women 30–39......	+1.0	+2.0	+3.0	+4.0	+4.0	+4.0
Women 40–49......	+3.0	+5.0	+6.0	+8.0	+9.0	+9.0
Women 50–59......	+5.0	+9.0	+13.0	+18.0	+20.0	+22.0

§ H. Fletcher, "Speech and Hearing in Communication," p. 131, D. Van Nostrand Company, Inc., Princeton, N.J., 1953.

¶ J. C. Steinberg, H. C. Montgomery, and M. B. Gardner, Results of the World's Fair Hearing Tests, *J. Acoust. Soc. Am.* **12**, 291–301 (1940); J. C. Webster, H. W. Himes, and M. Lichtenstein, San Diego County Fair Hearing Survey, *J. Acoust. Soc. Am.* **22**, 473–483 (1950).

TABLE 3k-4. DIFFERENTIAL THRESHOLD FOR INTENSITY, IN DECIBELS

Sensation level, dB above absolute threshold	Pure tones, frequency in Hz							
	35	70	200	1,000	4,000	7,000	10,000	White noise
5	....	....	4.75	3.03	2.48	4.05	4.72	1.80
10	7.24	4.22	3.44	2.35	1.70	2.83	3.34	1.20
20	4.31	2.38	1.93	1.46	0.97	1.49	1.70	0.47
30	2.72	1.52	1.24	1.00	0.68	0.90	1.10	0.44
40	1.76	1.04	0.86	0.72	0.49	0.68	0.86	0.42
50	....	0.75	0.68	0.53	0.41	0.61	0.75	0.41
60	....	0.61	0.53	0.41	0.29	0.53	0.68	0.41
70	....	0.57	0.45	0.33	0.25	0.49	0.61	
80	....	....	0.41	0.29	0.25	0.45	0.57	
90	....	....	0.41	0.29	0.21	0.41		
100	....	....	....	0.25	0.21			
110	....	....	....	0.25				

frequency difference than that reported here. Note also that individual listeners commonly show idiosyncrasies at particular frequencies.

Table 3k-5. Differential Threshold for Frequency, in $\Delta F/F$*

Sensation level, dB above absolute threshold	Pure tones, frequency in Hz						
	60	125	250	500	1,000	2,000	4,000
5	0.0252	0.0110	0.0097	0.0065	0.0049	0.0040	0.0077
10	0.0140	0.0060	0.0053	0.0035	0.0027	0.0022	0.0042
15	0.0092	0.0040	0.0035	0.0024	0.0018	0.0014	0.0028
20	0.0073	0.0032	0.0028	0.0019	0.0014	0.0012	0.0022
30		0.0032	0.0028	0.0019	0.0014	0.0011	0.0022

* J. D. Harris, Pitch Discrimination, *J. Acoust. Soc. Am.* **24,** 750–755 (1952).

3k-6. Masking. Masking refers to our inability to hear a weak sound in the presence of a louder sound. It is usually measured by the amount of change in the threshold of the weaker sound, i.e., how much more intense the weak sound must be made in order to be heard over the masking sound than it needed to be when the masking sound was not present. The masking of one pure tone by another is a complex function of the particular frequencies and of the absolute level of the respective tones. See any standard text on hearing for the curves describing this relationship.

The masking of a pure tone by a noise with a reasonably flat and continuous spectrum is a linear function (except at levels below 10 dB) of the total intensity within a "critical band" centered on the masked tone. The width of the critical band of frequencies whose total energy is just equal to the energy of the masked tone is given by Table 3k-6.

Table 3k-6. Width of "Critical Band" ΔF as a Function of Center Frequency F (10 log ΔF)*

	Frequency, Hz							
	100	250	500	1,000	2,000	4,000	8,000	10,000
ΔF, dB	19.4	17.1	17.1	18.0	19.9	23.1	27.7	29.2

* N. R. French and J. C. Steinberg, Factors Governing the Intelligibility of Speech Sounds, *J. Acoust. Soc. Am.* **19,** 90–119 (1947).

