*69. Zemen, K., "Speech Curves of 'hein,' Pronounced with Three Affective Accents," *Arch. f. d. ges. Psychol.*, 1928, 66, 493–496.
*70. Zipf, G. K., "Homogeneity and Heterogeneity in Language," *Psychol. Rec.*, 1938, 2, 347–367.
*71. ——, *The Psycho-Biology of Language*, Houghton Mifflin Co., Boston, 1935.
72. Zilboorg, G., "The Sense of Reality," *Psychoanalytic Quart.*, 1941, 10.

* References contain comprehensive bibliographies and/or descriptions of orientations in research in psychology and language.

LANGUAGE, THE PSYCHOLOGY OF.

—The present article will restrict itself to those phenomena of speech of which there are quantitative studies that indicate the existence of general dynamic principles.

I. The Problem of the Classification of Speech Entities

The scientific study of speech may be said to have started millennia ago when, for the sake of devising a system of writing, man first began to concern himself with the classification of the stream of speech into its significant entities. The history of writing (cf. article on "Alphabet," *Encyclopedia Brittanica*, 14th ed.), which is a part of the history of the psychology of language, reveals various more or less successful attempts in different regions to classify the local speech into its word elements, or its morphological elements, or its syllabic elements, or its phonetic elements, or even its morpho-semantic elements (e.g., Chinese). However, because the larger morphological, syllabic, and verbal elements of speech are intrinsically permutations of basic phonetic entities (e.g., *man* is a permutation of *m*, *a*, and *n*), the most practical classifications today are the alphabetic ones, of which the most extensive and accurate is the "narrow" *International Phonetic Alphabet* (of the International Phonetic Association). This alphabet classifies speech sounds essentially on an acoustic basis without concern about the particular phonetic differences in a given tongue that may be "significant" for that tongue in discriminating between its vocabulary of words.

As to the problem of determining "significant" phonetic differences within a given tongue, few solutions can compare with the simple method of *phonology* in the matter of operational neatness. This method consists of comparing pairs of words (i.e., *phonological oppositions*) that are homophonous except for different phonetic entities, or *phonemes*, in analogous positions (e.g., *pin, pan, pun,* or *pin, sin, tin*). By studying the *phonological oppositions* of a given vocabulary of words, one can disclose the vocabulary's stock of different *phonemes* which may be appropriately defined as "the minimal phonetic units that are used in the vocabulary to discriminate between otherwise homophonous pairs of words." In many languages this *phonological technique* may leave a small residue of ambiguous items for which no *phonological oppositions* can be found. In the present writer's experience this residue is statistically negligible.

Although *phonologists* in recent years have added enormously to our descriptive knowledge of many different languages (a running bibliography of studies on the topic can be found in *American Speech*), and although the discovery of the phoneme has greatly simplified the transcription of speech, nevertheless some claims made by some phonologists would seem to be unwarranted if taken literally in a truly scientific sense. Thus despite claims to the contrary, it has never been shown that phonemes represent *the* minimal units of distinctive significance in any sense other than that of keeping phonological oppositions apart. Second, it has been shown by Eberhard and Kurt Zwirner ("Phonometrischer Beitrag zur Frage der neuhochdeutschen Quantität," *Archiv für vergleichende Phonetik*, I [1937], 96–113) that despite the assertions of some phonologists two different phonemes in a given speech are not always kept rigidly apart (e.g., the phoneme, short *ă*, in German which differs from long *ā* only in the matter of duration, is sometimes pronounced longer than is long *ā;* in point of fact the difference between the two phonemes is solely in their statistical modes).

II. The Genetic Classification of Speech

Languages can also be classified on the basis of a common ancestry as ascertained by a comparison of forms and meanings. This work has long been conducted by *comparative philologists* (locally sometimes called *linguists*) who proceed on the basic assumption first empirically established by Karl Brugmann that *in a given language at a given time like phonetic entities under like conditions behave alike* (hereinafter referred to as *the orderliness of phonetic change*). For example, during a particular period

in Middle English all Old English words with a long *ā* became an open long *ō* which subsequently became the close long *ō* of today (e.g., Old English *stān, rāp, gāt* have become what we write today as *stone, rope, goat*). For most of the Indo-European languages, there are historical-comparative grammars and etymological dictionaries in abundance, of which perhaps the two best known are Karl Brugmann and B. Delbrück, *Grundriss der vergleichenden Grammatik der indogermanischen Sprachen,* 5 vols., 2nd ed. (Strassburg, 1897–1911), and Alois Walde and Julius Pokorny, *Vergleichendes Wörterbuch der indogermanischen Sprachen,* 3 vols. (Leipzig, 1927–1932). This vast store of carefully collected and objectively analyzed material has confirmed the principle of *the orderliness of phonetic change* beyond any doubt.

