Articulation Testing Methods

By H. FLETCHER and J. C. STEINBERG

This paper is chiefly concerned with the technique of making articulation tests. The construction of a syllabic testing list, the selection of a testing crew, the methods of comparing articulation data for various crews, and the significance of the test as a measure of the speech capabilities of a system are discussed. Various types of lists for different uses are also discussed.

THE transference of thought by means of speech is a very complicated, although common, process. So long as the process runs smoothly, its complications are forgotten. When an auditor fails to understand the speaker, however, inquiry into the reasons for the difficulty begins.

The production, the transmission, and the reception of speech constitute the three important elements of the process. To determine defects in any one of these, it is necessary to have a quantitative means of measuring the recognizability of the speech sounds that the auditor hears. The term "recognizability" as used here refers to correctness with which an auditor identifies a sound as being one, or some combination, of the fundamental speech sounds, when no meaning is associated with the sounds.

During the past few years methods of measuring the recognizability of speech sounds have come into greater and greater use both in this country and abroad. In order to compare the results obtained by various crews in various languages, it is desirable to standardize the methods of test and to set up reference circuits for purposes of calibration. It is the aim of this paper to discuss the methods that have been found the most useful, not only in determining defects in transmission, but defects in the production and reception of speech as well.

One needs only to tabulate the various devices that are used for transmitting speech to realize the importance of a quantitative method of rating their performance. There may be mentioned, for example, the various telephone and radio systems, the phonograph, sound pictures, rooms and auditoriums with various types of acoustic treatment, audiphone sets for the deafened, speaking tubes, etc.

Methods of measuring the recognizability of speech sounds have not been used so extensively for determining the ability of persons to speak properly. Such methods should be of value in the training of public speakers, actors, students of foreign languages or pupils in deaf schools who are learning to speak.

The rating of auditors by measuring the recognizability of speech sounds which they hear has been used to some extent. For example, such methods have been used to determine the ability of students to interpret a spoken foreign language. Also, the deafness of a person can be determined by such methods. In this case, however, the specialists have usually tried to vary transmission systems between the speaker and listener so as to compensate for the loss of hearing, the amount of such compensation being determined by measuring the recognizability of speech sounds.

The best method of determining the recognizability of the speech sounds naturally depends upon which of the things just enumerated is to be rated. In principle, the method in each case consists in the pronunciation of "selected speech sounds" by a speaker, the transmission of these sounds to an observer's ears, and the recording by the observer of the sounds which he recognizes. Such methods applied to telephone systems have been frequently referred to as articulation tests. The term "articulation" would be more logically used if it were applied only to cases where the speaking abilities of persons are being determined. However, it has been used so frequently in connection with rating transmission systems that it seems convenient to retain it.

The "selected speech sounds" which are ordinarily used in articulation tests are meaningless monosyllables. The percentage of the total number of spoken syllables which are correctly observed is called the syllable articulation. This percentage has frequently been called simply "the articulation."

A syllable is considered to be incorrectly observed, if one or more of the fundamental speech sounds which it contains are mistaken. It is frequently desirable to analyze these mistakes and determine the articulation of the speech sounds. The percentage of the total number of spoken sounds which are correctly observed is called the sound articulation. When the attention is directed toward a specific fundamental sound, such as "b" or "t" or "ā," etc., then the term "individual sound articulation" is used. For example, the individual sound articulation for "b" is the percentage of the number of times that "b" was called that it was observed correctly. Similarly, the terms "consonant articulation" or "vowel articulation" refer to the percentages of the total number of spoken consonant or vowel sounds which are correctly observed.

The articulation values as defined above are taken as the measures of the recognizabilities of the various phonetic units. English words and short sentences have also been used for testing purposes. When material of this kind is used, a new element enters, namely, the thought or meaning associated with the sentence or word. The criterion for the correct observation of words or sentences is also different from that used in the case of articulation tests. If the thought or meaning of a word or sentence is correctly understood, it is considered to be correctly received, even though the observer may not have correctly recognized each sound that was spoken. The terms "word articulation" or "sentence articulation," therefore, seem inappropriate when referring to the results of such tests. The term "intelligibility" has frequently been used in this sense. Since it has also been used in a more general sense, the terms "discrete word intelligibility" and "discrete sentence intelligibility" will be used when referring to the results obtained by using disconnected words or sentences for the testing material. They are defined as the percentage of the total number of spoken words and sentences, respectively, that are correctly interpreted according to the criterion given above.

Very early in the work of developing the telephone, words and sentences which were chosen in a haphazard way were used for testing purposes. Word lists of various sorts have been worked out and used with some success. Even in very recent years some of these word lists have been used to good advantage. The main objections which have developed, to the continuous use of such lists are: (*a*) it is hard to make the lists equally difficult without resorting to very long lists of words, (*b*) a very large number of lists are necessary in order to avoid memory effects.

Dr. G. A. Campbell [1] was one of the first to propose a system of syllabic speech sounds for testing the transmission characteristics of the telephone system. These syllables had no meaning and were constructed by combining the various initial consonants with the vowel "ee," such as bee, fee, etc. With these lists the consonant articulation was taken as a measure of the system.

Later Dr. I. B. Crandall [2] worked out a system which used both simple and compound consonant forms in a vowel-consonant and consonant-vowel type of syllable. All of the common vowels were used, and the combinations were formed in ways which are usually found in written speech. The sounds occurred with the same frequency as they occur in ordinary written material. As in the Campbell lists, the articulation was based on the consonant sounds alone.

[1] "Telephonic Intelligibility," G. A. Campbell, *Phil. Mag.*, Jan. 1910.
[2] "Composition of Speech," I. B. Crandall, *Phys. Rev.*, 10, p. 74, July 1917.

Several other lists which have not been published were proposed and used, the differences being in the choice of the fundamental speech sounds, in their arrangement into syllables, and in the relative frequency of occurrence, both of the different syllable forms and of the speech sounds in each form. There was a distinct effort to make the lists as nearly like speech as possible by using the syllable forms, and by using the particular combinations of fundamental sounds that occur frequently in English. Difficulties were encountered in testing, however, when this was carried too far in that enough different syllables could not be obtained for continuous testing. On the other hand, when random combinations of sounds were made, without regard to the particular combinations occurring in English, syllables that were very unusual and difficult to pronounce were obtained, unless the combinations were restricted to the simple syllable forms having only two or three sounds. In other words, testing lists must be selected with two things in mind; namely, they must be representative of speech and they must be suitable for making tests. The experience with these various lists also indicated that the results obtained with one system of lists could be calculated approximately from the results obtained with other systems by properly weighting the individual sound articulation values.

This experience led to the adoption by the Laboratories of a system of lists which has been used during the past ten years in studies of the effects of distortion upon the recognition of speech sounds.[3] These lists which have been referred to in the literature as the standard articulation lists were made up of only the con-vow, vow-con and con-vow-con syllable forms. The various fundamental sounds of English were combined at random into the syllables, such that each sound occurred approximately with uniform frequency.

During the past few years it has become evident that still further simplifications in the syllable forms used in the standard articulation lists might be made. Also our methods of making articulation tests and interpreting the results obtained have undergone considerable changes during this time. It is with these new methods that the present paper is chiefly concerned.

In order to distinguish between the old lists and the ones modified as described below, the prefixes "old" and "new" will be placed before the title "Standard Articulation Lists." When there is no chance for confusion, the new lists will be called simply the standard articulation lists, since they are the principal ones now being used in the work at Bell Telephone Laboratories.

[3] "Nature of Speech and Its Interpretation," H. Fletcher, *Journal Franklin Institute*, June, 1922.

New Standard Articulation Lists

In setting up any testing list it is necessary to classify and select a representative group of speech sounds. The National Phonetic Association uses a basic alphabet of 65 different sounds and also uses numerous modifiers which serve to distinguish slight variations in a given sound. Such a system is too complex for testing purposes. The revised scientific alphabet uses 48 simple sounds of which 24 are consonants, 19 vowels, and 5 diphthongs. Besides these fundamental sounds, connected speech contains certain recurrent combinations of them, such as st, ing, etc.

In speech these fundamental sounds are combined into syllables in a large variety of ways, but as mentioned before, in constructing a testing list it is desirable to adhere to very simple syllable forms. More complex forms which include the compound endings are either too few in number or involve unusual speech sound combinations. In either case they are soon memorized by a testing crew working with such lists. In the new lists, therefore, simplifications are made by omitting the con-vow and vow-con types of syllables, leaving only the con-vow-con type. In order to make syllables of this type it is obviously necessary to have the same number of vowels and consonants, provided that each consonant may be used in both the initial and the final position. Some consonants, however, can be used only in the former while others can be used only in the latter position.

With these facts in mind the sounds that are shown in Table I were adopted for these new lists. It will be noticed that all of the consonants are used in both the initial and final positions in the syllable, except h, w, and y, which are used only in the former, and zh, ng, and st, which are used only in the latter position. As was the case in the old standard lists, it will be seen that, in the new lists the vowel variants have been excluded. They occur infrequently in speech and phoneticians do not universely agree on their pronunciation. For this reason they are not included. Also, the diphthongs ī, ou, oi, and ew, which were used in the old lists, were omitted from the new lists. The last two of these diphthong sounds occur very infrequently in speech. Although the diphthongs, ī and ou, do occur quite frequently, it was felt that their essential properties were embraced by the properties of their constituent vowel sounds. By their omission and also by the introduction of the compound st as a final consonant, it is possible to construct any desired number of syllables of the con-vow-con type, from the speech sounds shown in the table.

TABLE I

SPEECH SOUNDS FOR NEW STANDARD TESTING LISTS *

Initial Consonant	I.P.A.	Key Word	Vowel	I.P.A.	Key Word	Final Consonant	I.P.A.
b			a	[ɑː]	father	b	
d			a			d	
f			ā	[eː]	fame	f	
g		go	ā			g	
k			a'	[æ]	fat	k	
l			a'			l	
m			e	[ɛ]	get	m	
n			e			n	
r			ē	[iː]	greet	r	
p			ē			p	
s			i	[ɪ]	tin	s	
sh	[ʃ]	ship	i			sh	
th'	[ð]	this	o	[ʌ]	but	th'	
th	[θ]	thin	o			th	
t			ō	[oː]	go	t	
v			ō			v	
ch	[tʃ]	church	u	[ʊ]	full	ch	
z			u			z	
j	[dʒ]	judge	ū	[uː]	rule	j	
h			ū			zh	[ʒ]
w			o'	[ɔː]	haul	ng	[ŋ]
y	[j]	yawl	o'			st	

Note: Final r and ng are used in the list only when they occur in combination with the following vowels:

a'r (as in carry, paragraph)
ar (as in are, far, tar)
er (as in bury, ferry, verify)
ir (as in spirit)
or (as in her, utter, fir)
o'r (as in for, lore)
ūr (as in your, sure)

a'ng (as in bang, sang)
eng (as in geng, e as in ten)
ing (as in sing, wring)
ong (as in sung, hung)
ung (as in gung, u as in took)
o'ng (as in long, wrong)

* The symbols for the sounds are those of the International Phonetic Association's alphabet. See Pronunciation of Standard English in America, Krapp, Oxford University Press, 1919. See also Revised Scientific Alphabet, Funk and Wagnall's Dictionary.

The testing syllables are formed in the following way. Cards upon which the initial consonants are written are placed in one box; others upon which the vowel sounds are written are placed in a second box; and those upon which the final consonant sounds are written are placed in a third box. A card from each box is drawn at random, thus forming the con-vow-con syllable. By drawing all of the sounds, a list of 22 syllables is formed. This process is repeated three times to obtain a list of 66 syllables which is a unit that has been found convenient to use. A list of syllables of about this length can be used without giving callers and observers a rest period. In such a list each initial consonant occurs three times, each vowel six times, and each final consonant three times.

In forming such syllables only those combinations involving final r and ng that are shown in Table I are included. Much confusion exists concerning the pronunciation of other combinations of these sounds. Syllables that represent slang in English are also omitted. These omissions are made by returning the card upon which the sound in question is written to its box and drawing another card. By combining the sounds at random in this manner any desired number of lists may be made which for practical purposes are all of equal difficulty.

In addition to containing a certain speech sound content, connected speech is characterized by inflection, accent, a rate of utterance, etc. In the earlier articulation studies the test syllables were called singly at intervals of about three seconds. When considered with respect to connected speech this procedure seems somewhat artificial. Comparative tests were made in which the syllables were called singly and as parts of introductory sentences. The tests showed the syllable articulation to be somewhat larger when the introductory sentences were used. The increase was due largely to the greater ease in interpreting the initial consonants of the syllables, when they were inserted in the introductory sentences. The effect was most noticeable for the stop and fricative consonants which have relatively short durations. In order to make the technique more nearly like connected speech the syllables are called in the short introductory sentences. The use of such introductory sentences also helps to insure that any element in the transmission system being tested, whose performance depends particularly upon their immediate past history, will be in the condition in which we are interested for determining speech transmission capabilities.

