Useful Numerical Constants of Speech and Hearing

By HARVEY FLETCHER

NOTE: The material given in this paper was prepared in a more condensed form for publication in the International Critical Tables. In order to make it available in convenient form for the use of telephone engineers it was deemed advisable to publish it in this journal. The author is indebted to Dr. J. C. Steinberg for able assistance in collecting and arranging the material.

I. Bibliography

A BIBLIOGRAPH of papers on Pitch Discrimination, Intensity Discrimination, Absolute Sensitivity of the Ear, Upper Limit of Audibility, Lower Limit of Audibility, Theories of Hearing and other miscellaneous works on Speech and Hearing are given in a paper by H. Fletcher, *Bell Tech. Jour.*, Vol. II, 4, pp. 178–180, Oct., 1923.

II. Absolute Sensitivity of the Ear

The sensitivity is the minimum audible rms pressure in dynes cm^{-2} in ear canal. The values below are the average of the results of Wien (*Arch. f. ges. Physiol.* 97, p. 1, 1903), Fletcher and Wegel (*Phys. Rev.*, 19, p. 553, June, 1922), and Kranz (*Phys. Rev.*, 21, p. 573, May, 1923) weighted 3, 72, and 14, respectively according to number of ears tested

TABLE I

Frequency (dv)[1]	64	128	256	512	1024	2048	4096
Sensitivity (dynes)	.12	.021	.0039	.001	.00052	.00041	.00042

III. Minimum Audible Power for a Normal Ear

The power in microwatts passing through each square centimeter in the wave front of a free progressive wave in air under average conditions is related to the rms pressure in dynes by the formula

$$p = 20.5\sqrt{J}.$$

The figures of Table I may be converted by this formula to minimum audible powers. It is thus seen that the minimum audible accoustical power is at frequencies between 2,000 and 4,000 vibrations per second and is equal to 4×10^{-10} microwatts per square centimeter

[1] The symbol dv is used to denote "double" or complete vibrations.

IV. Range of Audition in Frequency and Intensity

In Fig. 1 the lower curve is a plot of the average sensitivity values given in Table I. The upper curve gives the pressures that produce a sensation of feeling and serves as a practical limit to the range of auditory sensation. (Wegel, *Bell Tech. Jour.*, 1, p. 56,

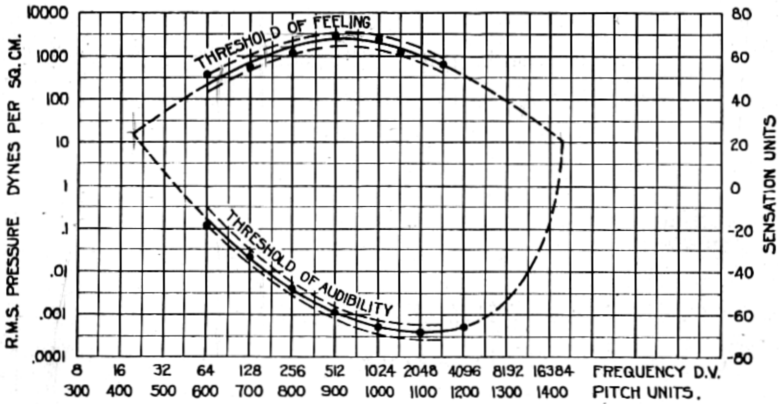


Fig. 1

November, 1922.) Investigators vary from about 8 to 40 dv for the lower pitch limit and from about 12,000 to 35,000 dv for the upper limit. (See I.) The values of 20 and 20,000 dv shown on the chart were taken as being most representative. Half of the observations lie within the dotted curves. The pitch is equal to 100 $\log_2 N$ and the sensation units equal to 20 log P where N is the frequency and P is the pressure. (Fletcher, *Jour. Frank. Inst.*, 194,

V. Minimum Perceptible Increase in Intensity and Frequency

(Knudsen, *Phys. Rev.* 21, p. 84, Jan., 1923)