This principle is of practical value to the social psychologist in helping to disclose the date of adoption of foreign loan-words. Thus the fact that we pronounce the vowels of the Romance loan-words *dame* and *fame* like the vowels of *make* means that the two words were adopted in Middle English times *before* a certain phonetic change had occurred (cf. H. C. Wyld, *A Short History of English,* 3rd ed., New York, 1927). The fact that an initial *d* in Latin, such as in *duo,* two, or *decem,* ten, appears quite generally in even present-day English as a *t,* such as in *two,* two, and *ten,* ten, suggests (simply stated) that the words are not loan-words but *cognates* that have been inherited from a common ancestry with differences ascribable to independent phonetic changes that were peculiar to the histories of the respective languages.

III. Dynamic Principles of Speech

From the earliest days of history down to the present, man has produced speech-theories of varying degree of gravity of utterance that have not always been free from a suggestion of "magic," even in the case of many of the *ad hoc* speech theories of comparative philologists. In recent years, however, the problem of speech-dynamics has been approached with increasing scientific rigor.

A. *The Dynamics of a Phonetic System.* The French psychologist, R. Bourdon (in his *Des Emotions et des Tendences dans le Langage,* Paris, 1892), was apparently the first to observe that in many different languages the dental stops (e.g., *t* and *d*) are more frequent than either the labial stops (e.g., *p* and *b*) or the velar stops (e.g., *k* and *g*).

Subsequently G. K. Zipf argued (i) that the comparative difficulty of utterance of a phoneme in a given speech-community was inversely related to the relative frequency of its occurrence, and (ii) that the orderliness of phonetic change results from the tendency to preserve or to restore a dynamic equilibrium between the difficulty of utterance of phonemes and their relative frequency of occurrence. Although this second argument (ii) is too extensive to be included here (cf. G. K. Zipf, *The Psycho-Biology of Language,* Boston, 1935), nevertheless the nature of the data and theory of the first argument (i) can be indicated quite simply:

Thus if we take, for example, the twelve consonantal stops of Peipingese Chinese which consist of the six *more difficult* aspirated stops, t^h, p^h, k^h, $cç^h$, $t\wr^h$, ts^h, and the six *less difficult* corresponding unaspirated stops, *t, p, k, cç, t*$\wr$, *ts,* we find in samples of 37,338 running phonemes (Zipf, *op. cit.*) that each of the six *more difficult* aspirated stops is markedly less frequent than its corresponding *less difficult* stop, as is shown by the percentages of frequency of the whole of Table I:

Although the above data refer only to six pairs of phonemes in one dialect, nevertheless the statistical analysis of the frequencies of other types of phonemes in other languages reveals the same correlation. Thus, for example, the *more*

Table I

VOICELESS ASPIRATED FORTES AND VOICELESS UNASPIRATED LENES STOPS IN PRESENT-DAY PEIPINGESE

(Percentages in reference to occurrences of all speech-sounds)

	t^h/t	p^h/p	k^h/k	$cç^h/cç$	$t\wr^h/t$	ts^h/ts
Aspirated (*more* difficult)	2.56%	.56%	1.02%	1.04%	1.23%	1.40%
Unaspirated (*less* difficult)	6.18%	2.37%	2.58%	2.69%	2.44%	2.63%

difficult long vowels in a given speech are almost without exception much less frequent than their *less difficult* corresponding short vowels (if they are present in the language). So, too, the *more difficult* voiced stops,[1] *d, b, g,* are practically always much less frequent than their *less difficult* corresponding voiceless stops, *t, p, k,* as becomes evident from the percentages of the whole in the following dozen languages of Table II:

TABLE II

PERCENTAGE OF OCCURRENCES OF VOICED–VOICELESS STOPS

(Diphthongs counted as one unit)

	t	d	p	b	k	g
Czechish	5.60%	3.73%	3.52%	1.86%	3.93%	.15%
Dutch	7.83%	4.67%	1.99%	1.20%	3.21%*	.09%*
English	7.13%	4.31%	2.04%	1.81%	2.71%	.74%
Hungarian	7.18%	3.30%	1.04%	1.71%	5.72%	2.45%
Lithuanian	5.76%	2.61%	3.71%	1.35%	4.61%	1.36%
North Russian	7.97%	1.52%	3.36%	1.01%	3.36%	.67%
South Russian.	7.05%	2.46%	2.79%	1.51%	3.97%	1.66%
Wendish	6.26%	3.02%	2.55%	1.56%	3.29%	2.41%
East Ukrainian	3.83%	3.24%	2.82%	2.11%	4.11%	not present
Bulgarian	7.54%	3.55%	2.82%	1.32%	2.98%	1.46%
Greek	7.58%	2.87%	3.38%	.49%	4.07%	1.74%
Sanskrit	6.65%	2.85%	2.46%	.46%	1.99%	.82%

* Variphone.

Studies of this type, which have been carried on extensively, show that the entities of a phonetic system behave as a system in which dynamic equilibrium is preserved between frequency and difficulty of utterance. In this connection it should be pointed out that in a brilliant study of the errors made by deaf mutes in pronouncing different phonemes, C. V. Hudgins and F. C. Numbers ("An Investigation of the Intelligibility of the Speech of the Deaf," *Genetic Psychology Monographs,* XXV [1942], 289–392) have shown (363 f.) experimentally that the errors in articulating phonemes *increase* as the relative frequencies of the phonemes *decrease* (the correlation being *.88, P.E. .02* for consonants, and *.92, P.E. .01* for vowels)—thereby establishing a "pragmatic *scale of difficulty*" which is inversely related to frequency of occurrence.

Such few experimental and statistical studies in phonetics (cf. bibliography in *American Speech*) as are cast in dynamic terms attest to the thoroughgoing orderliness of the phonetic process.

B. *The Generalized Law of Abbreviation.* The above inverse relationship between the comparative difficulty of articulation and the relative frequency of occurrence is not restricted to phonemes. On the contrary it is a conspicuous feature of all the entities of the speech-process.

Thus, when in a given language there are differences in the amount of stress-accent, whether between words in a sentence (e.g., *a mán*), or between the roots and affixes within words (e.g., under-*stánd*-ing), there is a marked tendency for the entities of comparatively lesser stress-accent to be coupled with a greater relative frequency of occurrence, although other factors are not absent (cf. Zipf, *Psycho-Biology, op. cit.,* Chap. 4).

So, too, in respect of the comparative lengths of words (*ibid.,* Chap. 2), or of their morphological parts (*ibid.,* Chap. 4), there is an unmistakable inverse relationship between length and frequency of usage which according to the research of E. L. Thorndike ("Studies in the Psychology of Language," *Arch. of Psych.,* No. 231, Sept., 1938, 67) applies even to words

[1] For experimental support of this statement of greater difficulty, cf. C. V. Hudgins and R. H. Stetson, "Voicing of Consonants by Depression of Larynx," *Archives Néerlandaises de Phonétique Expérimentale,* XI (1935), 1–28. For the statistics on these voiced and voiceless stops, cf. Zipf and Rogers, *ibid.,* XV (1939), 111–147, also Zipf, *Psycho-Biology, op. cit.*

whose frequencies are less than two in a million. As to the mechanisms whereby this inverse relationship is preserved, Zipf has argued (*Psycho-Biology, op. cit.*) that truncations of form (e.g., *gas* for *gasoline*, or *phone* for *telephone*) and substitutions of short words for long ones (e.g., *car* for *automobile*, or *current* for *electricity*) play an important role.

Whether the above inverse relationship between length and frequency (the Law of Abbreviation) applies also to the lengths and frequencies of phrases, clauses, sentences, and so on, has never been empirically tested. Nevertheless, A. C. Norwine and O. J. Murphy ("Characteristic Time Intervals in Telephonic Conversation," *Bell Telephone System Technical Publications, Monograph B-1074*, also in *The Bell System Technical Journal*, XVII [1938], 281–291) have shown that in telephone conversations the short "utterances" or "talk-spurts" are much more abundant than the longer ones with frequency rapidly decreasing as the size of the utterance increases. Eliot D. Chapple has confirmed this finding in his studies of non-telephonic conversations, and has further found, as we shall presently see, that there is a logarithmic linear relationship between the lengths and frequencies of a person's utterances and silences during his conversations ("'Personality' Differences as Described by Invariant Properties of Individuals in Interaction," *Proc. National Academy of Sciences*, XXVI [1940], 10–16).