The masking of a narrow-band noise by two tones, one higher and one lower than the noise, shows a similar relationship. The masking produced by the two tones overlaps unless the tones are separated by more than a "critical band," at which point the masking begins to fall off sharply. The critical band measured in this way is 3 to 4 dB wider than the values given in Table 3k-6.[1]

The masking of one continuous noise by another can be thought of as a case of differential sensitivity to change in the intensity of a noise (see last column of Table 3k-4). Thus, above 40 dB SPL, if a weak noise is more than 10 dB less intense than a very similar masking noise, the weak noise will not be heard; its presence or absence

---

[1] E. Zwicker G. Flottorp, and S. S. Stevens, Critical Band Width in Loudness Summation, *J. Acoust. Soc. Am.* **29,** 548–557 (1957). See especially the summary of the concept of critical bands, pp. 554–557.

does not produce a discriminable difference in intensity. If the spectral compositions of the two noises, masking and masked, are quite different, then the critical-band concept must be employed.

3k-7. Sounds of Short Duration. Acoustic disturbances of very short duration, i.e., less than 0.0001 sec, are heard only to the extent that they transmit energy to the ear. Short pulses at ultrasonic frequencies are generally not heard unless they are rectified. Impulse or step functions excite the ear, but not efficiently.

At the opposite extreme, tones, or continuous noise, of duration greater than from 0.2 to 0.5 sec are generally heard independently of duration. Between these limits relatively complex relations are found.[1]

As a first approximation for both tones and noise, the effective intensity of short sounds is a function of total energy integrated over the duration of the sound. More accurately, the threshold is defined by[2]

$$I_t = kIt^{0.8} \tag{3k-1}$$

For some short tones and for many types of impulse noise, account must be taken of the frequency distribution of energy. Inasmuch as the ear varies in sensitivity as a function of frequency, any change in the shape or duration of a short acoustic pulse will also change its effectiveness because of the altered spectral composition.

3k-8. Loudness. Loudness and pitch are ways in which a listener reacts to sounds. Furthermore, within limits, a listener can use numbers to describe how much of a response he makes to the sound. These numbers usefully describe how loud or how high in pitch a sound seems to be. It is then necessary to relate how loud it is (subjective response) to how intense it is in physical terms. The loudness of a pure tone of 1,000 Hz is described by the following relationship:

$$\log L = 0.0301N - 1.204 \tag{3k-2}$$

in which L is the loudness measured in sones and N is the loudness level in phons (equal to the sound pressure level of the tone in decibels above 0.0002 dyne/cm^2).[3] Another way of putting this is to say that loudness doubles for each 10-dB change in sound pressure level.

There is some evidence that the loudness of a noise grows more rapidly than that of a tone with an increase in sound pressure level, especially at low levels. The exact relations are less well known than those for a tone.

The loudness of tones at other frequencies than 1,000 Hz is given by determining the loudness level in the manner described below and converting to sones by Eq. (3k-2).

The loudness of noises can be measured by direct subjective comparison with a standard, such as a tone of 1,000 Hz, but such comparisons are difficult and need to be repeated by a number of judges. An approximation to the loudness of a noise can be calculated from measurements of the sound pressure level in a series of bands, usually a third-octave, a half-octave, or an octave in width, covering the audible spectrum.

The total loudness of the noise is given by the formula

$$L_t = S_m + F\left(\sum^{i} S_i - S_m\right) \tag{3k-3}$$

---

[1] S. S. Stevens, ed., "Handbook of Experimental Psychology," pp. 1020–1021, John Wiley & Sons, Inc., New York, 1951.

[2] D. B. Yntema, "The Probability of Hearing a Short Tone Near Threshold," Ph.D. Dissertation, Harvard University, 1954, 43 pp.

[3] S. S. Stevens, The Measurement of Loudness, *J. Acoust. Soc. Am.* **27**, 815–829 (1955).

The calculated loudness L_t should be qualified by the width of the bands used for its calculation. The terms S are empirical values of a loudness index shown as the parameter of the curves in Fig. 3k-1. The figure is entered with the geometric mean frequency of each band and the band pressure level as arguments. The loudness index S_i is estimated for each of the i bands. The band having the greatest index S_m is determined by inspection.