Sensation Level in Sensation Units or TU's	Per Cent Increase in Intensity to be Just Perceptible
10	23
20	14
30	12
40	11
50	10.6
60 to 100	10

Frequency	Per Cent Increase in Frequency to be Just Perceptible
64	.93
128	.59
256	.40
512	.32
768 to 4096	.30

p. 289, Sept., 1923.) The sensation level S of a sound is defined by $S = 20 \log \frac{P}{P_o}$ where P_o is the threshold pressure, or it is the number of sensation units above the threshold of audibility. These sensation units are the same as the transmission units used in telephone engineering.

The per cent increase in frequency to be just perceptible varies with sensation level in about the same way as does the per cent increase in intensity to be just perceptible. The values are for monaural reception the tones being heard successively.

VI. The Number of Double Vibrations Necessary to Determine Pitch

(Bode, *Psychol. Stud.*, 2, p. 293, 1907)

TABLE II

	Weak Tones		Medium Tones	
Freq. dv	Time (sec.)	No. of dv	Time (sec.)	No. of dv
128	0.0496	12.1		
256			0.06908	17.6
384	.0672	24.08	0.0445	17.1
512	.0579	29.64	0.04274	21.8

VII. The Masking Effect of One Sound Upon the Audibility of Another Sound

(Wegel and Lane, *Phys. Rev.*, 23, p. 266, Feb., 1924)

If the ear is stimulated by a pure tone of frequency N_1, it is in general rendered less sensitive to other pure tones. The tone that constantly stimulates the ear is called the masking tone. The tone that is heard in the presence of this stimulating tone is called the masked tone. The masking is measured in sensation units or TU's. It is equal to $20 \times \log_{10}$ of the ratio of the pressures necessary to perceive the masked tone with and without the presence of the masking tone. In other words it is equal to the number of units that the threshold has been shifted. Fig. 2 shows the amount of masking (ordinate) of tones of various frequencies as a function of the sensa-