C. *The Effect of Analogy.* Of particular interest to students of human behavior is the phenomenon of analogic action where a particular act of one class, say *a*, which has certain characteristics of another class, say *b*, alters its behavior so as to conform completely to the criteria of class *b*. The phenomenon of analogy is both frequent and transparent in speech. Thus, though we are taught in school that the plural of *ox* is made by adding *-en* to *oxen* (class *a*), whereas the plural of most other nouns is made by adding (*e*)*s*, as in the case of *boxes, foxes, cats, dogs, houses* (class *b*), nevertheless we frequently hear the plural *oxes*, apparently out of respect for what is the plural-pattern in the overwhelming majority of nouns. Expressed as a ratio, we may say that *ox:oxes::box:boxes*. The same argument applies to the past tense of *to dive*, which today is *dived* (out of respect for the prevailing pattern) instead of the older *dove*. There are also cases of analogic changes in accentual and phonetic structure (cf. *Psycho-Biology, op. cit.*, Chaps. 3 and 4); and so too in the syntax of phrases (e.g., the more traditional *different from* becomes *different than*, presumably in analogy with *other than*).

In the analogic changes of speech we see perhaps the economy of reducing the number of different classes, or "stereotypes," in terms of which speech-action occurs.

D. *The Frequency-Distribution of Words.* In studying stenographic problems, J. B. Estoup (*Gammes Sténographiques*, 4th ed., Paris, 1916) observed the general hyperbolic relationship between the number of new and different words in successive samples of a thousand French running words on the one hand, and the cumulative diversity of vocabulary on the other. Since then innumerable frequency studies have been made.

In 1928, E. V. Condon ("Statistics of Vocabulary," *Science*, LXVII [1928], 300) presented graphically the frequency-distribution of the n different words in a large sample of speech, ranked in the order of decreasing frequency. In this he found that the r-rank of a word, when multiplied by its f-frequency of occurrence, approximated the equation[2] $r \times f = C$. In Figure I we find two examples of this distribution, with r measured logarithmically on the abscissa and f logarithmically on the ordinate for (I) the vocabulary of James Joyce's *Ulysses* (M. L. Hanley, *Word Index to James Joyce's Ulysses*, Madison, Wis., 1937), and (II) the 43,989 running words of samples of American newspapers according to R. C. Eldridge (*Six Thousand Common English Words*, Buffalo, 1911).

However, if we consider the curves of Figure I as particular cases where $p = 1$ of the generalized harmonic series:

$$F.Sn = \frac{F}{1^p} + \frac{F}{2^p} + \frac{F}{3^p} + \ldots\ldots + \frac{F}{n^p}$$

in which the denominators refer to the respective ranks of the n different words in the sample (and where $F = n^p$), we have a more useful equation in mathematically describing a "saturated" sample of approximately $F.Sn$ running words. Thus, in the polysynthetic American In-

[2] Zipf's subsequent and independent observation that the N-number of different words of like f-frequency under certain conditions approximates the equation $N \times f^2 = C$ is corollary to the above (cf. *Psych. Record*, II [1938], 347–367).