A list of sentences which is used for this purpose together with a sample record of articulation data is shown in the articulation test record of Table II. For calling purposes, the syllables on the cards are written in the spaces under the columns marked "called" of the test record. These sentences are called uniformly at the rate of 15 per minute. When the syllables in the first column are called, the sentences are repeated using the syllables in the second column and then those in the third column.

The observers are provided with blank articulation test record sheets. They write the sounds which they hear in the corresponding "observed" columns. When the test is completed the observed and called sheets are compared and the various articulations obtained.

For good results it has been found advisable to use a testing crew of ten people—5 men and 5 women. Eight people are ordinarily

TABLE II

Articulation Test Record

DATE 3-16-28 SYLLABLE ARTICULATION 51.5%

TITLE OF TEST Practice Tests CONDITION TESTED 1500~ Low Pass Filter

TEST NO. 10 OBSERVER W.H.S. Jr.

LIST NOS. 5-9-37 CALLER E.B.

NO.		OBSERVED	CALLED	OBSERVED	CALLED	OBSERVED	CALLED
1	THE FIRST GROUP IS	máv	náv	póz	pó'th́	Kōb ✓	Kōb
2	CAN YOU HEAR	pōch ✓	pōch	nēz	nēzh	shēth́	siz
3	I WILL NOW SAY	seng ✓	seng	jóch ✓	jóch	fūch ✓	fūch
4	AS THE FOURTH WRITE	chūd ✓	chūd	th́ám ✓	th́ám	thōl ✓	thōl
5	WRITE DOWN	run ✓	run	hab ✓	hab	poth́ ✓	poth́
6	DID YOU UNDERSTAND	chiz	Kiz	def	doth	wám ✓	wám
7	I CONTINUE WITH	foz	fozh	chech	chej	gūm	gūn
8	THESE SOUNDS ARE	ló'l ✓	ló'l	lun	lon	nāsh	nāth
9	TRY THE COMBINATION	jās	zhāth	shāl ✓	shāl	vóg ✓	vóg
10	PLEASE RECORD	th́áth	th́ásh	muz ✓	muz	lung	long
11	WRITE THE FOLLOWING	wūr ✓	wūr	lĕd	bēd	diz	dizh
12	NOW TRY	yāp ✓	yāp	wif ✓	wif	KaK	taK
13	THIRTEEN WILL BE	mad	maj	gōst ✓	gōst	th́ár	zhár
14	YOU SHOULD OBSERVE	bēch	bēK	thāv	sāv	must ✓	must
15	WRITE CLEARLY	gēm	dēm	Kōf ✓	Kōf	yód ✓	yód
16	NUMBER 16 IS	th́eb	veb	rág ✓	rág	jet ✓	jet
17	YOU MAY PERCEIVE	jok	jost	thip ✓	thip	rēp	rēj
18	I AM ABOUT TO SAY	gaf ✓	gaf	yar ✓	yar	th́ēp	hēp
19	TRY TO HEAR	hus ✓	hus	zhūt ✓	zhūt	—	chuv
20	PLEASE WRITE	hiv	thith́	Kūk	tūk	th́ef	th́esh
21	LISTEN CAREFULLY TO	tōg ✓	tōg	fung ✓	fung	bās ✓	bās
22	THE LAST GROUP IS	shót ✓	shót	th́ev	vesh	th́óf	shaf

E-1843 (6-28)

employed in a test, the remaining two being held for emergencies in order to keep the work going. One member of the crew calls at a time, and the remaining members act as observers. Ordinarily, eight callers are used with four observers recording simultaneously for each caller, although as many as eight or nine observers may be used. The order is arranged such that the various members are equally represented in the test.

Each observer's sheet (Table II) furnishes a value of syllable articulation (the percentage correctly observed), corresponding to a particular caller-observer pair and to 66 syllables or 198 speech sounds called. These values of syllable articulation are recorded in the form shown in Table III. The average of each column gives the average articulation for each observer. The averages of the rows give the callers' articulation.

TABLE III

Articulation Test Result Record

FROM 2/29/28 TITLE Practice Tests

TO 3/21/28 CONDITION 1500 Cycle Low Pass Filter

REFERENCE MM-2186 - 8/27/28 REMARKS

CONDITION

Caller \ Obs.	E.B.	W.H.S.	F.S.	C.M.	H.C.	A.H.	E.L.F.	M.W.	P.H.		Ave.
E.B.		51.5	51.5	50.0	48.5	57.5	59.0		45.5		51.9
W.H.S.	72.5		62.0	69.5	59.0	66.5	60.5		56.0		63.7
F.S.	47.0	54.5		56.0	48.5	57.5	50.0		42.5		50.9
C.M.	51.5	62.0	53.0		42.5	65.0	57.5		38.0		52.8
H.C.	48.5	56.0	48.5	36.5		50.0	54.5	30.5			46.4
A.H.	41.0	53.0	50.0	53.0	35.0		47.0	42.5			45.9
E.L.F.	51.5	51.5	51.5	45.5	53.0	57.5		41.0			50.2
M.W.	50.0	48.5	50.0	36.5	51.5	59.0	63.5				51.3
Ave.	51.7	53.9	52.4	49.6	48.3	59.0	56.0	38.0	45.5		51.6

When the articulation values are near 100 per cent or 0 per cent, then a group of values will not distribute itself symmetrically about the arithmetic mean or average value. For the high values, this is due to the fact that one cannot get a higher value than 100 per cent. To some observers, the 100 per cent mark may be obtained very easily and to others it may be obtained only with considerable effort. This difference in difficulty cannot be registered in the percentages obtained. A similar reason exists for the unsymmetrical grouping for values near zero. Our experiments have shown that this grouping is symmetrical in the range from 20 per cent to 80 per cent. For the range from 80 per cent to 100 per cent, the average value is less than the most frequent value, and for the range from 0 per cent to 20 per cent, the average value is greater than the most frequent value. These differences are of the order of 1 or 2 per cent. From an extensive series of tests, the averaging factor curve of Fig. 1 was constructed which enables the data to be averaged, in a way, such that the average value is approximately equal to the value which would be most frequently observed in a large number of tests. To do this, each observed articulation, based on 66 syllables, is converted into an averaging factor by means of the above curve. These factors are then averaged and the average value reconverted into average articulation by means of the above curve. The value so obtained is taken as the average syllable articulation for the test.

The articulations for the various sounds are recorded on the Articulation Test Analysis Record of Table IV. In making the analysis the total errors for each caller are counted. The occurrences per caller are the products of the number of times the sounds are spoken by the caller, and the number of observers. One analysis sheet contains the results for a complete test.

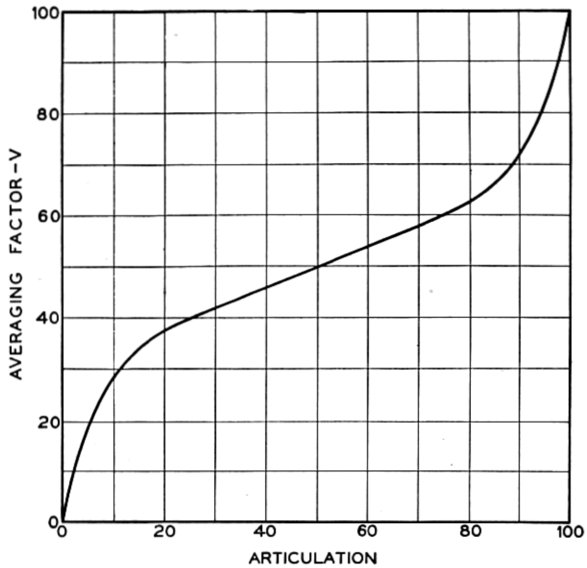


Fig. 1—Averaging factor curve

In dealing with the syllable articulation the unit is 66 called syllables, or the result of one caller and one observer. Hence a number of syllable values are obtained for one test which are averaged by means of the averaging factor curve. When we deal with the articulation of the individual speech sounds, it is advisable to use a larger unit since each sound occurs only six times in a 66 syllable unit. In Table IV the errors are shown for each caller as a unit, and since there were 7 observers per caller, each sound occurs 42 times. This is a unit of sufficient size to qualitatively compare various voices. However, in drawing conclusions as regards the effects of the circuit upon the transmission of the various sounds, it is best to use eight callers as a unit, so that each sound, depending upon the number of observers, occurs of an order of 200 to 300 times. In this case the articulation of an individual sound has a precision that is comparable with the precision of the syllable articulation when based on 66 called syllables, as will be discussed in a later paragraph.

TABLE IV

ARTICULATION TEST ANALYSIS RECORD

DATE 2/29/28 to 3/21/28 TITLE PRACTICE TESTS

REFERENCE MM-2186 8/27/28 CONDITION 1500~ LOW PASS FILTER

OBSERVERS EB, WHS, FS, CM, HC, AH, ELF, MW, PH. SOUND ARTICULATION 79.3 SYLLABLE ARTICULATION 51.6

Letter	Occur. per Caller	Errors per Caller: EB	WHS	FS	CM	HC	AH	ELF	MW	Total Errors	Total Occur.	Ind. Sound Art.
a	42	3	2	10	1	15	2	2	3	38	336	88.7
ā	42	1	1	4	0	7	2	0	2	17	336	94.9
a′	42	2	0	8	7	20	22	27	11	97	336	71.1
e	42	2	4	4	4	22	19	20	15	90	336	73.2
ē	42	0	0	2	0	1	4	0	1	8	336	97.6
i	42	2	0	2	0	9	8	9	3	33	336	90.2
o	42	12	1	6	5	12	5	4	10	55	336	83.6
ō	42	1	0	0	0	0	1	0	0	2	336	99.4
o′	42	0	0	5	2	0	0	0	0	7	336	97.9
u	42	5	2	19	1	8	3	4	4	46	336	83.6
ū	42	1	1	1	1	4	4	2	1	15	336	95.5
TOTAL	462	29	11	61	21	98	70	68	50	408	3696	89.0
b	42	7	0	11	10	13	1	4	6	52	336	84.5
ch	42	13	4	11	13	2	8	8	7	66	336	80.4
d	42	9	3	6	8	8	7	10	14	65	336	80.7
f	42	6	11	7	12	25	23	21	22	127	336	62.2
g	42	9	4	8	13	2	3	1	2	42	336	87.5
h	21	2	0	0	0	2	0	1	1	6	168	96.4
j	42	17	4	10	9	3	6	3	2	54	336	83.9
k	42	17	6	6	11	3	12	11	6	72	336	78.6
l	42	1	0	3	0	0	2	3	0	9	336	97.3
m	42	1	2	4	6	2	5	7	0	27	336	92.0
n	42	4	5	0	10	7	6	9	15	56	336	83.3
ng	21	0	0	2	3	1	1	4	5	16	168	90.5
p	42	11	11	4	5	19	21	14	8	93	336	72.3
r	42	3	0	0	2	4	3	4	2	18	336	94.6
s	42	18	5	16	9	23	22	18	17	128	336	61.9
sh	42	25	21	28	26	14	25	17	20	176	336	47.6
st	21	6	0	8	2	7	6	2	1	32	168	81.0
t	42	15	8	15	13	33	24	25	13	146	336	56.5
th	42	22	24	17	22	27	16	22	27	177	336	47.3
t́h	42	20	24	25	19	17	12	20	15	152	336	54.8
v	42	17	7	12	10	4	9	5	14	78	336	76.8
w	21	0	0	0	0	0	0	0	0	0	168	100.0
y	21	1	0	0	0	1	0	1	0	3	168	98.2
z	21	6	4	11	4	12	9	9	13	68	168	59.5
zh	42	33	24	32	35	27	27	23	26	227	336	32.4
TOTAL	924	263	167	236	242	256	248	242	236	1890	7392	74.4

E-1848 (6-28)

SELECTION OF TESTING PERSONNEL

It is necessary to set up the technique of testing, such that the values of syllable articulation can be reproduced within acceptable limits. The limits depend upon the control of the auditory and vocal characteristics of the testing crew, the control of numerous haphazard factors, and the control of the practice or experience that the crew acquires in the testing of circuits.

The departure from normal, in acuity of hearing of prospective crew members, should be measured with a good audiometer. In our laboratories the 2-A audiometer is used for this purpose. Only those individuals whose average hearing loss departs from normal in the speech range of frequencies (100 to 8,000 cycles per second) by less than 5 db (decibels) are selected. Although of normal hearing, some observers of a crew usually obtain higher values of syllable articulation than do others. The averages of the columns of Table III are a typical set of results for nine observers who have passed such a hearing test and who have had a year or more of experience in observing.