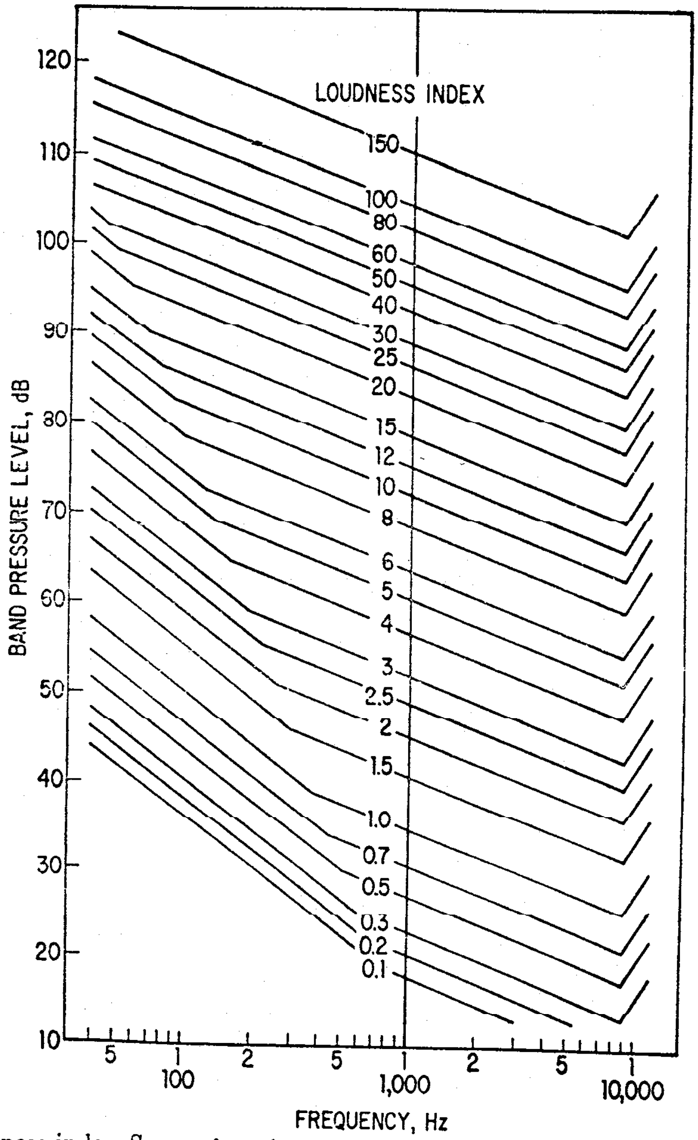

FIG. 3k-1. Loudness index S, as a function of geometric mean frequency of band measured and band pressure level (sound pressure level in third-octave, half-octave, or octave band under measurement). (*Taken from S. S. Stevens, "Procedure for Calculating Loudness," Mark VI, Psycho-Acoustic Laboratory Report PNR-253, Mar. 1, 1961.*)

As the formula indicates, total loudness is linearly additive except for a constant factor F that represents the reduction due to mutual masking of all bands except the loudest. The value of F depends on the width of the bands used. It has the value of 0.15 for third-octave, 0.2 for half-octave, and 0.3 for octave bands.

The loudness L_t can be converted to loudness level by Eq. (3k-2).

3k-9. Loudness Level. The loudness level of a tone of 1,000 Hz, expressed in phons, is defined as the sound pressure level in decibels above the reference level of 0.0002 dyne/cm^2.

The loudness level of tones of other frequencies is given by the empirical relations in Table 3k-7.

TABLE 3k-7. LOUDNESS LEVEL AS A FUNCTION OF SOUND PRESSURE LEVEL AND FREQUENCY*

Sound pressure level	Frequency, Hz							
	125	250	500	1,000	2,000	4,000	8,000	10,000
10	.....	.....	.....	10.0	18.0	18.0		
20	.....	6.3	16.0	20.0	28.0	28.0	11.0	
30	4.0	18.0	26.5	30.0	37.0	36.5	20.5	17.0
40	17.0	31.0	38.5	40.0	45.5	45.0	29.5	26.0
50	34.0	45.5	52.0	50.0	55.0	54.0	38.0	35.0
60	52.0	59.5	64.5	60.0	64.0	63.5	47.0	43.5
70	70.0	72.5	76.0	70.0	73.5	72.5	56.0	53.5
80	86.0	84.5	86.0	80.0	84.5	83.0	66.0	63.5
90	98.0	95.5	96.0	90.0	95.0	94.5	77.0	73.5
100	108.0	105.5	105.0	100.0	106.0	106.0	88.0	85.5
110	118.0	115.5	113.0	110.0	117.0	117.5	101.5	98.0

* American Standard for Noise Measurement, ASA Z24.2—1942.