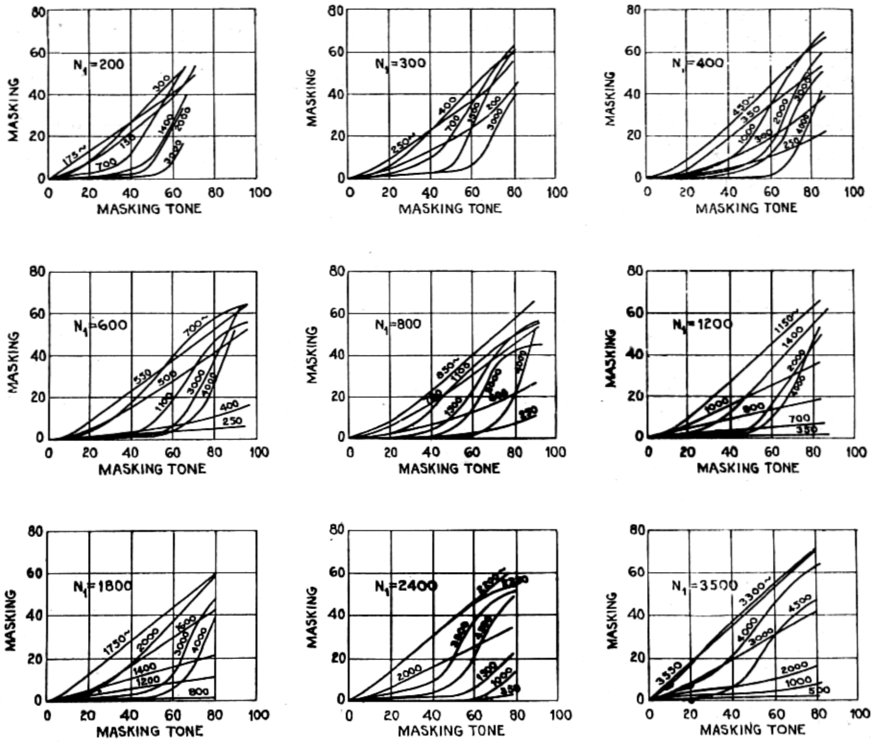


Fig. 2—Masking for Tones in Same Ear

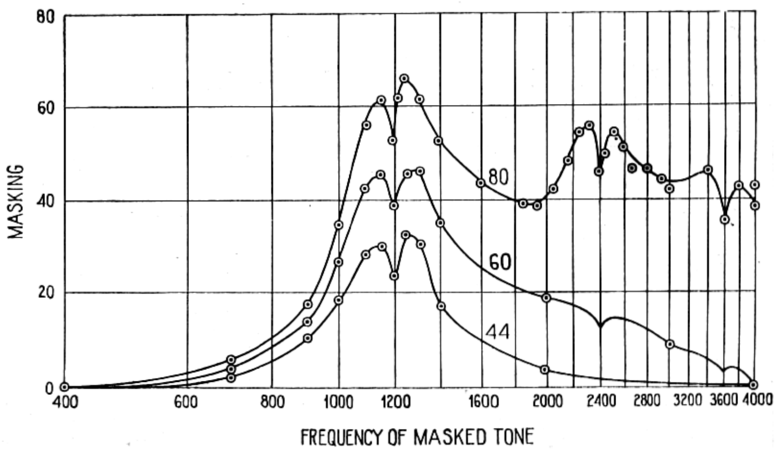


Fig. 3—Masking of Various Frequencies by 1,200 Cycles at Sensation Levels of 80, 60, and 44 Units, Respectively

tion level (abscissa) and frequency N_1 of the masking tone. In Fig. 3 data for a masking tone of 1,200 dv is plotted in which the frequencies of the masked tones are plotted on the abscissa. In order to get satisfactory curves of this kind it is necessary to take more

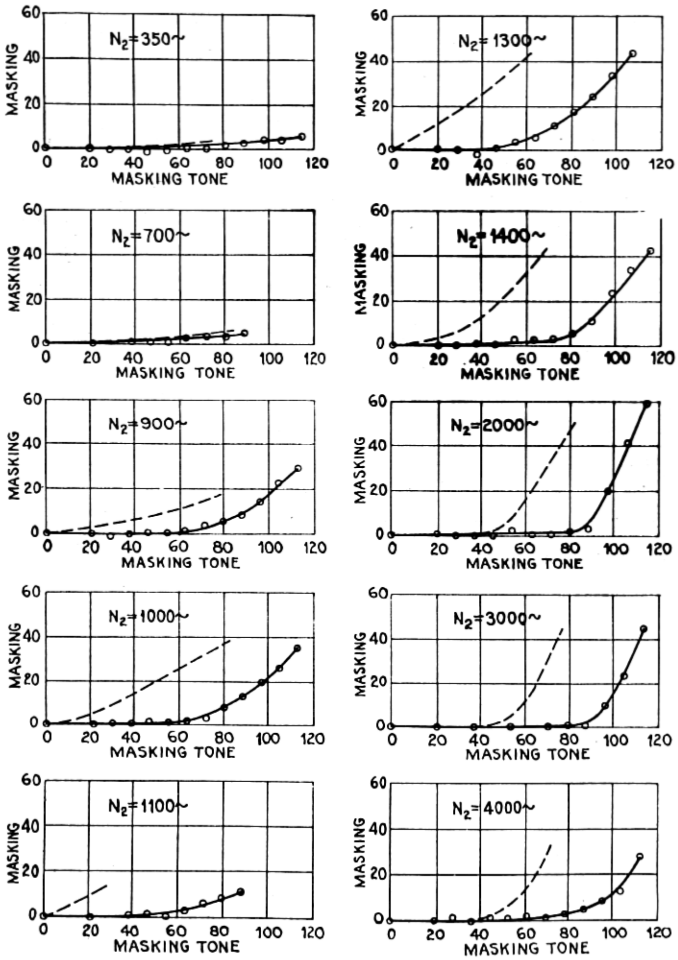


Fig. 4—Masking Data. Tones in Opposite Ears. Masking Tone 1,200 Cycles

comprehensive data than that shown in Fig. 2. The solid curves of Fig. 4 show the masking when the masked and masking tones are introduced into opposite ears. The dotted curves were taken from Fig. 2.