Observer A. H. obtained the highest percentage, namely 59, and observer M. W. obtained the lowest percentage, namely 38. In general this order would be preserved in a series of tests, although haphazard variations in a single test might change the order. The spread in observations is of an order of 20 per cent. More extended tests have shown that the spread tends to decrease as the observed percentages approach 0 or 100. In order to make a replacement in the observing personnel from time to time without causing a probable change of more than 2 per cent in the average percentage, it is necessary to use an observing crew of 8 to 10 persons. Our experience has shown that men and women show no characteristic difference when acting as observers.

The ability of prospective crew members to enunciate the sounds in a normal way is determined in the following manner. An extensive series of tests on various voices have yielded data which are arbitrarily used as a basis for determining normalcy. These tests were made with a simplified list consisting of common English words which will be described in a later paragraph (see Table XVII). Tests were made under three conditions; namely, direct transmission through the air in a quiet, well damped room, transmission over a circuit which uniformly transmitted the frequency range from 100–4,500 cycles, transmission over a circuit having a carbon transmitter. A diagram of this latter circuit is shown in Fig. 7. The sounds were observed by a crew of experienced observers. Table V gives the results of tests that were made upon 21 male and 23 female voices, the personnel being selected from various departments of the Laboratories. The average articulations of the simple consonant sounds are shown. The data are given separately for men and women.

TABLE V

NORMAL ENUNCIATION

Speech Sound	Air Transmission		Band 100 to 4,500~		Carbon Transmitter Circuit	
	Av. % Articulation of Sounds		Av. % Articulation of Sounds		Av. % Articulation of Sounds	
	Men	Women	Men	Women	Men	Women
b	98.7	98.0	96.2	90.1	95.0	91.5
ch	100.0	99.3	98.4	98.0	98.7	98.4
d	99.4	100.0	98.7	98.3	91.9	88.6
f	97.5	84.8	96.5	87.7	79.6	65.2
g	100.0	100.0	99.4	94.6	70.6	63.0
h	100.0	98.0	99.4	98.0	94.4	97.8
k	100.0	100.0	99.1	99.7	78.1	82.1
l	100.0	100.0	98.1	97.7	96.0	87.5
m	99.2	91.3	97.5	95.5	86.5	86.0
n	99.4	100.0	99.4	99.1	95.5	94.5
ng	100.0	100.0	99.7	99.4	93.7	100.0
p	96.9	95.8	97.2	96.6	80.0	73.5
r	100.0	100.0	100.0	99.7	96.9	75.5
s	97.5	99.1	95.0	68.0	91.2	67.8
sh	100.0	100.0	100.0	99.8	98.3	93.1
th	90.4	80.6	75.2	56.3	60.1	50.2
th'	97.5	87.0	93.3	87.7	74.1	77.2
t	98.3	100.0	99.4	96.6	92.9	72.1
v	93.3	78.9	96.2	83.6	87.1	75.0
w	100.0	99.3	100.0	98.9	97.5	84.8
z	100.0	100.0	95.9	70.8	91.2	65.2
Aver.	98.2	94.6	96.4	90.2	87.5	80.5

For our work, a prospective crew member is required to call such a list of syllables to a crew of experienced observers. If the observed articulations of the sounds are reasonably close to those indicated in Table V for each circuit, and if no obvious irregularities are noticed in the speech, the prospect is considered satisfactory for testing work. Measurements are also made upon the individual's speech power, but it has not been found necessary to use the information in the process of selection.

Aside from the practical application to the methods of testing, the table is interesting in showing characteristic differences between the voices of men and women. In general, woman's speech is more difficult to interpret than man's, particularly, in the case of the sibilant and fricative consonants. This is probably due to the fact that in woman's speech, these sounds are not only fainter, but occupy higher frequency ranges than in man's speech. The frequency range from 6,000 to 8,000 cycles for the former, is approximately equivalent to the range from 4,000 to 6,000 cycles for the latter. In the case of the voiced sounds, woman's speech has only one half as many components as man's, which also may cause greater difficulty in interpreting the former.

With respect to the vowel sounds, the crew members are instructed in the correct manner of enunciation. Only those vowels which have definite differences have been included in the testing lists, so that, slight differences in enunciation do not seriously affect the observed results.

The object of the selection process is to determine in a broad but definite way the normalcy in speech of prospective members, and to eliminate those individuals who have speech characteristics which are not readily reproducible should it be necessary to change the testing personnel. The row averages of Table III show a typical set of results for 8 callers who were selected in the above way, and have had a year or more of experience in calling.

The spread in results is of an order of 20 per cent, so that, if a crew of 8 to 10 callers is used, a replacement may be made in the calling personnel without causing a change in the average percentage of more than 2 per cent. Owing to inherent differences in the voices of men and women, they are equally represented on the testing crew. Individuals who have the equivalent of a high school education, and whose ages range from 18 to 23 years, are usually selected for this work.

Control of Haphazard Factors

Haphazard factors arise from various sources, some of which can be controlled reasonably well. The observers work in a sound-proof room, so that extraneous noises will not affect the articulation results. The calling is ordinarily done in a sound-proof booth that has been especially treated with sound absorbing material so as to reduce the reverberation time to an order of a few tenths of a second. Ordinarily the crew does not test more than two to four hours during the day, and the schedule is usually arranged so that this is not done continuously.

The intensity level of each caller is also measured during the test, as small variations in intensity level may cause rather large variations in articulation. Ordinarily the various callers are permitted to call at the intensity level most natural to them, although in some tests the callers all attempt, by watching an indicator, to call at the same level. Various instruments have been used for measuring the intensity levels during tests. The volume indicator [4] has proven quite satis-

[4] This instrument depends for its readings, essentially upon the syllabic powers of the vowel and semi-vowel sounds, so that the reading of the instrument is determined largely by the amplitudes in the frequency range from 100 to 2,000 cycles.

factory and is the instrument ordinarily used for this work. It has the advantage over some of the other instruments that were tried, of being in much more general use on speech circuits.

Control of other haphazard factors of a more or less psychological character, may best be obtained by taking enough data so as to average out their effects. This involves the number of syllables that are called by each speaker and the number of caller-observer pairs that are used in the test. The variability of caller-observer pairs for a calling unit of 66 syllables may be seen from Table III. The probable error [5] in percentage articulation of a single observation (ϵ_s) i.e., one caller-observer pair as taken from the data in the table, is $\pm$ 9. The probable error of the average articulation ($\epsilon_{av.}$) of the 56 caller-observer pairs is $\pm$ 1.2.

It has been found from a large number of tests that the probable error of a number of crews, each consisting of one caller and one observer, is of an order of $\pm$ 12 (per cent articulation) for a 66 syllable unit when the syllable articulation is around 50 per cent. This value tends to decrease with increasing experience in testing, and with increasing or decreasing values of syllable articulation. The use of 36 caller observer pairs obviously reduces the probable error to an order of $\pm$ 2 in percentage articulation, which is about the order of magnitude of the errors involved in maintaining the testing personnel over a period of time.

Since as will be shown in a later paragraph, the syllable articulation is equal to the cube of the sound articulation, the probable error in the sound articulation [6] for one caller and one observer, or a unit of 198 sounds, is of an order of $\pm$ 6 when its value is around 80 per cent. Since each individual sound is called only six times, the probable error for each individual sound for a single caller-observer pair is of an order of $\sqrt{\frac{198}{6}} \times 6 = \pm 35$. If a test comprises 4 observers per caller and 8 callers, each sound is called 192 times, which reduces the probable error for the articulation of each sound to $\pm$ 6. Under

[5] $\varepsilon_s = .67 \frac{\Sigma_d^2}{n-1}$ and $\varepsilon_{av.} = \frac{\varepsilon_s}{\sqrt{n}}$; where, n = number of caller-observer pairs; d = difference between the articulation of a caller-observer pair and the average articulation of n caller-observer pairs.

[6] $\varepsilon_s = \frac{\partial S}{\partial L}\varepsilon_L = 3L^2\varepsilon_L$

$\varepsilon_s = \pm 12,\ S = .5,\ L = S^{1/3},$

$\varepsilon_L = \frac{\varepsilon_s}{3S^{2/3}} = \pm 6.$

ε_s = prob. error in syl. art. for one caller-observer pair.
ε_L = prob. error in sound art. for one caller-observer pair.

the same circumstances, the probable error for the average syllable articulation is ± 2, and that for the sound articulation is ± 1.

Control of Practice Effects

The third factor entering into the reproducibility of articulation results is practice and experience. The practice effect manifests itself in various ways. An increase in articulation takes place as the observers become more familiar with the vocal characteristics of the speakers. Similar effects are observed as they become more accustomed to a given technique, or to a particular type of distortion. In general, these effects become smaller as the testing crew becomes more experienced.

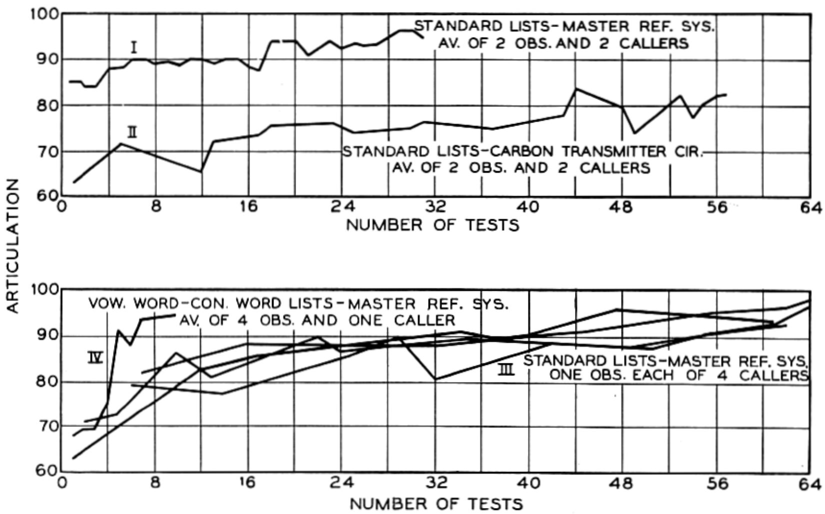


Fig. 2—Typical growth curves

Fig. 2 shows several typical growth curves that were obtained in the process of training new crew members. In this process the new members observe continually on various circuits until the results compare favorably with the results that are obtained by experienced observers. In such tests, experienced speakers are used. The averages for two new observers, of the results that were obtained on a high grade circuit, are shown by Curve I. Two speakers were used in these tests. A limited amount of testing was done by the observers prior to the above tests. Upon the completion of the tests of Curve I about 30 or 40 additional tests were made on various circuits. A series of tests, in which several speakers were used, were then undertaken on a carbon transmitter circuit. In Curve II the averages for the two observers, of results on two voices, are shown. Three to four weeks' time was spent by the observers in making the various tests mentioned above.

The curves under III show similar data that were taken at a later date by one new observer, for several voices. All of the above tests

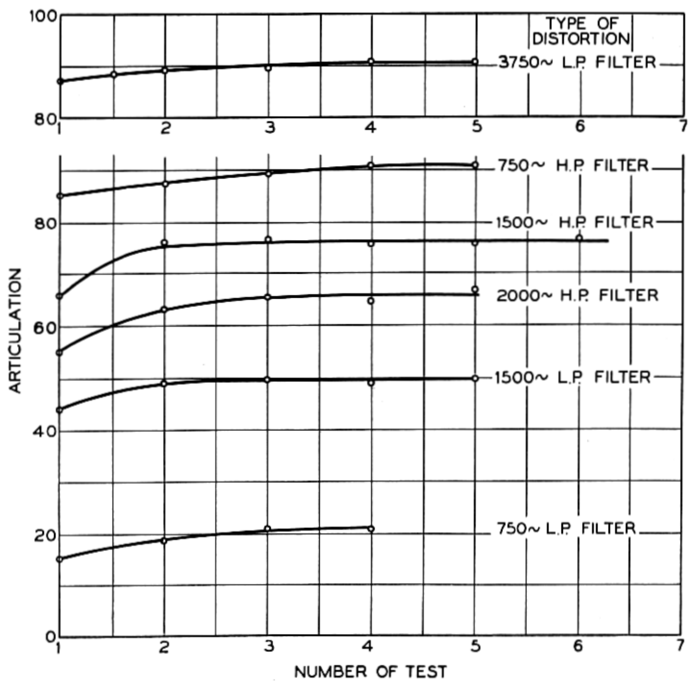


Fig. 3—Practice effects for an experienced crew

were made with the standard lists. In Curve IV, data are shown that were obtained with the vowel and consonant word lists (see Table XVII). In these tests four new observers were used and no preliminary training was given. It is evident that with the word lists, the results reach a state of saturation very quickly.