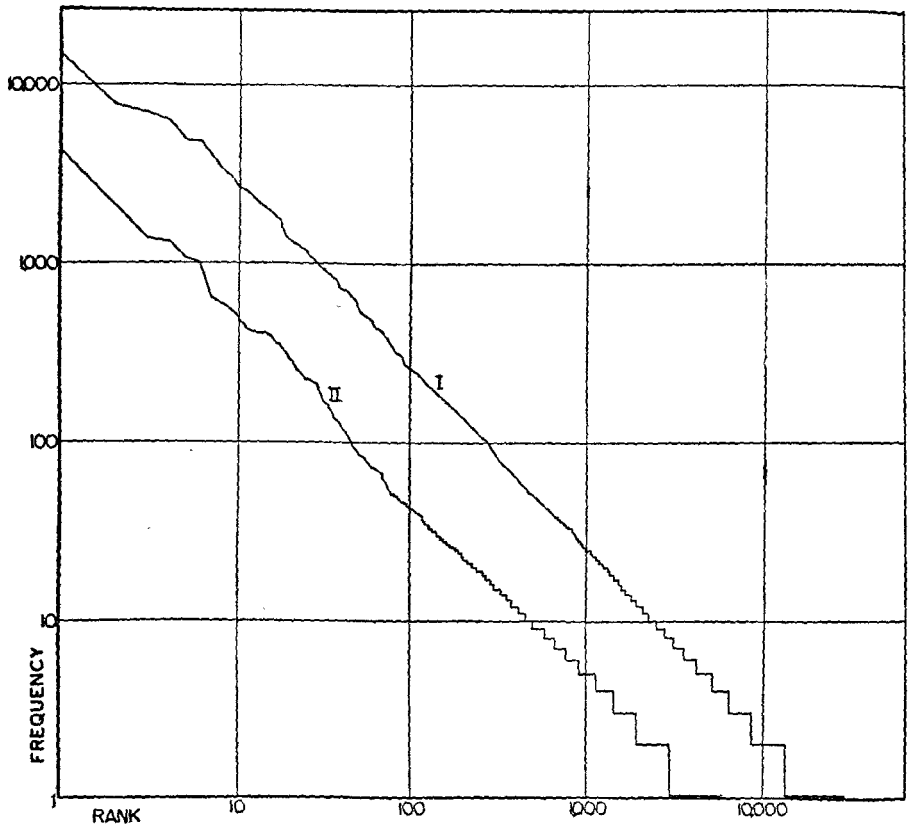

FIGURE I

dian languages of *Nootka*, etc., Zipf has observed values of p that are less than 1.[3] J. C. Whitehorn and Zipf ("Schizophrenic Language," *Arch. Neur. and Psychiatry*, XLIX [1943], 831–851) have observed a value of p greater than 1 in the letters of a female paranoid schizophrene—a finding that was subsequently confirmed independently by Dr. James J. Miller in his hitherto unpublished study of the conversation of a schizophrenic patient. Nor is this observation inconsistent with S. W. Cook's and B. F. Skinner's study of certain verbal association data of the insane ("Some Factors Influencing the Distribution of Associated Words," *Psych. Record*, III [1939], 178–184).

As to the size of n (and Sn) in the above equation Zipf has noted in his studies of children's speech (preliminary report, *Science*, XCVI [1942], 344–345; complete report to appear in *The Principle of Least Effort*) that there is a positive correlation between n (and Sn) and the chronological age of the child, as is to be expected.

But although the above equation of the generalized harmonic series may be of great descriptive value, it can scarcely be the primary equation since it tells us nothing about the length of the intervals, I, between the repetition of words, a problem whose preliminary exploration was kindly undertaken by my former seminar student, Dr. Alexander Murray Fowler, and later extended and published with detailed theoretical and mathematical treatment (G. K. Zipf, "The Repetition of Words, Time-Perspective,

[3] To be reported in his forthcoming book, *The Principle of Least Effort*.