VIII. Conduction of Skull Between the Two Ears

A comparison of the two curves in Fig. 4 shows that the attenuation introduced by the skull from one ear to the other when the tone is introduced by a telephone receiver is between 40 and 50 sensation units corresponding to an intensity ratio of from 10^4 to 10^5. This becomes 7 TU greater when rubber caps are interposed between the head and the receiver cap.

IX. Localization of Pure Tones as a Function of the Phase Difference at the Two Ears

(G. W. Stewart, *Phys. Rev.*, 25, p. 425, May, 1920)

The experimental results can be represented by the formula

$$\frac{\Phi}{\Theta}=0.0034N+.8 \text{ (approx.)}$$

Φ is the phase difference in degrees of the tones at the two ears.

Θ is the number of degrees to the right or left of the median plane that an observer locates the source of sound. The direction of location is toward the ear leading in phase.

N is the frequency of the tone in dv. The relation applies only for frequencies of 100 to 1,000 dv., inclusive.

X. Constants Used in the Computation of the Loudness of a Complex Sound

(Fletcher and Steinberg, *Phys. Rev.*, 24, p. 306, Sept., 1924)
(Steinberg, *Phys. Rev.* To be published soon)

If L be the loudness as judged by an average normal ear, then

$$L=3.33 \log_{10}\left[\sum_{n=1}^{n=k} (W_n p_n)^{\frac{2}{r}}\right]^{\frac{r}{2}}$$

where

p_n = rms pressure of the n^{th} component,

W_n = a weight factor for the n^{th} component (Fig. 5)

r = a root factor (Fig. 5)

The sensation levels (See IV) given in the chart are for the complex tone.

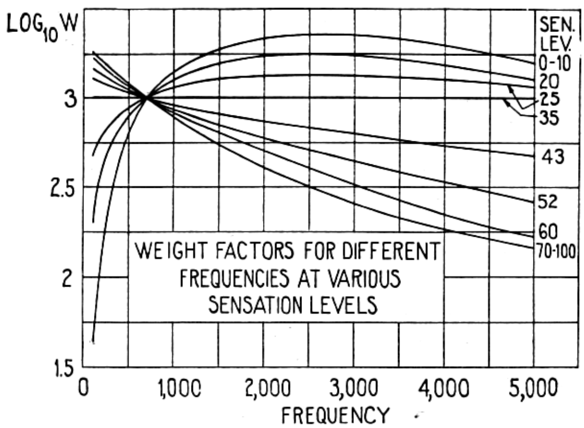
XI. Dynamical Constants of the Hearing Mechanism

(Howell, W. H., "A Textbook of Physiology"

(Wrightson, Sir Thomas, "Analytical Mechanism of the Internal Ear")

(*a*) Ear Canal

Length, 2.1–2.6 cm.
Volume, 1 cm^3.
Area at Opening, .33 to .50 cm^2.



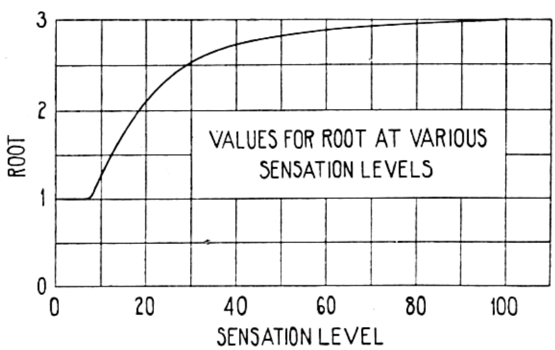


Fig. 5

(*b*) Drum

Vertical Diameter, .85 cm.
Horizontal Diameter, 1.00 cm.
Area, .65 cm^2.