After a crew has spent several months in testing, its performance becomes largely mechanical. Under such circumstances the practice effects are rather small for types of distorted speech with which it has had experience. When the speech distortion is unusual, however, rather large practice effects may be obtained. Fig. 3 shows such practice effects for several types of distortion for a crew of eight people. All six of the circuits were tested on each test before going on to the following test. The first three tests were made successively and covered a period of about two months. In each test the types of distortion were interspersed. In other words half of the first test was completed with the filters in one order, and the other half with the filters in the reverse order. The fourth test was made about three

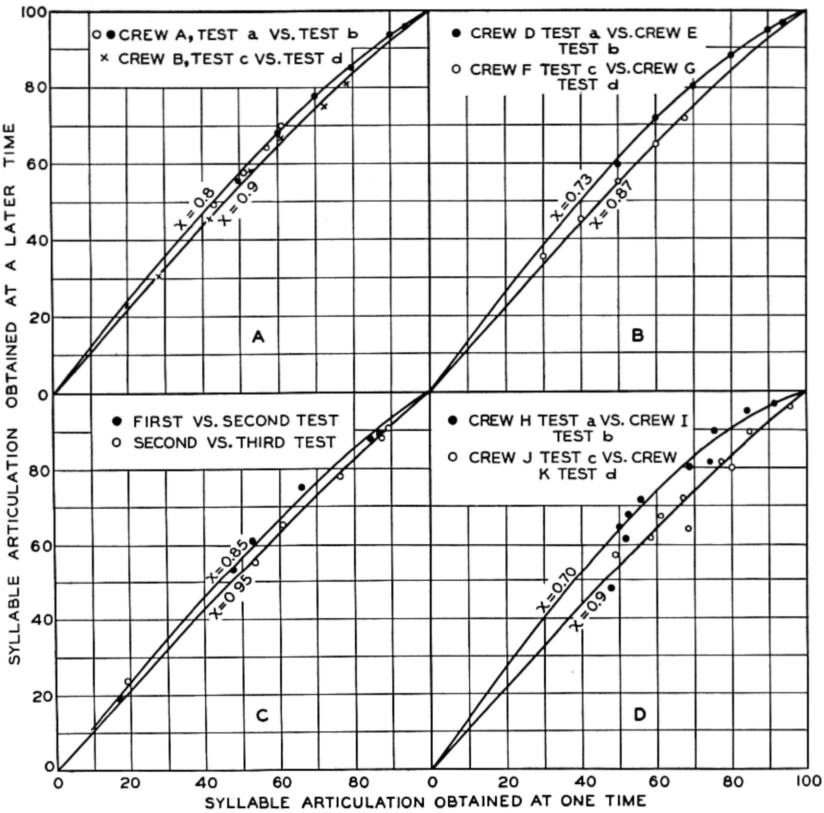


Fig. 4—Practice effects

months later, and the fifth test was made approximately six months later. Although the crew had been testing various circuits for about a year, and were thoroughly accustomed to the routine, these particular circuits had not been previously tested, so that, the crew's experience with these types of speech distortion was small. It is evident that the articulation of an experienced crew reaches saturation very quickly. It is probable that practice effects for a crew of several years' experience with distorted speech would be negligible.

Several procedures are followed in order to correct, in so far as possible, for practice effects. In comparative tests, whenever it is possible, the circuits to be compared are interspersed so as to average out practice effects. If it is desired to compare the articulation of a very new or unusual circuit (from the standpoint of the speech distortion) with one of common experience, several successive tests are made upon the new circuit until no further increase in articulation with practice appears. When it is impossible to intersperse the tests, the data may be corrected to a given state of practice by means of curves which were obtained in the following way. Although as will be seen, this procedure is valid only under certain restrictions, which will be discussed, such a correction will always tend to correct the data to a more comparable basis.

In Fig. 4-*a* a practice curve is shown that was obtained for a crew, from two series of tests that were separated by an interval of three months. The dots represent tests that were made upon a circuit which uniformly transmitted a frequency range from 100 to 5,500 cycles. The circles were obtained from a circuit of the type shown in Fig. 7 involving the carbon transmitter. In both cases the various articulation values correspond to different received speech levels. The crosses represent similar results that were obtained with a different crew on the latter type of circuit.

In Fig. 4-*c* the data of the first three tests in Fig. 3 are shown. In this case the distortion was varied and the received speech level held constant. As previously stated, in so far as was known, the crew had no previous experience with these types of speech distortion so that the practice for the various types of distortion ought to be comparable.

All of the solid curves are graphs of the following equation

$$(1 - S') = (1 - S)^x \cdots, \tag{1}$$

where S' = decimal value of syllable articulation obtained on a given circuit at one stage of a crew's career,

S = the value obtained on the same circuit at a later stage of the crew's career,

x = a number called the practice factor.

The values of the practice factor x that were necessary in order to fit the observed values are shown in the figure. It is impossible to state definitely that a crew has uniform practice with various types of distortion for the reason that experience is cumulative. A crew's experience with one type of distortion may be of aid in the understanding of some other type of distorted speech. With this in mind it will be seen that a constant value of x fits the data for the various types of distortion reasonably well. In the case of changing speech levels with a constant type of speech distortion, where the question of uniformity of experience is not so important, the fit is even better.

It is reasonable to suppose that an inexperienced observer must make a greater mental effort than an experienced observer to obtain the same articulation values. In other words the element reflected

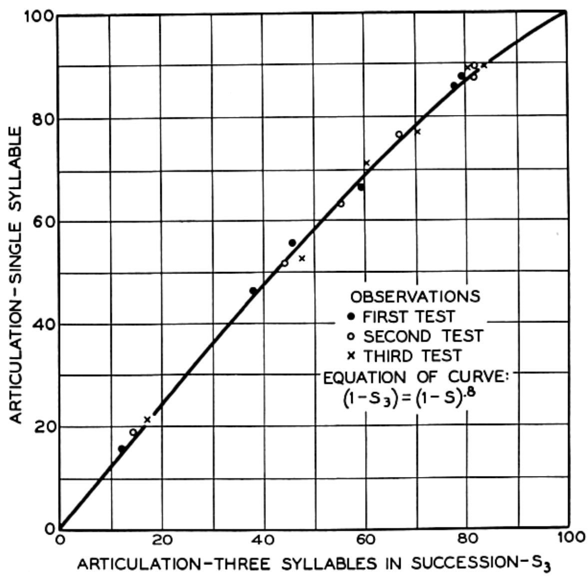


Fig. 5—Relation between techniques

by these curves is closely associated with the burden or strain upon the observer. A somewhat analogous situation obtains when tests are made with two different techniques which differ primarily in the burden imposed on the observer. In making the filter tests described above, two techniques which differed in this respect were used. One was the standard technique, in which one syllable was called with the introductory sentences. In the other, the syllables were called in groups of three (three in succession) with the sentences. The syllables were uttered as nearly in the manner of a three syllable word as was possible.

The results that were obtained with the two techniques are shown in Fig. 5. When the syllables are called in groups of three, the articulation values are smaller than when they are called singly.

It was found, however, that the type of relation shown in Eq. 1 also relates the data obtained with the two techniques. In this case the relation may be expressed as follows:

$$(1 - S_3) = (1 - S)^{.8 \cdots}, \tag{2}$$

where S_3 = decimal value of syllable articulation when called connectedly,

S = decimal value of syllable articulation when called singly.

The curve of Fig. 5 is a graph of the above equation.

In this case uniformity of experience with the various types of distortion does not enter, as the tests with the two techniques were made simultaneously. The only difference in the techniques was that in the three-syllable case the observer listened to three syllables before writing them down. It seems reasonable to conclude, therefore, that when a crew has the same experience with different types of distortion, then the results obtained by it at one time may be compared with the results obtained by it at some other time by using such a relation. No doubt other types of functions could be found which would also fit the above data. The relation shown here was chosen because it fit both the practice data and the data that were obtained with the two different techniques and is very convenient to use in making such corrections.

It is evident that in order to use the practice curves it is necessary to set up a reference circuit in order to obtain an appropriate value of x. Theoretically, one reference condition should be sufficient, provided that the practice of the crew had the same relative distribution over various types of speech distortion. Since this is usually not the case, it is necessary to use several reference circuits representing various types of speech distortion. When it is desired to correct data for practice effects, the appropriate value of x is determined by making tests upon the reference circuits having types of distortion similar to the circuits for which the corrections are desired. A description of several reference or control circuits which have been found useful with the values of sound and syllable articulation as obtained with the testing crew of five men and five women as previously described, is given below.

(*a*) *Air Transmission. Master Reference System for Telephone Transmission.*—The air transmission tests were made in a quiet, well damped room having a volume of approximately 1,000 cubic feet. The observers faced away and were located at an average distance of 30 inches from the speaker. Sound articulation "L" 99.1 per cent.

Syllable articulation "S" 97.5 per cent. Practically identical results were obtained with the "Master Reference System"[7] with the system set for optimum received speech level, i.e. a sensation level of 70 db, average distance from lips to transmitter 1.5 inches.

(*b*) *Auxiliary Circuit of the Master Reference System.*—The auxiliary circuit of the master reference system consists of networks which are

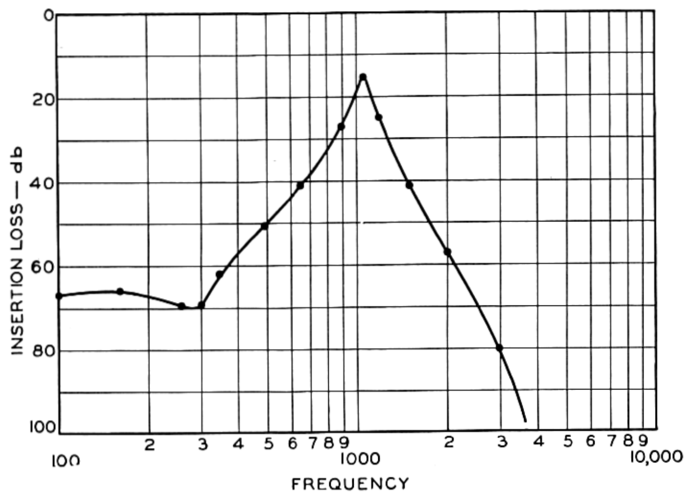


Fig. 6—Insertion loss of auxiliary networks

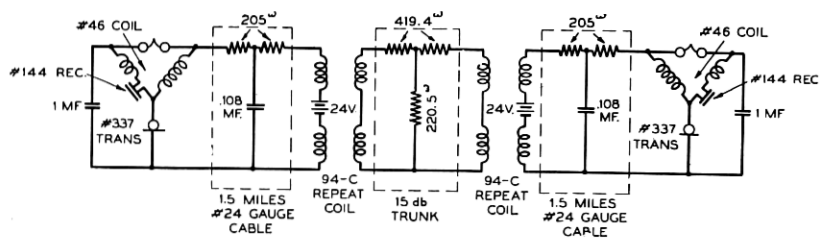


Fig. 7—Carbon transmitter circuit

inserted into the otherwise distortionless reference system, to give it a frequency resonance around 1,100 cycles. The insertion loss of the networks is shown in Fig. 6. This loss is approximately equal to the combined losses of the No. 1 transmitter and receiver distortion networks of the Master reference system. Sensation level 74 db $L = 89.2$ per cent, $S = 72$ per cent.

[7] L. J. Sivian, "A Telephone Transmission Reference System," *Electrical Communication*, 3, Oct. 1924. M. Cohen, "Apparatus Standards of Telephonic Transmission." W. H. Martin and C. H. G. Gray, "Master Reference System," *Bell Tech. Jour.*, July, 1929.

(*c*) *Carbon Transmitter Circuit* (see Fig. 7).—The average values for five transmitters are $L = 93$ per cent, $S = 81$ per cent. In these tests the sensation level of the received speech was 75 db, and the calling level as measured by a volume indicator bridged across the line side of the input repeating coil was $-$ 12.5 db.

(*d*) *Master Reference System Plus Filters.*—System set for a sensation level of 70 db.

3,750~ Low Pass Filter		$L = 96.7\%$	$S = 91.0\%$
750~ High " "		$L = 96.7\%$	$S = 91.0\%$
1,500~ Low " "		$L = 77.7\%$	$S = 49.5\%$
1,500~ High " "		$L = 91.0\%$	$S = 76.0\%$

The foregoing discussion has been concerned with methods of correcting the articulation results obtained by a given crew at different times to an arbitrary stage of practice or experience. To do this it is necessary to calibrate the crew for types of distortion that are similar to those of the systems for which the corrections are desired. The method has been described in detail because there are times when it is necessary to make such corrections. However, it has been our experience that such practice effects become negligible with a crew that has been set up in accordance with the methods previously described, when the crew's experience with types of distortion is diversified and when unusual circuits are tested successively until no further increase in articulation with practice occurs.