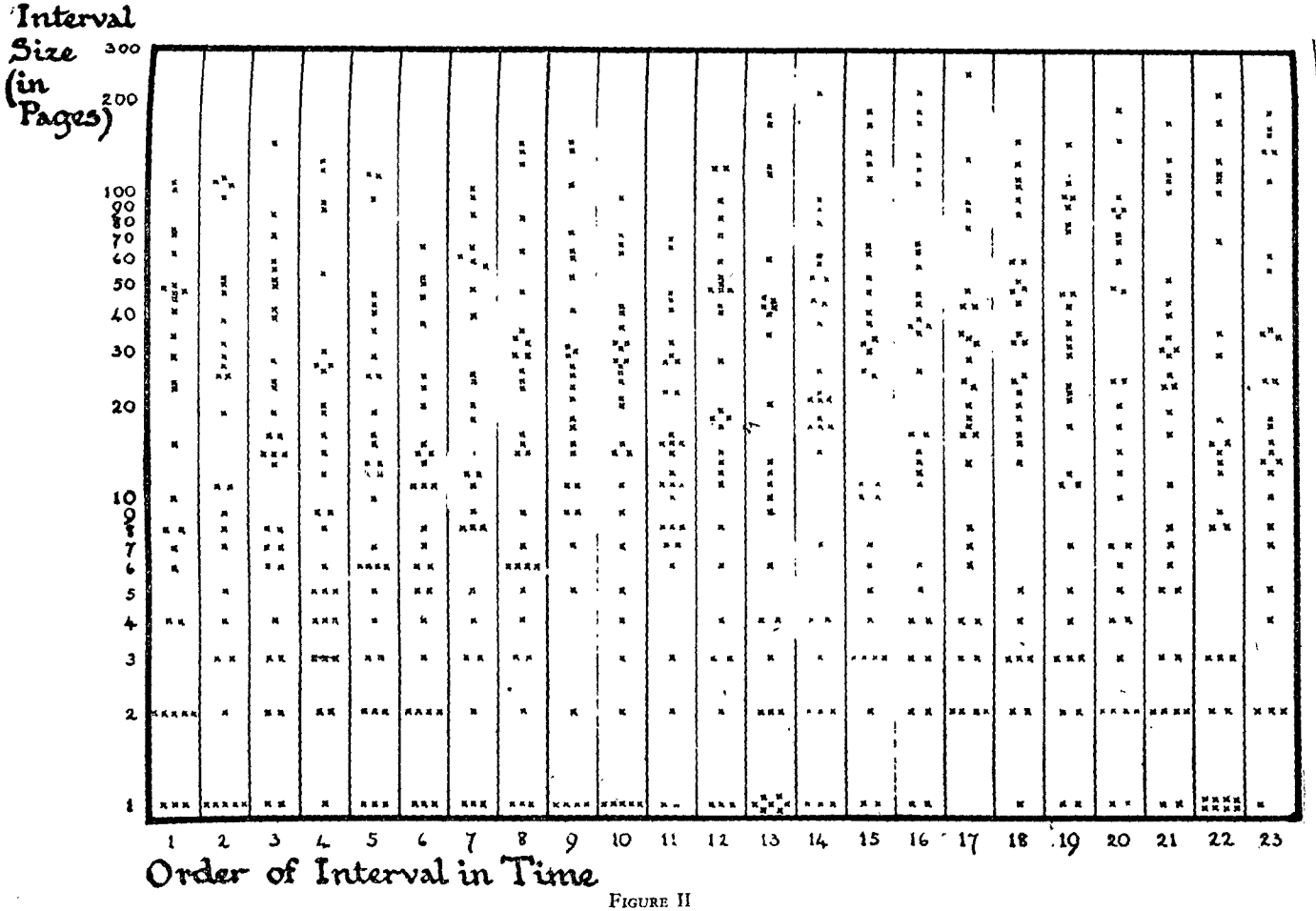
Interval Size (in Pages)
300
200
100
90
80
70
60
50
40
30
20
10
9
8
7
6
5
4
3
2
1
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
Order of Interval in Time

FIGURE II

and Semantic Balance," *Jour. General Psych.*, XXXII [1945], 127–148, with 28-item bibliography). According to the data of the Hanley word-index to Joyce's *Ulysses* (*op. cit.*), the N number of different I_f intervals of like length (in pages) between the $f - 1$ repetitions of all words in the *Ulysses* of like f-frequency approximates the simple equation:

$$N \cdot I_f = a \text{ constant.}$$

As a typical example of this distribution we find in Figure II the actual data for the I_{24}-intervals between the repetitions of all the 34 different words occurring 24 times in the *Ulysses;* arithmetically on the abscissa are noted the 23 (i.e., $f - 1$) successive intervals, and logarithmically on the ordinate are indicated with crosses the I interval-sizes in pages for the 782 (i.e., 34 x 23) different intervals between repetitions. An inspection of this scatter-diagram reveals not only that the I-intervals of varying lengths are distributed quite evenly over the entire book, but also, and even more significantly, that they are similarly distributed in each of the successive intervals. Zipf has argued (*ibid.*) that this indicates the even distribution of minimalized work over time, and that the foregoing equation is primary to the others which are corollary to it.

E. *The Problem of Distributing the "Meanings" of Words.* Inasmuch as words are used to convey "meanings" (i.e., to evoke more or less stereotyped responses in the interlocutor) the question presents itself as to what these clear-cut linear word-frequency distributions may indicate about the distribution of different "meanings" among the words. Although the theoretical problem is too extensive to discuss here,[4] Zipf deduced theoretically that the m_f number of different "meanings" of a word of f-frequency would tend to approximate the square root of f, or:

$$m_f = f^{.5}$$

For the purpose of testing this theoretical equation Zipf, with the help of his students, ascertained from the data of the *Thorndike-Century Senior Dictionary* the average m-number of different living meanings per word (according to the Lorge-Thorndike semantic count, cf. *ibid.* preface) for the twenty successive sets of thousand most frequent words of the E. L. Thorndike list of 20,000 most frequent English words (E. L. Thorndike, *A Teachers Word*