(*c*) Hammer
Length, .8 to .9 cm.
Weight, 23 mg.

(*d*) Anvil
Weight, 25 mg.

(*e*) Stirrup
Weight, 3 mg.

(*f*) Mechanical Impedance of the Ear Drum
(Data by Wegel and Lane, Bell Telephone Laboratories)

The order of magnitude is 20 to 30 mechanical ohms (cgs units) over the frequency range from 200 to 4,000 dv.

XII. Speech Energy

A. Speech Power

(Data furnished by C. F. Sacia and L. J. Sivian, Bell Telephone Laboratories)

1. The average speech power delivered by an average speaker is about 10 microwatts. In the process of obtaining the average the silent intervals were included. If they are excluded the average increases about 50%. The peak power frequently rises to 2,000 microwatts.

2. Variation of average speech power delivered by different persons during conversation. (Fig. 6.)

B. Energy Frequency Distribution of Average Speech

(Crandall and MacKenzie, *Phys. Rev.*, 19, p. 221, March, 1922) (Fig. 7)

C. Acoustic Power in Vowel Sounds

(Data furnished by C. F. Sacia of the Bell Telephone Laboratories.

This data together with a description of the apparatus and methods used in obtaining it will be given in a paper soon to be published.)

Table III contains data on the power of individual vowels obtained from analyzing the vowel portions of the syllables shown in the key-word. The first two columns give the average power in microwatts of 8 males and 8 females during the particular cycle of the fundamental containing the maximum energy for unaccented vowels. A rough estimate of the corresponding figures of typical accented

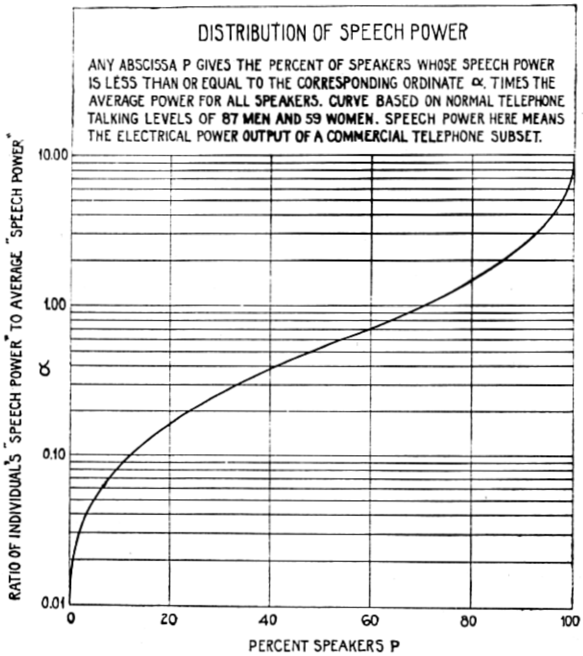


Fig. 6

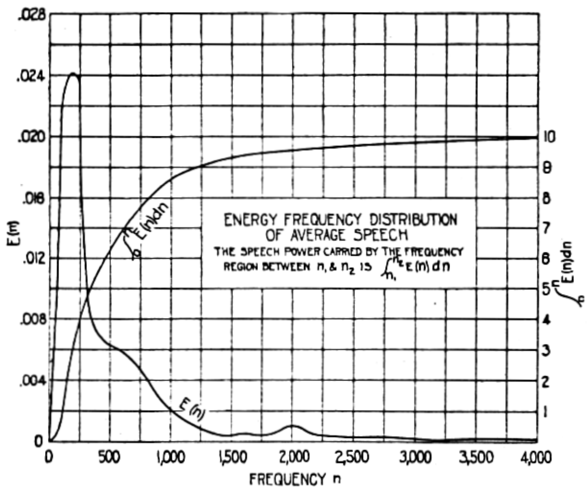


Fig. 7

vowels may be obtained by multiplying these values by a factor of 3. The third and fourth columns give peak factors which convert the power figures of the first two columns into maximum instantaneous powers. Columns 5 and 6 give the maximum values of these peak factors found among the male and female voices, respectively.