These methods may also be used to correlate the articulation data of various crews and various techniques, provided that the only essential difference between the crews and techniques is in the demand or burden that is placed upon the observer. This means that the crews must have similar vocal characteristics and similar hearing abilities, and that the testing lists must have similar speech sound content. It has been found, for example, that a crew of women callers obtain a considerably higher articulation than men callers on a circuit which eliminates all frequencies below 1,500 cycles and vice versa on a circuit which eliminates all frequencies above 1,500 cycles. It is obvious, therefore, that the methods described above could not be used to correlate the two crews for such circuits. Similarly, the methods could not be used for comparing two crews, if the hearing level of one is 10 db below the other, or to compare two techniques, one of which is made up entirely of vowel sounds and the other entirely of consonant sounds. As shown in Fig. 4-*b*, data have been obtained with various crews on various circuits which can be correlated very well by means of the above curves. Fig. 4-*d* gives data that were obtained with other crews which show very poor correlation. In neither case are the characteristics of the crew well enough known to satisfactorily account for the observed differences. At the time the work was done the significance of these factors was not so apparent, so that they were not given the attention they now receive. During the past two years a crew of 10 people has been used almost continuously in testing work. During this time numerous changes in personnel have taken place and only five of the original members are now on the crew. The data obtained during this time appear to be strictly comparable. In some cases it is necessary to use the practice curves. In other cases (circuits that are frequently tested), practically identical results are obtained. For this reason, it is believed that if a similar crew of 10 different people were to be selected as previously described, comparable articulation results would be obtained. It seems reasonable to expect that crews testing in various languages should also obtain comparable results provided that the crews were similar in the sense used here and that the lists were phonetically similar. It seems desirable, therefore, to standardize on the factors which affect the comparison of data, such as, the size and type of crew, the type of list, and the type and number of reference circuits. Best results are likely to be obtained when the crews do not differ by amounts which correspond to values of x less than 0.7. Smaller values indicate that the crews have not had sufficient testing experience, or have speech and hearing characteristics which are essentially different, or that the phonetic content of the testing lists are appreciably different. In the latter case the results may be correlated by means of statistical relations that will be given in a later paragraph.

Relation of Articulation to the Transference of Thought by Speech

The foregoing paragraphs have been concerned with the practical problems of setting up a suitable testing technique and correlating the observed articulation results. The procedure that has been discussed enables us to measure the percentages of the various speech sounds which are correctly recognized when they are spoken in a simple con-vow-con syllable. We desire at this point to consider the broader significance of this measure. In other words, how is the articulation result related to the transference of thought by means of speech? This relationship involves many psychological factors which are difficult to evaluate so it must not be expected that a comprehensive answer can be given here, but it is important to understand as fully as possible those parts of the problem that can be evaluated. Such a relation involves two questions, (*a*) how do the articulations of the sounds as measured with the testing lists compare with their articulations as they are used in speech, (*b*) how should the articulation values be weighted in order to obtain an index of the speech capabilities of a system.

In the first place, certain fundamental sounds of speech were omitted from the above lists. The most important of these are the consonant compounds. The majority of these sounds may be regarded as the product of a very few combining consonants acting as modifiers to the rest of the consonant alphabet. Since the combining consonants or modifiers occur over and over in combination with various consonants, it might be expected that the interpretation of the compounds would depend primarily upon the interpretation of the various consonants, and not upon the modifiers. In other words, the compounds would be interpreted as simple consonant sounds. The tests discussed below show that this is true on the average, although notable exceptions may occur in individual cases.

The testing lists were made up from the sounds shown in Table VI.

TABLE VI

Consonants		*Vowels*
Initial	*Final*	
b, br,	rb, b,	a'
d, dr,	rd, d,	a
g, gr,	rg, g,	e
p, pr,	rp, p,	i
k, kr,	rk, k,	o
t, tr,	rt, t,	
f, fr,	rf, f,	
th, thr,	rth, th,	
s, sl,	nd, d,	
b, bl,	nj, j,	
g, gl,	nz, z,	
p, pl,	nk, k,	
k, kl,	nt, t,	
f, fl,	ns, s,	
r, l,	n, r,	

These sounds were combined at random into syllables of the con-vow-con-con and con-con-vow-con form. Ten lists of 90 syllables each were made, and a crew of 10 callers, with 5 observers per caller, was used. With this number of tests the probable error in the per cent articulation for each sound is approximately 5 per cent. The tests were made on the auxiliary circuit of the master reference system. The sensation level of the received speech was about 80 db. The results are shown in Table VII.

TABLE VII

ARTICULATION OF CONSONANT COMPOUNDS

Initial				Final			
Sound	% Art.	Sound	% Art.	Sound	% Art.	Sound	% Art.
b	89.0	br	85.3	rb	63.3	b	77.0
d	98.0	dr	88.7	rd	60.0	d	91.0
g	96.3	gr	90.0	rg	57.3	g	88.7
p	78.3	pr	52.0	rp	39.3	p	66.7
k	95.3	kr	93.3	rk	52.7	k	90.7
t	79.3	tr	89.3	rt	58.7	t	88.7
f	56.3	fr	60.7	rf	39.3	f	53.3
th	71.3	thr	88.0	rth	42.0	th	52.0
ave.	83.0	ave.	81.0	ave.	51.5	ave.	76.0
s	57.3	sl	72.7	nd	92.7	d	91.0
b	89.0	bl	95.3	nj	91.3	j	96.0
g	96.3	gl	77.3	nz	82.0	z	76.7
p	78.3	pl	68.0	nk	92.7	k	90.7
k	95.3	kl	86.7	nt	84.7	t	88.7
f	56.3	fl	86.7	ns	72.7	s	44.7
ave.	78.7	ave.	81.1	ave.	86.0	ave.	81.3
Init. Ave.	81.1		81.0	Fin. Ave.	66.3		78.3

Ave. Simple Cons.................79.7
" Comp. " ....73.5 Exclusive of r () 82.7

The articulation of the consonant compounds as a class does not differ appreciably from the articulation of the corresponding simple consonants. The final r compound is seen to be an exception to this general rule. The errors for combinations containing this sound were caused by the large number of omissions of the modifier. For example, if "barb" were called, "bab" would be recorded. When the final r is combined with a consonant, the tendency is to shorten its duration and to stress it less than is done when it occurs as a simple final. Also, as mentioned before, the r sound materially modifies the vowel preceding it and usually in such a way that the vowel and r sounds are spoken as a vowel. For these two reasons, it escapes detection more readily than when used as a simple final. It will be noticed from the table that f is definitely more difficult to recognize than fl, while p is definitely less difficult to recognize than its compounds pl and pr. There is also a large difference between the results for s and ns. Although these differences are large, some tend to increase and others to decrease the average articulation. It is seen from the table that if the r compounds are omitted, the averages for the simple consonant sounds and for their compounds are approximately equal. Since this class of sounds comprises less than 15 per cent of the speech sounds, the results obtained by using a list in which the consonant compounds are omitted will be very closely the same as those obtained by lists in which such sounds occur. In view of this, and also because their inclusion would greatly extend the time needed for testing, compound consonants have been omitted.

In conversational or written speech some of the sounds are used much more frequently than others, whereas in the testing lists each sound is used the same number of times. Does this procedure lead to essentially different articulation values, for the various sounds, from those obtained by using the sounds in proportion to their frequencies of occurrence in speech?

TABLE VIII

ARTICULATION OF SOUNDS OF EQUAL VS. UNEQUAL OCCURRENCE

Sound	Equal Occ.		Unequal Occ.	
	No. of Occur.	Art.	No. of Occur.	Art.
a	300	91.0	550	88.0
ā	300	98.3	600	96.5
a'	300	88.3	400	82.3
e	300	91.0	1200	87.4
ē	300	100.0	850	98.8
i	300	96.0	1150	94.5
o	300	86.0	500	90.6
ō	300	98.3	500	97.6
o'	300	97.7	450	96.2
u	300	88.3	150	92.7
ū	300	97.3	250	97.6
Ave.		93.8		92.9
b	300	87.0	400	91.0
ch	300	97.3	150	98.0
d	300	91.0	700	89.1
f	300	73.7	400	67.8
g	300	92.0	450	94.2
h	150	100.0	200	100.0
j	300	98.0	150	97.3
k	300	87.3	900	89.8
l	300	98.7	1200	96.2
m	300	95.3	900	81.3
n	300	93.7	1300	95.5
ng	150	93.3	400	90.0
p	300	75.3	500	71.2
r	300	97.0	1250	97.4
s	300	72.3	850	76.3
sh	300	99.3	300	99.7
st	150	96.0	200	96.5
t	300	83.7	1550	86.6
th	300	60.3	200	60.5
th'	300	70.7	150	68.0
v	300	78.7	250	71.6
w	150	96.7	400	97.8
y	150	99.3	100	94.0
z	300	67.0	250	72.0
zh	150	97.3	50	100.0
Ave.		88.0		87.3

Table VIII gives the results of articulation tests that were made with two such types of lists. In both cases the sounds were combined at random into syllables of the con-vow-con type. The tests were made on the auxiliary circuit of the master reference system.

Realizing that the probable error in the articulation value given for each sound is $\pm$ 5, there do not appear to be any outstanding differences in the articulations of the various sounds with the two types of list. The average articulations for the two lists differ by less than the probable error. The test indicate, therefore, that lists having uniform occurrence of sounds give the same individual sound articulation values as lists having the frequencies of occurrence of the sounds proportional to their frequencies of occurrence in speech. At least this is true within the accuracy usually attained in making such tests. The testing advantages of the former type of list have already been pointed out.

It is important to notice that the average sound or the average syllable articulation may not be the same for the two types of lists. even though the articulation for each sound is the same. The averages shown in the table were obtained by assigning equal weights to the articulation for each fundamental sound. If weights which are proportional to frequency of occurrence of the sounds in speech be assigned, the averages obtained will, in general, be slightly different. For the particular circuit corresponding to the data of Table VIII, the averages obtained in the two ways did not differ by more than the observational error. Our data have shown that this is also true for a large class of circuits ordinarily used in telephone work. However, those transmission systems which have a specific effect upon certain consonant or vowel sounds, for example, upon s which occurs 850 times in one list compared to 300 times in the other, would obviously have different values for the sound articulation by using the two methods of obtaining the average.

In speech, certain combinations of sounds occur more frequently than others. In other words, some consonants precede certain vowels more frequently than they do other vowels, and similarly, some consonants follow certain vowels more frequently than others. For example, the combination "es" is used much more frequently than the combination "us" (u as in foot). Since the testing lists are made by random selection, the various con-vow and vow-con combinations occur with uniform frequency. In order to determine how this difference influences the interpretation of the sounds, articulation data on various circuits were examined. Attention was focused first on the final consonant sounds. One hundred errors for each consonant, or 2,200 consonant errors were selected at random from the articulation data and the number of these 2,200 errors that occurred after each vowel sound ascertained. Similarly, the number of vowel errors, out of a total of 1,100 errors, that occurred after each of the consonant sounds, was determined.

Probability studies indicate that the distribution of these errors as shown in Table IX is of the same order as that to be expected on

TABLE IX

Distribution of Vowel and Consonant Errors

Distribution of Vow. Errors		Distribution of Fin. Con. Errors	
Preceding Con. Sound	No. of Vow. Errors	Preceding Vow. Sound	No. of Fin. Con. Errors
b	58	a	188
ch	73	ā	264
d	52	a'	232
f	49	e	166
g	50	ē	214
h	55	i	172
j	54	o	192
k	60	ō	192
l	45	o'	208
m	49	u	196
n	54	ū	176
p	68		
r	39		
s	50		
sh	53		
th	38		
t	52		
v	40		
w	52		
y	49		
z	60		

the basis that the distribution of errors is due entirely to chance. Since, from the way the lists were constructed, the occurrence distribution is due to chance, it is evident that the errors in recognition of the sounds do not depend upon the particular sounds that they follow. Although the analysis was not made, it would be expected that a similar situation obtains for initial consonant and vowel errors. These data may be interpreted to mean that the consonant articulation, the vowel articulation, the sound articulation, or the syllable articulation, is approximately independent of the particular sound combinations, when a wide variety of combinations are used. The results obtained with these lists, therefore, are as representative of speech as the results that would be obtained with lists employing particular sound combinations in proportion to their frequencies of occurrence in speech. The analysis was not extensive enough to draw conclusions as to the effects of particular sound combinations upon the articulation of individual speech sounds.

Approximately 40 per cent of the syllables that occur in English are of the con-vow-con type. About 34 per cent are of the con-vow, and vow-con type. The syllables, including the compounds, such as, con-con-vow, vow-con-con, con-vow-con-con, and con-con-vow-con, make up about 16 per cent of the syllables of English. Since, as pointed out above, the interpretation of the consonant compounds depends primarily upon only one of the consonants, the latter syllables may be grouped in the two former classes, which then constitute some 90 per cent of English. Of the remaining syllables, 7 per cent consist of a single vowel, so that the more complex syllable forms constitute only 3 per cent of English.[8] Since 97 per cent of the syllables of English are included in the one, two and three letter forms, there is little reason to include the more complex syllable forms in order to represent speech, when as has been previously stated, they are undesirable from a testing standpoint. As will be shown in a later paragraph, one, two and three-letter syllables all yield equal values of articulation for the various speech sounds. Since the three-letter syllables require a smaller testing time for a given number of called sounds, the other syllable forms were excluded from the testing lists.