[4] Cf. G. K. Zipf, "The Meaning-Frequency Relationship of Words," *Jour. Gen. Psych.*, XXXIII (1945), 251–256. To be treated more fully in *The Principle of Least Effort.*

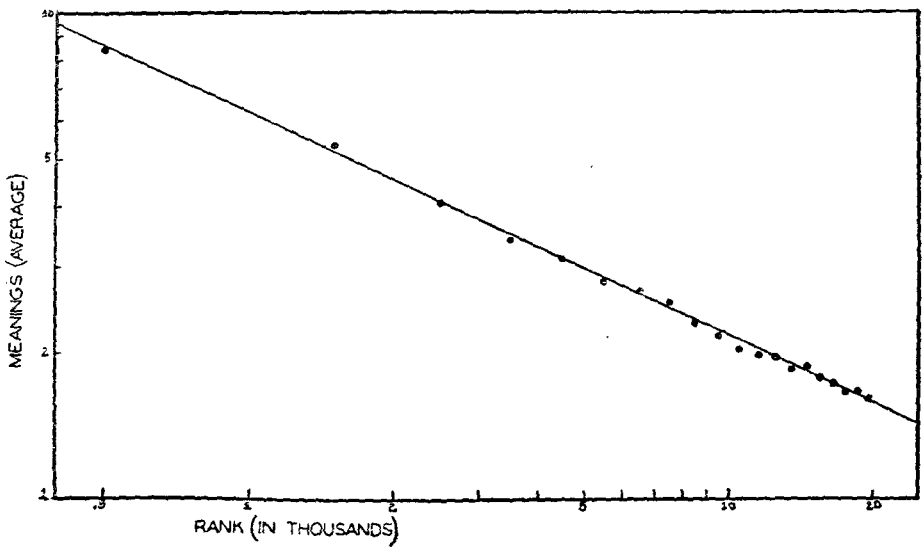

FIGURE III

Book of the Twenty Thousand Words, etc., revised, New York, 1932).

The results of this investigation are presented graphically in Figure III, with the ranked thousands of words measured logarithmically on the abscissa, and with the average meanings per word measured logarithmically on the ordinate. By least squares these points of Figure III have a negative slope of .4605 (± .0083). And *if we may assume* that the rank-frequency distribution of the underlying words would approximate the equation of the generalized harmonic series with $p = 1$, then we may say that these data can be approximately described as

$$m_f = f^{.4605}$$

In any event the linearity of the data of Figure III, which far surpasses even the fondest hopes of the present writer, shows that there is indeed a law of word-meaning distribution.

F. *Summary of General Speech-Dynamics.* Although the foregoing data represent only a small sample of the present-day resources of statistical findings, nevertheless they suffice to show that the speech-process is dynamically structured according to fundamental principles that apparently operate quite generally with extreme rigor. And that seems to be the case whether we view the sizes and frequencies of usage of words, or their rate of repetition, or their respective "loadings" with "meanings," or the phonetic minutiae of which they consist. Since it is difficult to imagine that the speech-process is not completely integrated with the rest of the personality and with human social relations, these findings would seem to impose important restrictions upon all theories of the personality and of human relations.

IV. Style and Personality Invariants

Each person's speech is to some extent characteristic of him, whether because of its tone, timbre, and the like, or because of his choice and arrangement of words and the speed with which he talks, or because of his general loquacity or taciturnity—or because of a combination of all these. And to some extent a person's written style may be characteristic of him. Indeed, since earliest times, holy men have scrupulously guarded the *precise* structure of sacred texts, while the reconstruction of the *original* texts of past authors has provided a livelihood for innumerable philologians who have also not hesitated to ascribe texts to one author or another on the basis of generally subjective criteria of style (though sometimes they may naïvely strive to be objective). Perhaps the most articulated of these subjective (or quasi-objective) tests of style is that of E. Sievers (*Ziele und Wege der Schallanalyse*, Heidelberg, 1924), which, though no longer in vogue today, is nevertheless typical of the "Pentecostal" (or "Whitsuntide") methods of analysis that are employed in such problems by European and American experts in the humanities today.

In the meantime, however, with the general advance of scientific interest, clear thinkers, unencumbered by the "Pentecostal" traditions of linguistics and the humanities, have turned with considerable success to the problem of the characteristics of individual speech style.