Note that this table is based on the ASA standard and presumes the "free-field" measurement of sound pressure. This requires a measurement of a plane progressive wave at the listener's position before the listener is placed in the field. More meaningful measurements would doubtless be obtained from pressure measurements at the ear. For this purpose, apply the corrections contained in Table 3k-3*b* to the ear canal pressures before entering Table 3k-7.

To enter the table with sound pressure levels measured under other conditions, first add the corrections in Table 3k-3*b*, then subtract rather than add corrections in Tables 3k-3*a* and 3k-3*c*. Note, however, that corrections given for presbycusis in Table 3k-3*d* may give quite misleading results because of recruitment at high frequencies in some elderly people.

3k-10. Pitch. The relation between frequency and the subjective magnitude of perceived pitch is shown by Table 3k-8. By definition, the pitch of a tone of 1,000 Hz at 40 dB SPL is 1,000 mels.[1]

3k-11. Localization of Sound. The localization of complex sounds is primarily a function of time differences of arrival at the two ears, and to a first approximation, such differences can be calculated by assuming the ears on either end of the diameter of a sphere of 7.5 cm radius.

The localization of tones of low frequency (below 1,500 Hz) is possible on the basis of phase differences, which may be interpreted in terms of time differences.

The localization of tones of high frequency is possible on the basis of intensity differences resulting from the sound shadow of the head. Exact measurements here are difficult at best.

[1] S. S. Stevens and J. Volkmann, The Relation of Pitch to Frequency: A Revised Scale, *Am. J. Psychol.* **53**, 329–353 (1940).

TABLE 3k-8. PITCH OF A PURE TONE, IN MELS, AS A FUNCTION OF FREQUENCY

Frequency	Mels	Frequency	Mels	Frequency	Mels
20	0	350	460	1,750	1,428
30	24	400	508	2,000	1,545
40	46	500	602	2,500	1,771
60	87	600	690	3,000	1,962
80	126	700	775	3,500	2,116
100	161	800	854	4,000	2,250
150	237	900	929	5,000	2,478
200	301	1,000	1,000	6,000	2,657
250	358	1,250	1,154	7,000	2,800
300	409	1,500	1,296	10,000	3,075

Sound localization is greatly aided when the head or body can be rotated or moved about in the sound field while the observer hears the appropriate sequence of sounds.[1]

Sound localization in reverberant rooms or with so-called "stereophonic-sound sources" depends critically upon a "precedence effect," by which the localization determined by the primary sound or sound from the nearer of two sound sources is overriding in its effect.[2]

In experiments where time differences are used to balance out intensity differences in the opposite direction, 1.0×10^{-5} sec priority offsets a 6-dB difference in intensity; 2.3×10^{-5} sec offsets a 14-dB difference in intensity between the two ears.[3]

3k-12. Speech Power. The total radiated speech power, averaged over a 15-sec interval for a sample including both men and women at conversational levels used for telephone talking, has been estimated as 32 microwatts.

When measured at the face of a telephone transmitter, this power produces the sound pressure levels given in Table 3k-9 for different distances from the mouth of the speaker.[4]

TABLE 3k-9. AVERAGE SOUND PRESSURE LEVEL PRODUCED BY CONVERSATIONAL SPEECH AS A FUNCTION OF DISTANCE FROM LIPS TO MICROPHONE

	Distance, cm								
	Touching	0.5	1.0	2.5	5.0	10.0	25.0	50.0	100.0
Sound pressure level.....	104	102	99	95	90	85	78	72	66

A second source of variability lies in the essentially statistical distribution of speech power in time. If speech power is measured in successive $\frac{1}{8}$-sec intervals (a time slightly shorter than a syllable and slightly longer than a phoneme), a distribution is obtained with the mean values given in Table 3k-9 and variability that can be

---

[1] H. Wallach, Ueber die Wahrnehmung der Schallrichtung, *Psychol. Forsch.* **22**, 238–266 (1938).

[2] H. Wallach, E. B. Newman, and M. R. Rosenzweig, The Precedence Effect in Sound Localization, *Am. J. Psychol.* **62**, 313–336 (1949).