TABLE III

Acoustic Power in Microwatts of the Vowel Sounds

Vowel	Key	(1) P_m 8 males	(2) P_m 8 fem.	(3) Av. Peak Factor 8 males	(4) Av. Peak Factor 8 fem.	(5) Max. Peak Factor 8 males	(6) Max. Peak Factor 8 fem.
ū	tool	27	41	2.6	2.8	3.8	3.4
u	took	32	49	4.0	3.1	4.9	3.4
ō	tone	33	44	4.1	3.4	6.4	4.9
o′	talk	37	49	4.5	3.3	5.7	3.6
o	ton	29	38	4.6	3.9	6.8	5.7
a	top	50	48	4.2	3.6	4.2	4.7
a′	tap	43	39	5.4	4.7	7.4	5.2
e	ten	25	30	5.6	3.8	6.3	4.6
ā	tape	21	30	5.3	4.5	6.0	5.1
i	tip	25	31	4.1	3.8	5.8	5.7
ē	team	32	23	4.7	2.6	5.8	3.6

XIII. Frequency of Occurrence of English Speech Sounds

(Table IV contains data from a book by Godfrey Dewey, "The Relative Frequency of English Speech Sounds," Harvard University Press)

TABLE IV

Relative Frequency of Occurrence of English Speech Sounds

Speech Sound	Key	Rel. Freq.	Speech Sound	Key	Rel. Freq.
a	top	3.3	g		0.74
ā	tape	1.84	h		1.81
a′	tap	3.95	j		0.44
e	ten	3.44	k		2.71
ē	eat	2.12	l		3.74
er	term	0.63	m		2.78
i	tip	8.53	n		7.24
ī	dike	1.59	ng	hang	0.96
o	ton	6.33	p		2.04
ō	tone	1.63	r		6.88
o′	talk	1.35	s		4.55
u	took	0.71	sh	shell	0.87
ū	tool	1.89	th	(thin)	.37
ou	our	0.59	th	then	3.43
b		1.81	t		7.13
ch	chalk	0.52	v		2.28
d		4.31	w		2.08
f		1.84	y		0.60
			z		2.97

XIV. Interpretation of Speech

(Fletcher, H., *Jour. Frank. Inst.*, 193, 6, June, 1922)

A measure of the interpretation of speech was obtained by means of articulation tests. Meaningless syllables were pronounced and observers were required to record the syllables. The articulation is