Having shown that the standard technique gives, for the various sounds, data that are representative of speech, the question now arises as to the best figure that may be computed from the data obtained with this technique, in order to best represent the speech transmission ability of the system under test. Before discussing this, it is necessary to consider some probability relations existing between the quantities entering into the calculation of such a figure.

Statistical Relations

The syllable articulation S when expressed as the ratio of the number of successes (correct interpretations of the syllables) to the number of trials (syllables called) is the chance of perceiving a syllable correctly. Also, if a similar ratio is used for the sound articulation L, the vowel articulation V, and the consonant articulation C, then these letters represent the probability of perceiving correctly a fundamental sound, a vowel sound or a consonent sound, respectively. If a syllable contains only one fundamental sound, then it is obvious that

$$S = L. \tag{3}$$

[8] These data were obtained from Godfrey Dewey's book "Relative Frequency of English Speech Sounds," Harvard University Press, 1923.

If a syllable has two letter sounds, then the chance of perceiving them both correctly is the same as the chance of perceiving the syllable correctly or

$$S = L^2. \tag{4}$$

Similarly, for a syllable containing m sounds

$$S = L^m. \tag{5}$$

Or if A_1, A_2, A_3, $A_4 \cdots A_m$ give the per cent of syllables in the list containing 1, 2, 3, 4, $\cdots$ m sounds, respectively

$$S = A_1L + A_2L^2 + A_3L^3 + \cdots A_mL^m. \tag{6}$$

Similarly, the chance of perceiving a syllable of the type con-vow or vow-con is VC; of the type con-vow-con, con-con-vow or vow-con-con is VC^2; of the type con-con-vow-con, con-vow-con-con, vow-con-con-con, or con-con-con-vow, is VC^3, etc.

For the old standard articulation lists these formulæ reduced to

$$S = \frac{1}{5}VC + \frac{4}{5}VC^2 = \frac{1}{5}L^2 + \frac{4}{5}L^3. \tag{7}$$

For the new standard articulation lists they reduce to [9]

$$S = VC^2 = L^3. \tag{8}$$

If a list of N syllables is used, then the letter errors and syllable errors will be $3N(1 - L)$ and $N(1 - L^3)$, respectively, or the number of letter errors per mistaken syllable, for the new standard lists, will be

$$m = \frac{3}{1 + L + L^2}. \tag{9}$$

It is seen that m approaches 3 as L becomes small, and unity as L approaches unity. For $L = .30$, $m = 2.06$; for $L = .50$, $m = 1.71$;

[9] When derived from the probability formulæ

$$VC^2 = L^3.$$

However, from the definition of V, C and L,

$$L = (2C + V)/3 \quad \text{so that} \quad VC^2 = L^3.$$

The difference is

$$d = \frac{(V + 8C)(V - C)^2}{27}.$$

Actually V and C are not wholly independent of each other and when values as obtained in tests are substituted in the above equation, the difference turns out to be small.

and for $L = .80$, $m = 1.23$. When observed values of m become consistently greater or less than this theoretical value, it must be concluded that the assumptions underlying this statistical theory are not valid.

All of the above statistical relations are dependent upon the tacit assumption that the chance of perceiving any sound correctly is entirely independent of the other sounds present and also independent of the number of other sounds present. It was shown in the previous section that the articulation of the various sounds is, on the average, independent of the other sounds in the syllables. On the other hand, experiments have indicated that the articulation does depend upon the number of sounds in the syllable. The sound articulation becomes smaller when the number of sounds in the syllables increases beyond three per syllable.

The data from which this conclusion was drawn were taken from three different experiments. In the first, three different transmission systems were tested by using first the standard articulation lists and then the vowel-consonant lists which are described in the last section. When using the vowel list, the vowels only are considered and when using the consonant list the consonants only are considered. These lists together, then may be considered as composed of syllables having only one sound. The syllable and sound articulations are the same when using such lists. The comparison of the results obtained with the two types of lists is shown in Table X. It will be seen that there

TABLE X

Articulation for One- and Three-Sound Syllables

	Freq. below 1,000~ only			Freq. below 1,950~ only			Freq. above 1,500~ only		
	Vow. Art.	Cons. Art.	Sound Art.	Vow. Art.	Cons. Art.	Sound Art.	Vow. Art.	Cons. Art.	Sound Art.
One-sound Syllables	71.5	62	65	98.5	82.5	88	81	96	91
Three-sound Syllables.	69.5	61.5	64	96	83	87	80	96.5	91

is only a slight tendency for the sound articulation to be lower for the three-sound syllable when compared with the one-sound syllable. The differences are within the observational error in testing.

In the second experiment a new list was constructed using the syllable forms con-vow and vow-con. The auxiliary circuit of the master reference system was tested with this list and also with the standard articulation list. The results are shown in Table XI. The

TABLE XI

ARTICULATIONS FOR TWO- AND THREE-SOUND SYLLABLES

	V	*C*	*L*
Two-sound Syllables	95	88	90
Three-sound Syllables	94	87	89

number of syllables used of each type for determining these averages was 1,344. The average sound articulation in each case was determined by giving equal weights to the articulation for each sound. For the three-sound syllables this is done by dividing the number of sounds correctly recognized by the total number of sounds called. For the two-sound syllable the procedure is not so simple. Since each vowel sound occurs twice as often as each consonant sound, it is necessary to obtain an average for the vowels and consonants

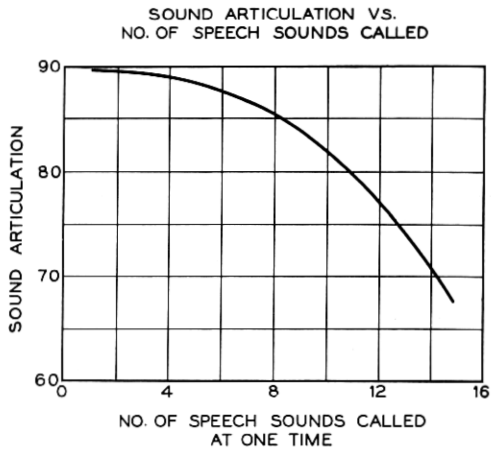


Fig. 8

separately. The final average value for the sound articulation is obtained by assigning weights of 1 and 2 to the vowel and consonant articulations, respectively. It is seen that there is no appreciable difference between the values of L obtained by the two types of lists.

In the third experiment the standard lists were used to test the auxiliary circuit but the syllables were called in groups of 1, 2, 3, 4, or 5 at a time. The results of these tests are shown in Table XII.

TABLE XII

Number of Sounds Called at One Time	Sound Articulation
3	89.0
6	87.5
9	84.0
12	77.0
15	67.0

From these three sets of data and from other available data which could be applied to this problem, the curve shown in Fig. 8 was constructed. It gives the sound articulation which would be obtained for a circuit such as the auxiliary circuit of the master reference system, when the number of sounds that are spoken at a time, that is, before the observer starts writing, is represented by the abscissa. It is evident from the shape of this curve that the assumptions underlying the statistical formulæ are valid for syllables having three or less sounds per syllable, and that they will break down for the more complex types of syllables. These assumptions might be expected to break down also, for certain extreme types of distortion.

Definite relations between the vowel, consonant, sound, and syllable articulations for both the old and the new techniques, have been derived by statistical theory. An experimental relationship between these quantities is shown in Figs. 9 and 10. These were obtained by an analysis of the errors of a large number of tests with widely different types of distortion, the data in Fig. 9 being taken with the old and the data in Fig. 10 with the new technique.

In the figures observed values of sound articulation have been plotted against the corresponding observed syllable articulation values. The solid curves in the two cases were calculated from Equations 7 and 8, respectively. The observed values agree reasonably well with the theoretical curves.

There is very little correspondence between the vowel and syllable or consonant and syllable articulation. The table below shows that

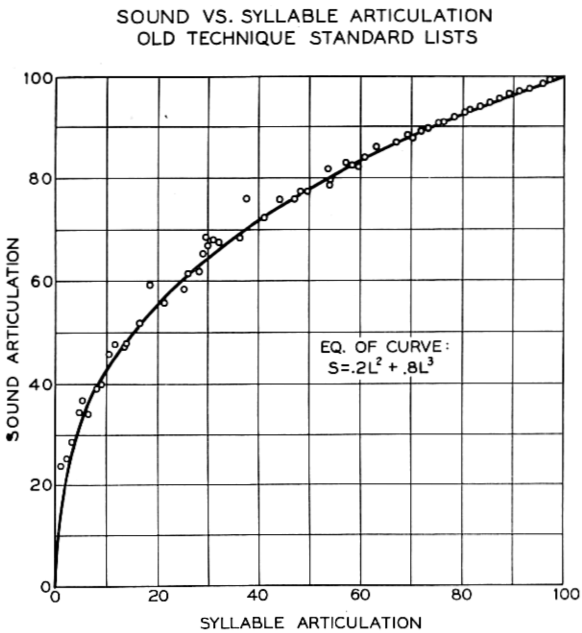
TABLE XIII

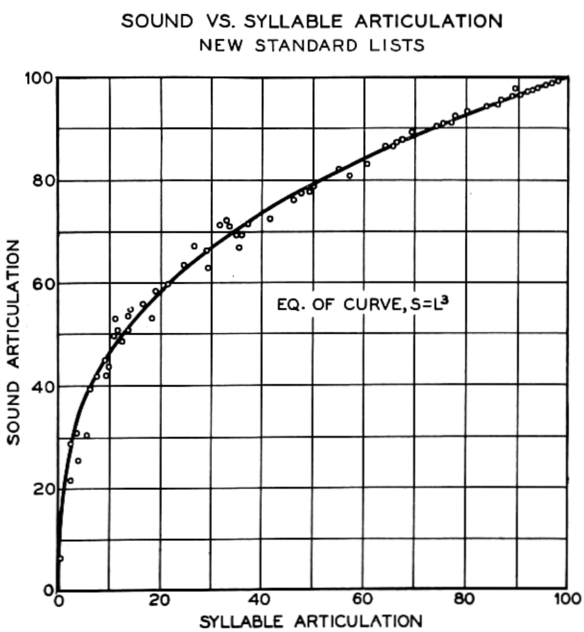
VOWEL, CONSONANT, AND SYLLABLE ARTICULATIONS

	V	C	S	VC^2
3,750 L.P.F.	98.7	95.4	90.8	89.8
750 H.P.F.	93.0	98.6	90.9	90.4
2,850 L.P.F.	98.5	92.9	86.3	85.4
1,000 H.P.F.	89.0	98.4	87.0	86.2

a circuit which discriminates against the vowels may have a syllable articulation equal to another circuit which discriminates against the consonants. However, it is seen that the product VC^2 is equal to S as the statistical theory indicates. If then *the sound, or vowel and consonant articulations are known, it is possible to calculate the syllable articulation*, for the case of two- and three-sound syllables.

We are now in a position to consider the figure which best represents





Figs. 9 and 10—Relation between sound and syllable articulation the capabilities of systems to transfer thought by means of speech. For giving a complete picture, it is necessary to give the articulation values for each speech sound. Since this involves 36 articulation values, it is difficult to compare various systems. To combine these values into averages raises the question of how such an average shall be taken. At first thought it might seem obvious that the weights assigned to each sound articulation value should be proportional to the frequency of occurrence of that sound in English speech. Many of the most frequently occurring words, however, such as the, of, and, to, in, that, etc., do not carry much of the thought, so that it seems reasonable to exclude the effects of such words in the weighting process. It is evident that many sets of weighting factors could be evolved depending upon how far the exclusion process is carried and depending upon whether written or spoken English is used, in determining the frequencies of occurrence of the sounds. After excluding the twenty or twenty-five most common words, however, further exclusion does not appreciably change the calculated articulation value. The table below gives a set of factors obtained from the frequencies of occurrence used in Table VIII. They are based upon the studies of Messrs. French and Koenig [10] on the frequencies of occurrence of speech sounds in spoken English. The effects of the more common parts of speech, such as, personal pronouns, definite articles, conjunctions, and prepositions have been excluded.