[3] J. H. Shaxby and F. H. Gage, Studies in the Localization of Sound. A. The Localization of Sounds in the Median Plane: An Experimental Investigation of the Physical Processes Concerned, *Med. Research Council (Brit.) Spec. Rept. Ser.* no. 166 (1932), 32 pp.

[4] M. H. Abrams, S. J. Goffard, J. Miller, F. H. Sanford, and S. S. Stevens, The Effect of Microphone Position on the Intelligibility of Speech in Noise, *OSRD Rept.* 4023 (1944), 16 pp.

attributed to time sampling equal to a standard deviation of 7.0 dB.[1] The distribution is badly skewed, so that the value 7.0 dB indicates only a rough order of magnitude. The variability is also greater when particular frequency bands are measured.

A third source of variability is the variation in effort expended by the person who is talking. As a rough approximation, a raised voice level is 6 dB above conversational level, the loudest level that can be maintained is 12 dB above conversational level, and the loudest shout is 18 dB above conversational level. In the other direction, a whisper may be 20 dB below conversational level.

3k-13. Speech Sounds

TABLE 3k-10. CHARACTERISTICS OF SOUNDS IN GENERAL AMERICAN SPEECH

Symbol	Example	Power,* dB re long time average†	Relative frequency of sound, %‡	Formant frequencies for men and women¶					
				First		Second		Third	
				M	W	M	W	M	W
u	cool	+0.6	1.60	300	370	870	950	2,240	2,670
ʊ	cook	+2.3	0.69	440	470	1,020	1,160	2,240	2,680
o	cone	+2.5	0.33	500	. . .	820			
ɔ	talk	+4.1	1.26	570	590	840	920	2,410	2,710
ɒ	cloth	+3.7	2.81	730	850	1,090	1,220	2,440	2,810
ɑ	calm		0.49						
a	ask	+2.5	3.95	660	860	1,720	2,050	2,410	2,850
æ	bat								
ɛ	bet	+1.6	3.44	530	610	1,840	2,330	2,480	2,990
e	tape	+1.4	1.84						
ɪ	bit	0.0	8.53	390	430	1,990	2,480	2,550	3,070
i	beet	0.0	2.12	270	310	2,290	2,790	3,010	3,310
ɚ	bird	−0.5	0.53	490	500	1,350	1,640	1,690	1,960
ə	sofa	. . . . .	4.63						
ʌ	bun	+2.9	2.33	640	760	1,190	1,400	2,390	2,780
eɪ	laid	+1.4	see e						
aɪ	bite	+2.5	1.59						
ju	you	+0.6	0.31						
oʊ	soap	+2.5	1.30						
aʊ	about	+2.3	0.59						
ɔɪ	boil	+3.0	0.09						

* The power measurements do not represent the peak instantaneous power but the average over the sustained portion of the phoneme where such a period can be defined. In this case, as with the formant frequencies, the absolute values are highly variable, but intercomparisons among the various sounds are generally more reliable.

† H. Fletcher, "Speech and Hearing in Communication," p. 86, D. Van Nostrand Company. Inc., Princeton, N.J., 1953.

‡ G. Dewey, "Relative Frequency of English Speech Sounds," Harvard University Press, Cambridge, Mass., 1923.

¶ E. G. Richardson, ed., "Technical Aspects of Sound," pp. 215–217, Elsevier Press, Inc., New York 1953.

---

[1] H. K. Dunn and S. D. White, Statistical Measurements on Conversational Speech, *J. Acoust. Soc. Am.* **11**, 278–288 (1940).

TABLE 3k-10. CHARACTERISTICS OF SOUNDS IN GENERAL AMERICAN SPEECH (*Continued*)