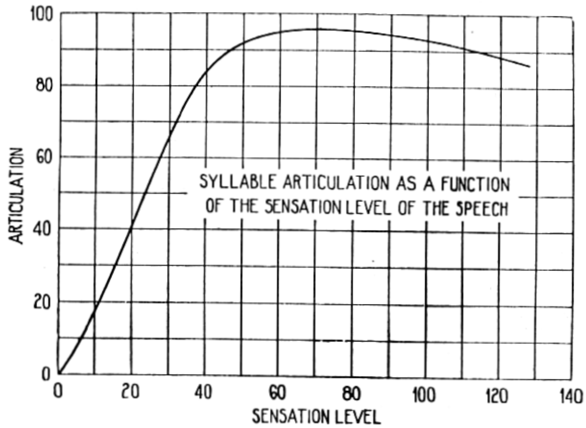


Fig. 8

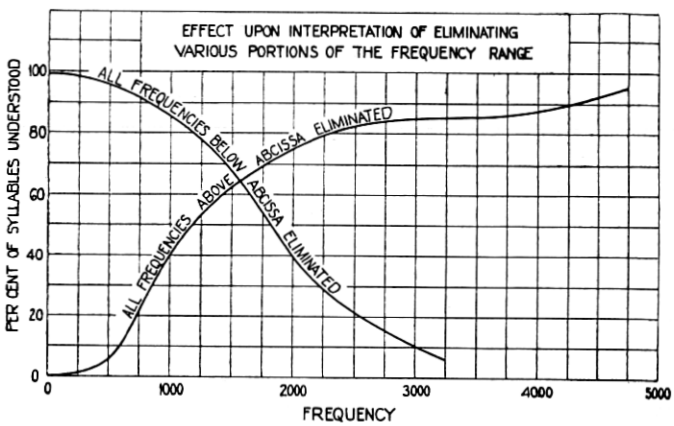


Fig. 9

the per cent of syllables that were correctly recorded. The articulation depends upon the sensation level of the speech (Fig. 8), and upon the width of the frequency band transmitted (Fig. 9).

The syllables that were recorded in these tests were analyzed to show the articulation of the fundamental speech sounds. Fig. 10

shows these articulations as functions of the sensation level of the speech. In Fig. 11 they are shown as functions of the width of the transmitted frequency band. It should be noted that the term articulation as here employed denotes only the correct interpretation of unrelated speech sounds and is not a measure of voice naturalness which is also an important factor in the telephonic transmission of speech.

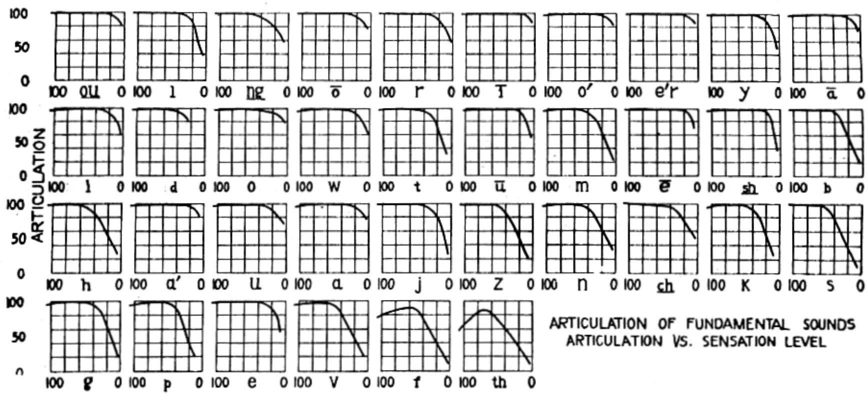


Fig. 10

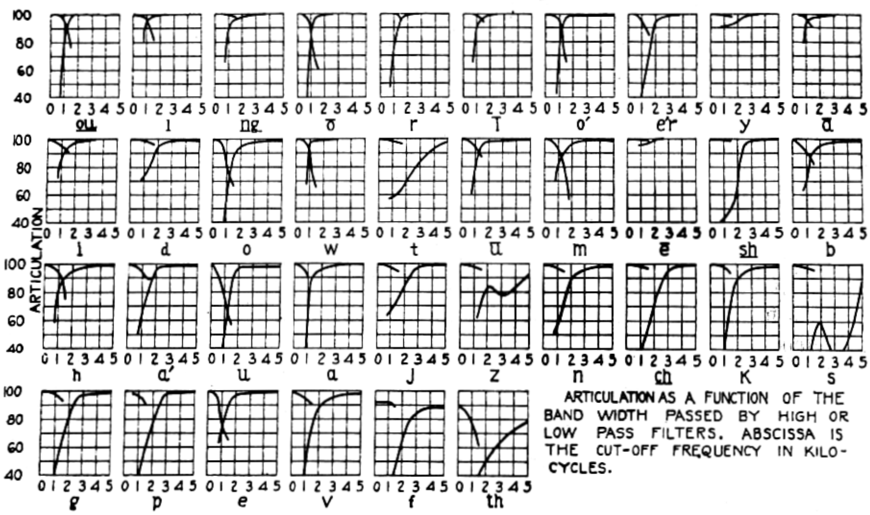


Fig. 11