TABLE XIV

Group I	Weight	Group II	Weight	Group III	Weight	Group IV	Weight	Group V	Weight
ā	3.0	i	5.8	r	6.3	d	3.5	z	1.3
ē	4.3	o	2.5	l	6.1	t	7.8	s	4.3
ō	2.5	a'	2.0	ng	2.0	b	2.0	v	1.3
ū	1.3	u	.8	n	6.6	p	2.5	f	2.0
o'	2.3	e	6.0	m	4.5	g	2.3	zh	.3
a	2.8	y	.5			k	4.5	sh	1.5
		w	2.0			j	.8	th'	.8
						ch	.8	th	1.0
						h	1.0	st	1.0
Total Weight	16.2		19.6		25.5		25.2		13.5

It will be noticed that the speech sounds are arranged in five groups. The sounds in each group have very similar characteristics, so instead of dealing with 36 articulation values for a circuit, it is only necessary

[10] "Frequency of Occurrence of Speech Sounds in Spoken English," N. R. French & W. Koenig, *Proc. Acoustical Society of America*, 1929.

to deal with the average value for each of the five groups. The average for the first group is designated V_l, signifying long-vowel index; for the second group V_s, signifying short-vowel index; for the third group C_n, signifying nasalized-consonant index; for the fourth group C_s, signifying stop-consonant index; for the fifth group C_f, signifying fricative-consonant index. If the articulation obtained from any test for each sound be designated by the phonetic symbol for that sound, then,

$$\left.\begin{aligned} V_l &= .19\ \text{ā} + .27\ \text{ē} + .15\ \text{ō} + .08\ \text{ū} + .14\ \text{o}' + .17\ \text{a} \\ V_s &= .30\ \text{i} + .13\ \text{o} + .10\ \text{a}' + .04\ \text{u} + .31\ \text{e} + .02\ \text{y} + .10\ \text{w} \\ C_n &= .25\ \text{r} + .24\ \text{l} + .08\ \text{ng} + .26\ \text{n} + .17\ \text{m} \\ C_s &= .14\ \text{d} + .31\ \text{t} + .08\ \text{b} + .10\ \text{p} + .09\ \text{g} + .18\ \text{k} + .03\ \text{j} \\ &\qquad + .03\ \text{ch} + .04\ \text{p} \\ C_f &= .10\ \text{z} + .32\ \text{s} + .10\ \text{v} + .15\ \text{f} + .02\ \text{zh} + .11\ \text{sh} \\ &\qquad + .06\ \text{th}' + .07\ \text{th} + .07\ \text{st}. \end{aligned}\right\} \qquad (10)$$

The sound index is related to these values by

$$i = .162\ V_l + .196\ V_s + .255\ C_n + .252\ C_s + .135\ C_f. \qquad (11)$$

For obtaining the most representative single value for the syllable index I, the equation given below is used.

$$I = .5\ i^2 + .5\ i^3. \qquad (12)$$

This equation is based upon the frequency of occurrence of the syllable forms in English speech. As pointed out before, if the compound consonants be considered as simple sounds, then there are less than 10 per cent of syllable forms other than the two- and three-sound type. The frequency of occurrence of these two types is approximately equal.

Similar formulæ to the above may be used to relate articulation results in English to articulation results in a different language. To do this it is necessary to select the fundamental sounds of the different languages that correspond to the 36 fundamental sounds of English, where the correspondence is based on similar phonetic characteristics and similar positions of the vocal organs in producing the sounds. When this is done, the coefficients in Eqs. 10, 11 and 12, must be modified to correspond with the frequencies of occurrence of the sounds and syllables in the language.

Observed values of individual sound articulation are thus reduced to a single index, or for a more comprehensive picture, to five indices corresponding to the five groups of speech sounds. In order to compare the indices obtained by a given crew with those of a reference crew, it is necessary to correct the data in accordance with Eq. 2 for the effects of practice. To do this, as previously discussed, articulation tests are made upon one or more of the reference circuits by the crew in question. If I' is the syllable index so obtained, the practice factor for the crew is given by the relation

$$x = \frac{\log (1 - I')}{\log (1 - I)}. \tag{13}$$

The practice factors for the other indices may be obtained also, by substituting the appropriate indices for the syllable index in Eq. 13. In Table XV the reference values for the various indices are given for the reference circuits that were previously described.

TABLE XV

REFERENCE VALUES

Circuit	V_l	V_s	C_n	C_s	C_f	i	I
Master Reference System	98.5	98.9	99.6	99.2	98.8	99.3	98.0
Auxiliary Circuit of Master Ref. Sys.	95.0	95.0	96.5	88.5	66.5	90.0	77.0
Carbon Transmitter Circuit	97.0	97.0	96.5	93.5	82.0	94.0	85.5
Master Ref. Sys. plus 3,750~ L.P.F.	99.0	99.5	99.5	99.0	86.5	97.6	94.0
" " " " 750~ H.P.F.	96.0	92.5	99.0	99.5	98.5	97.1	93.0
" " " " 1,500~ L.P.F.	93.5	86.5	90.5	76.0	52.5	80.2	58.0
" " " " 1,500~ H.P.F.	85.0	82.5	96.0	97.5	97.0	92.0	81.0

If the values for the sound index be compared with the sound articulation values based on uniform weighting, that were given under the section on practice effects, it will be seen that for these circuits there is very little difference, between the two sets of values. In other words, the average sound articulation is very nearly equal to the average that is obtained when the individual sound articulations are weighted according to the frequencies of occurrence of the sounds in English.

Similar comparisons have been made for a large number of other transmission systems. They showed similar small differences between the weighted and unweighted averages. For this reason we consider it unnecessary to use the weighted average when great accuracy is not required, for example, in a great deal of our routine work where comparisons are being made between circuits which have similar characteristics. This means that when testing an unknown circuit, having an electrical characteristic similar to one of the reference circuits, the syllable index I can be calculated from the observed syllable articulation S (as obtained with the new standard lists) by means of the equation,

$$I = .5\ S^{2/3} + .5\ S. \tag{14}$$

This value must now be reduced to the reference condition of practice by the methods which have already been described. In such cases it is thus possible to obtain the syllable index from the observed syllable articulation values, and it is unnecessary to analyze the data for the individual sound articulation values.

The weighted average, however, is the more logical way of obtaining a single index and should be used when it is suspected that it might give results which are essentially different from the unweighted average.

It is possible to carry the probability relations a step further and apply them to cases of English words and sentences. In order to do this it is necessary to make assumptions as to how the thought or meaning of the words affects the interpretation of the sounds. These assumptions are not only somewhat uncertain, but owing to psychological factors in testing are difficult to verify experimentally. In general, the meaning associated with words makes them easier to interpret than meaningless words. For single-syllable words, these effects are small. Two-syllable words are easier to interpret than single-syllable words. The interpretation of words containing from three to five syllables, and short sentences, depends almost entirely upon interpreting those parts which are not indicated by the thought or meaning.

Other Testing Methods

For most articulation studies it has been found desirable to use the standard testing technique which has been described, but it is frequently necessary, in special cases, to use other techniques. In such cases it is desirable, if possible to interpret the results in terms of the standard technique. In the course of research work, several different articulation testing methods have been used which give information on the type of correlation between them that may be expected.

The probability relations have been made use of in constructing two other types of lists which are called vowel-consonant and vowel word-consonant word lists. These lists are designed to give the same values of sound articulation as given by the standard lists. The former lists are shown in Table XVI. The various vowels are combined with the same consonant, and the various consonants with the

TABLE XVI

VOWEL LIST

Sound to be Graded	Testing Syllables in the List	
a	at	ta
ā	āt	tā
a'	a't	ta'
e	et	te
ē	ēt	tē
i	it	ti
o	ot	to
ō	ōt	tō
o'	o't	to'
u	ut	tu
ū	ūt	tū

CONSONANT LIST

Sound to be Graded	Testing Syllables in the List					
b	bū	ūb	ba	ab	bē	ēb
d	dū	ūd	da	ad	dē	ēd
f	fū	ūf	fa	af	fē	ēf
g	gū	ūg	ga	ag	gē	ēg
k	kū	ūk	ka	ak	kē	ēk
l	lū	ūl	la	al	lē	ēl
m	mū	ūm	ma	am	mē	ēm
n	nū	ūn	na	an	nē	ēn
r	rū	ūr	ra	ar	rē	ēr
p	pū	ūp	pa	ap	pē	ēp
s	sū	ūs	sa	as	sē	ēs
sh	shū	ūsh	sha	ash	shē	ēsh
th'	th'ū	ūth'	th'a	ath'	th'ē	ēth'
th	thū	ūth	tha	ath	thē	ēth
t	tū	ūt	ta	at	tē	ēt
v	vū	ūv	va	av	vē	ēv
ch	chū	ūch	cha	ach	chē	ēch
z	zū	ūz	za	az	zē	ēz
j	jū	ūj	ja	aj	jē	ēj
h	hū		ha		hē	
w	wū		wa		wē	
y	yū		ya		yē	
zh		ūzh		azh		ēzh
ng		ūng		ang		ēng
st.		ūst		ast		ēst

same vowel. The technique of using the list is the same as that previously described, except that the vowel articulation and consonant articulation are measured separately. Only the vowel errors are counted when using the vowel list and only the consonant errors when using the consonant list. These lists have the advantage that they can be used over and over by merely changing the sequence of the syllables.

Table XVII shows two lists similar to the above except that they are made up entirely of common English words. They are designated as vowel word and consonant word lists. This list is used in the same way as the vowel and consonant lists. In using either of these lists the testing crews should be familiar with the syllables or words in the lists.

TABLE XVII

VOWEL WORD LIST (ENGLISH WORDS)

Sound to be Graded	English Words in the List	
a'	bat	back
ā	bait	bake
e	bet	beck
ē	beat	beak
i	bit	bit
ī	bite	bike
o	but	buck
o'	bought	balk
ō	boat	boat
u	book	book
ū	boot	boot

CONSONANT WORD LIST (ENGLISH WORDS)

Sound to be Graded	English Words in the List	
b	by	by
ch	which	which
d	die	die
f	fie	whiff
g	guy	wig
h	high	high
j		
k	wick	wick
l	lie	will
m	my	whim
n	nigh	win
ng	wing	wing
p	pie	whip
r	wry	wry
s	sigh	sigh
sh	shy	wish
th'	thy	with
th	thigh	thigh
t	tie	wit
v	vie	vie
w	why	why
y		
z	whiz	whiz
st	sty	whist

Note: The h following w is not pronounced in such words as whim, whip, etc.

It usually requires a training period of a month or more for a testing crew to thoroughly master the technique of using the standard lists, that is, to reach a stage where the phonetic symbols are spoken and recorded almost mechanically. The vowel consonant lists require less time, since it is only necessary for the observers to fix their attention on one sound in the syllable. With the word lists this training period is reduced to a minimum. Phonetic symbols are avoided, and attention is given to only one sound in the words. As may be seen from Fig. 2, after a few tests they practically reach a degree of uniform proficiency. In using the lists, the words are recorded with the English spelling. Only errors in the vowel and consonant sounds of the left-hand column of the above table, are counted. Since only one sound in each syllable is utilized, the above lists require a somewhat greater testing time for a given precision than do the standard lists where all three sounds of the syllables are used.

Table XVIII below, shows data that were obtained with the three types of lists, namely, the standard lists, the vowel consonant lists, and the vowel word consonant word lists. The vowel consonant

TABLE XVIII

ARTICULATION RESULTS WITH VARIOUS LISTS

Circuit	Cons. Vow. List		Cons. Word Vow. Word List		Standard List			Corrected
	C	V	C_w	V_w	S	VC^2	$V_wC_w^2$	$V_wC_w^2$
Air Transmission	99.5	99.0	99.0	99.5	98.0 ± .4	98.0	98.0	97.5
Master Ref. System + 5500 L.P. Fil.	99.0	99.0	99.0	99.5	96.5 ± 1.0	97.0	97.5	97.0
M.R.S. + 3750 L.P. Fil.	95.5	98.5	96.5	99.0	90.5 ± 2.0	90.0	92.0	90.0
M.R.S. + 1950 L.P. Fil.	82.5	98.5	85.5	98.5	67.0 ± 2.0	67.5	72.0	68.0
M.R.S. + 1000 L.P. Fil.	62.0	71.5	67.0	82.0	29.0 ± 3.0	27.5	37.0	33.0
M.R.S. + 1500 H.P. Fil.	96.0	81.0	97.0	86.0	75.0 ± 2.0	75.0	81.0	77.0
Carbon Transmitter Circuit	91.0	96.0	91.5	99.0	81.0 ± 1.5	80.0	83.0	80.0

lists lead to syllable articulations that are essentially the same as obtained with the standard lists. It is evident from the data that the word lists lead to slightly higher values of syllable articulation. The explanation is to be found in the make up of the word list. It was not possible to arrange the lists so that all sounds are equally probable, and still use English words. For instance, an observer would not record "bik" for "beck," nor "wiv" for "with," etc. For this reason, the observers do not make errors which occur frequently in the other two types of lists, where any sound is possible. Hence the observed articulations are somewhat higher with the word lists.

It was found, however, that the word lists may be correlated with the standard lists by means of the relation given by Equation 1. The value of x for this case is 0.9 so that, the word technique may be corrected to the standard technique, by means of the equation.

$$S = 1 - (1 - V_wC_w^2)^{.9}, \tag{15}$$

where S = syl. art. of standard lists expressed as a ratio,
V_w = vowel art. of vowel word lists expressed as a ratio,
C_w = cons. art. of cons. word list expressed as a ratio.

The corrected values are also shown in the above table.