Symbol	Example	Power,* dB re long time average†	Relative frequency of sound, %‡	Formant frequencies for men and women¶			
				First	Second	Third	Fourth
l	lip	−3.0	3.74	450	1,000	2,550	2,950
m	me	−5.8	2.78	140	1,250	2,250	2,750
n	nip	−7.4	7.24	140	1,450	2,300	2,750
ŋ	sing	−4.4	0.96	140	2,350	2,750	
w	we	0.0	2.08				
r	rip	−1.0	6.35	500	1,350	1,850	3,500
j	yes	0.0	0.60	270	2,040		
p	pie	−15.2	2.04	. . .	800	1,350	
t	tie	−11.2	7.13	. . .	1,700	2,450	
k	key	−11.9	2.71	. . .	Variable		
b	by	−14.6	1.81	140	800	1,350	
d	die	−14.6	4.31	140	1,700	2,450	
g	guy	−11.2	0.74	140	Variable		
v	vie	−12.2	2.28	140	1,150	2,500	3,650
f	foe	−16.0	1.84	. . .	1,150	2,500	3,650
θ	thin	−23.0	0.37	. . .	1,450	2,550	
ð	then	−12.6	3.43	140	1,450	2,550	
s	sip	−11.0	4.55	. . .	2,000	2,700	
z	is	−11.0	2.97	140	2,000	2,700	
ʃ	shy	−4.0	0.82	. . .	2,150	2,650	
ʒ	measure	−10.0	0.05	140	2,150	2,650	
h	hit	−13.0	1.81				
tʃ	chop	−6.8	0.52				
dʒ	Joe	−9.4	0.44				

* The power measurements do not represent the peak instantaneous power but the average over the sustained portion of the phoneme where such a period can be defined. In this case, as with the formant frequencies, the absolute values are highly variable, but intercomparisons among the various sounds are generally more reliable.

† H. Fletcher, "Speech and Hearing in Communication," p. 86, D. Van Nostrand Company, Inc., Princeton, N.J., 1953.

‡ G. Dewey, "Relative Frequency of English Speech Sounds," Harvard University Press, Cambridge, Mass., 1923.

¶ E. G. Richardson, ed., "Technical Aspects of Sound," pp. 215–217, Elsevier Press, Inc., New York, 1953.

3k-14. Articulation Index. The articulation index is a set of numbers that makes possible the prediction of the efficiency of some types of voice-communication systems by the addition of suitably chosen values. The operations involve (1) dividing the speech spectrum into a series of bands having an equal possible contribution ΔA to the total efficiency, and (2) determining what proportion of the ΔA each band will contribute under the particular noise and speech conditions being tested.

Under (1) it is customary to use no more than 20 such bands. The frequency limits of 20 such bands are given in Table 3k-11.

TABLE 3k-11. TWENTY FREQUENCY BANDS CONTRIBUTING EQUALLY TO EFFICIENCY OF SPEECH COMMUNICATION*

Band No.	Frequency range	Band No.	Frequency range	Band No.	Frequency range
1	395	8	1,250-1,425	15	2,930-3,285
2	395-540	9	1,425-1,620	16	3,285-3,700
3	540-675	10	1,620-1,735	17	3,700-4,200
4	675-810	11	1,735-2,075	18	4,200-4,845
5	810-950	12	2,075-2,335	19	4,845-5,790
6	950-1,095	13	2,335-2,620	20	5,790
7	1,095-1,250	14	2,620-2,930		

* H. Fletcher, "Speech and Hearing in Communication," D. Van Nostrand Company, Inc., Princeton, N.J., 1953.

For conditions where substantial wide-band noise is present, the second requirement may be approximated by the formula

$$w_i = \tfrac{1}{30}(S_i - N_i + 6) \tag{3k-4}$$

in which w_i is a weight having a maximum value of 1.0, S_i is the signal level in band i in decibels, N_i is the noise level in the same band i in decibels referred to the same base as S_i.[1]

TABLE 3k-12. ARTICULATION SCORES AS A FUNCTION OF ARTICULATION INDEX*

Articulation index	CVC syllables, %	Monosyllabic words (PB lists), %
0.10	7	7
0.20	22	22
0.30	38	40
0.40	55	61
0.50	68	77
0.60	79	87
0.70	87	93
0.80	93	96
0.90	96	98
1.00	98	99

* E. G. Richardson, ed., "Technical Aspects of Sound," Elsevier Press, Inc., New York, 1953. "CVC syllables" are estimated from sets of words that vary the initial consonants, the vowel, and the final consonant separately. PB words are lists of monosyllables phonetically balanced so that the proportion of phonemes roughly equals that in general speech.

The articulation index A is then described by the summation

$$A = \frac{1}{n} \sum_{i=1}^{i=n} w_i \tag{3k-5}$$

Articulation scores are related to the articulation index according to the Table 3k-12.

---

[1] N. R. French and J. C. Steinberg, Factors Governing the Intelligibility of Speech Sounds, *J. Acoust. Soc. Am.* **19**, 90-119 (1947).