It is frequently necessary to test very poor systems where the standard lists giving an articulation of a few per cent, are not satisfactory. The vowel consonant lists are somewhat more satisfactory under these circumstances. Lists of sentences have also been found to be very useful for such purposes. The sentences were of the interrogative or imperative form containing a simple idea. They were designed to test the observer's acuteness of perception rather than his intelligence. Tests were made with these sentences and the standard lists on various circuits, involving carbon transmitter circuits and various filter systems. The data are shown in Fig. 11. The

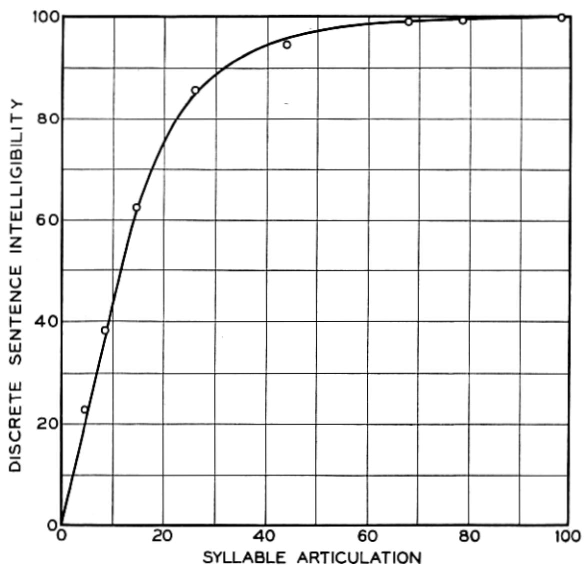


Fig. 11—Discrete sentence intelligibility vs. articulation

sentences were considered to be understood if the observer either recorded the sentence correctly or recorded an intelligent answer. As stated earlier, the percentage correctly observed is called the discrete sentence intelligibility.

It will be seen that for changes in distortion, the changes in the discrete sentence intelligibility will be small for systems having syllable articulations greater than 30 per cent, but very large for systems having syllable articulations below 20 per cent. It is for systems in this latter class that these test sentences are useful. A case in point is the measurement of the degree of secrecy obtained in sound proofing telephone booths, or in dealing with cross-talk. The sentences have also been found to be useful in making quick qualitative tests of the goodness of an audiphone set for a particular case of deafness.

Because of their general usefulness for these purposes, the complete lists of sentences are given in the appendix. Due to memory effects a set of sentences can be used with the same personnel only a very few times. The psychological factors are also more prominent with sentences than with simple syllable.

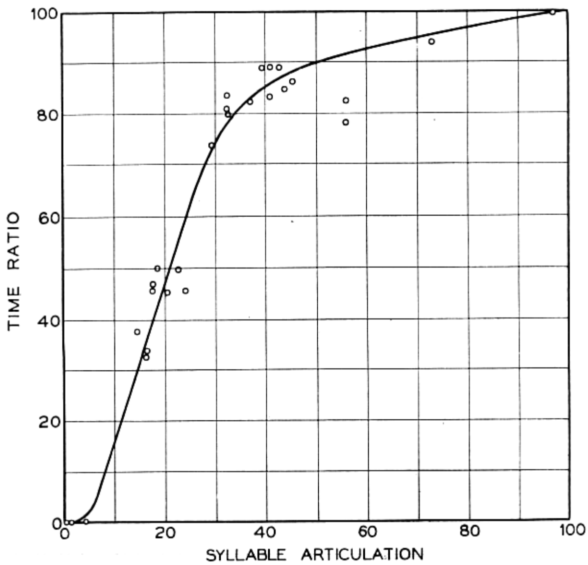


Fig. 12—Articulation vs. time ratio

Sentence lists of the above type have also been used to obtain a notion of how the time taken to transmit an idea correctly over a system depends upon the articulation. To do this, the observer was instructed to reply orally to the question. If the reply indicated that the observer failed to understand, the speaker repeated the question. Both speaker and observer tried to carry out the test in a normal conversational manner. The observer could ask the speaker to repeat, reword or spell out difficult parts of the sentence.

The tests were made on a variety of systems of known syllable articulation. The results that were obtained are shown in Fig. 12. The ordinates of the curve give the ratio of the time required to transmit correctly one of these test sentences over an ideal system to the time required over the system under test. With the crew used in making these tests, and with an ideal transmission system, it required an average time of 5.2 seconds after the speaker started to pronounce the sentence before the observers grasped the idea. It will be seen from the curve that for systems having approximately 20 per cent articulation, the time required is twice as great. Fig. 11 shows that one out of every four of the sentences is mistaken for this value of articulation. If it is assumed that an observer asks that only sentences which he fails to understand be repeated, it can be shown that this time ratio is equal to the discrete sentence intelligibility.[11]

It is evident from Figs. 11 and 12, that the observed time ratio is appreciably less than the discrete sentence intelligibility. This difference may be taken to indicate that an observer not only asks that sentences which he fails to understand be repeated, but also that sentences about which he is uncertain be repeated. In other words, the time element reflects both factors, the understandability and the uncertainty.

As has been previously mentioned, tests have been made with various types of English word lists. Because of the manner in which the words were selected, and also due to uncertain psychological factors entering into the tests when such words are used repeatedly, it is difficult to compare the results so obtained with syllable articulation results.

However, it was found that if a definite rule were followed in selecting words from a newspaper, consistent results could be obtained with lists containing 500 or more words per list. The method of selection was to take the first word from every third line of a newspaper column. In this selection all proper names and the following six most frequent words of English were excluded, the, of, and, to a, in. When a word was hyphenated from the previous line, the whole word was used. Each of eight callers called a list of 66 words to four observers in the manner of an ordinary standard articulation test. Tests were made with the carbon transmitter circuit and the six circuits indicated in Fig. 3. The data were analyzed to give the discrete word intelligibilities for the one, two, three, four, and five-syllable words occurring in the lists, as well as for the lists as a whole. The lists on the average contained 46.3 per cent one-syllable, 29 per cent two-syllable, 16.8 per cent three-syllable, 6.4 per cent four-

[11] "A Theoretical Study of Articulation and Intelligibility of a Telephone Circuit," John Collard, *Electrical Communication*, 7, page 168, January, 1929.

syllable, and 1.5 per cent five-syllable words, and an average number of two syllables per word. The discrete word intelligibility vs. syllable articulation as obtained with the standard lists is shown in Fig. 13. The dashed curves indicate the relations for the various types of words, and the solid curve for the word lists as a whole. The data for two syllable words practically coincided with the solid curve. Owing to the small amount of data, the curves for the four- and five-syllable words are less reliable than those for the other types. Curves of the above type, both for words and sentences, depend very

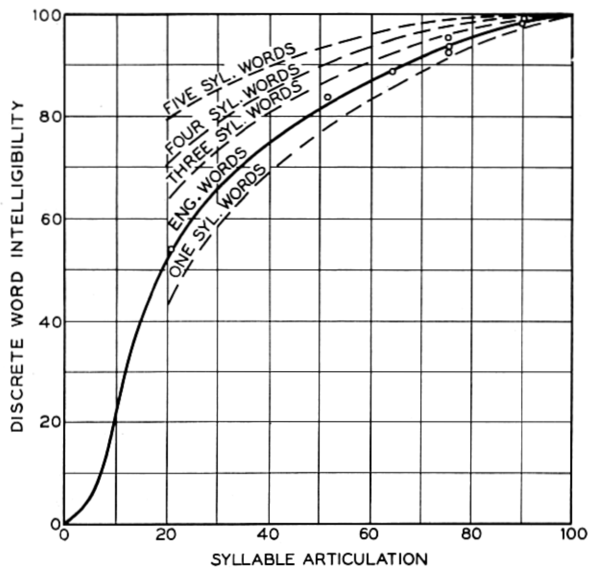


Fig. 13—Discrete word intelligibility vs. syllable articulation

much upon the way the speech material is selected. If, for example, only "different" words had been included in the word lists, appreciably smaller values of discrete word intelligibility would have been obtained.

Tests have also been made with lists made up of the following numbers, 1, 2, 3, 4, 5, 6, 8. These numbers were combined at random into groups of three and called in the manner of an ordinary articulation syllable. The distinguishing characteristic of each of the above numbers is a vowel sound, so that, they are interpreted primarily from recognizing the vowel. Such lists, therefore, do not give a very good picture of the speech capabilities of a system which distorts speech. They are, however, very useful in measuring the deafness of an observer, for the reason that the number articulation decreases very rapidly as the sounds approach the threshold of hearing. As may be seen from Fig. 14, the number articulation passes from practically 100 per cent to 0 per cent in the short range of 10 or 15 db. It is evident that such lists give a critical measure of the point at which

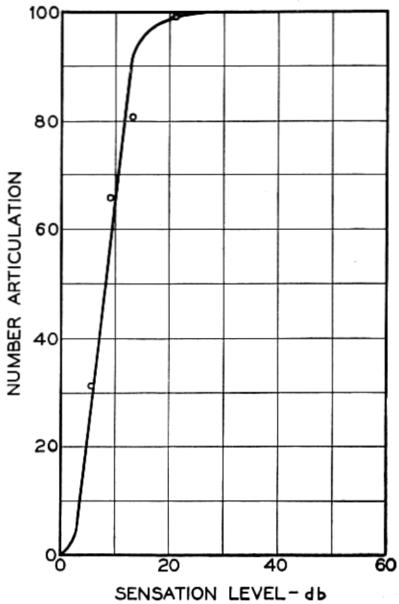


Fig. 14—Number articulation vs. sensation level

an observer fails to hear the sounds. Lists of this type have been used extensively in testing the hearing of school children.

Summary

The standard testing technique is primarily a means of determining the articulation or recognizability of the individual speech sounds when they are spoken in a way that is representative of conversational speech, and in a way which facilitates the carrying out of articulation tests. The articulations of the individual sounds may be converted into an index which indicates the speech capabilities of a system. Other types of lists which yield either the recognizability of speech sounds, or the intelligibility of discrete English words and sentences containing thought, have been described and experimentally correlated with the syllable technique.

It should be emphasized that there may not be a one to one correspondence between all of these measured quantities for all types of speech distortion. The data entering into the sound, vowel, consonant, and syllable articulation curves were very extensive with respect to types of distortion and testing personnel. The theoretical equations relating them seem to rest upon assumptions with few uncertainties. For these reasons, it is felt that these relations can be used with considerable confidence, especially for values of syllable articulation greater than 30 per cent. The curves dealing with English words and sentences are based upon less diversified data and should be regarded as indicative only of the correlation and the type of relation.

During the past few years articulation testing methods have been used more and more, both in this country and abroad. It is felt that in order to compare the results obtained by various crews in various tongues, *it is desirable to use techniques that operate on the same basic principles and to calibrate various crews on similar reference circuits.*

INTELLIGIBILITY LIST

List 1

1. Name a prominent millionaire of the country.
2. How large is the sun compared with the earth?
3. Why are flagpoles surmounted by lightning rods?
4. Give the abbreviation for January and February.
5. Name the tree on which bananas grow.
6. How often does the century plant bloom?
7. What description can you give of the bottom of the ocean?
8. Explain the difference between a hill and a mountain.
9. What is the chief purpose of industrial strikes?
10. Describe the shoes of the native Hollander.
11. Name some uses to which electricity is put.
12. What would cause the air to escape from a bicycle tire?
13. Where is more grain raised, in the East or the West?
14. Tell what is meant by an Indian Reservation.
15. For what invention is Thomas Edison noted?
16. Name a state which has no seacoast.
17. Write the Roman numeral ten.
18. Explain the difference between export and import.
19. Explain why a corked bottle floats.
20. What substance is a good conductor of electricity?
21. Explain why Indians were afraid of firearms.
22. Explain the purpose of fire drills.
23. At what time do ocean waves become dangerous?
24. What medicine would you take to remedy indigestion?
25. What knowledge is covered by the study of astronomy?
26. Name a good restaurant in this vicinity.
27. What is the importance of large windows in stores?
28. Explain why a giraffe eats the foliage of trees.
29. How are the pages of a magazine held together?
30. Explain why the name string-bean is appropriate.
31. Name a nearby city in which there is a shipyard.

32. Name a fruit which grows in bunches.
33. Which of our Presidents went to South Africa?
34. Why are wire springs used in beds?
35. Why are books bound in stiff covers?
36. Why did the home people conserve food during the war?
37. Name an insect that has a hard shell.
38. What symbol on the United States money stands for liberty?
39. What weapons did the Indians use in warfare?
40. In what kind of weather does milk sour?
41. What streets in this city have Dutch names?
42. How does turning a ship's wheel steer the ship?
43. What nation aided us in the Revolutionary War?
44. What are some personal characteristics of the people of Japan?
45. What candy is black and good for colds?
46. Name a famous Indian Tribe.
47. Why is this building lighted by reflected light?
48. Why are most lighthouses situated on rocks?
49. Give some ingredients used in soap.
50. Why is a house built of stone superior to others?