Thus, in 1938, G. Udny Yule published ("On Sentence-Length as a Statistical Characteristic of Style in Prose; with Application to Two Cases of Disputed Authorship," *Biometrika*, XXX [1939], 363–390) his quantitative investigation of sentence-length as a statistical characteristic of style in prose. Dividing sentence-lengths into classes of 1–5 words, 6–10, 11–15, and so on, Yule analyzed two sizable samples each from Bacon's *Essays*, Coleridge's *Biographica Literaria*, Macaulay's *Essays*, and Lamb's *Elia* and *Last Essays of Elia*. In each case he observed that the two samples from the same author agreed fairly closely, while differing from those of other authors. He then applied his technique to writings of disputed authorship.

In 1944, Yule reported (*Statistical Study of Literary Vocabulary*, Cambridge, England, 1944) the results of his statistical analysis of the diversity and frequency of nouns in several authors. Despite the admitted incompleteness of Yule's work, which may have resulted to some extent from Yule's strange unawareness (p. 32) that anyone else had compiled and studied word distributions,[5] the analysis itself is of great didactic value both for its statement of statistical problems related to the topic and for its application of statistical methods to some of those problems.

A study of alliteration and of certain types

[5] For a fairly complete bibliography of the now enormous literature on word-frequency distributions, cf. C. C. Fries and A. A. Travers, *English Word Lists*, American Council of Education, Washington, D. C., 1940.

sound-patterning in poetry by B. F. Skinner ("A Quantitative Estimate of Certain Types of Sound-Patterning in Poetry," *Amer. Jour. of Psych.*, LIV [1941], 64–79) has also yielded interesting results—though partly of the nature of a negative proof—and is of considerable value in helping to state objectively the entire problem of literary and stylistic criticism. Moreover, B. F. Skinner's invention of the verbal summator ("The Verbal Summator and a Method for the Study of Latent Speech," *Jour. of Psych.*, II [1936], 71–107), together with W. K. Estes' invention of the visual form of the summator ("A Visual Form of the Verbal Summator," *Psych. Record*, IV [1940], 174–180), offer distinct possibilities of studying quantitatively a person's own verbal associations, which, more than anything else, would seem to be peculiar to his personality.

As far as the diversity of a person's vocabulary is concerned—a topic inferentially broached by Estoup (*op. cit.*) and latterly treated in respect of nouns by G. U. Yule in his recent book (*op. cit.*)—J. B. Carroll, who has studied the problem both quantitatively and theoretically with great care, believes: "An index of diversity might also be used to differentiate linguistic materials with respect to stylistic and other characteristics." (J. B. Carroll, "Diversity of Vocabulary and the Harmonic Series Law of Word-Frequency Distribution," *Psych. Record*, II [1938], 379–386.) Carroll's equation for vocabulary diversity would seem to be of considerable practical value (*if*, as he points out, the equation of the harmonic series is known to apply to the entire sample).

Addressing himself to the topic of speech and the personality, F. H. Sanford restricted his investigation to a comparative study of the speech-action of two young men, and is able to report on the basis of extensive and scrupulously detailed data (F. H. Sanford, "Speech and the Personality: A Comparative Case Study," *Character and Personality*, X [1942], 169–198): "By means of 234 'mechanical,' grammatical, 'psychogrammatical,' and lexical categories, samples of oral speech from two subjects were subjected to an intensive statistical analysis. The quantitative data yielded by this analysis lend themselves to conceptualization in terms of linguistic traits, one group of traits for one subject, another group for the other."

At about the same time the above studies were undertaken, Eliot D. Chapple, likewise working independently, studied the ratio of an individual's speech-actions to his speech-silences in conversations. Chapple discovered (*op. cit.*, p. 15): "The measurement of the interaction of individuals provides us with an opportunity to find out whether any unique property of an individual, ordinarily called 'personality,' manifests itself when two people are talking together. When a series of observations is made, the frequency distributions of the durations of actions and silences are fitted to the exponential equation, $F = ae^{-bt} + ce^{-dt}$, and the plot of $log\ b_a/b_s = -log\ d_a/d_s$ with a slope of 1." (I.e., of -1, *ed. note.*) "The position of this curve as defined by the intercepts is invariant for each individual, since it does not shift when the individual interacts with different individuals. The range of the curve also may be invariant for each individual, being delimited at the lower end by the absence of a d silence slope and at the upper end by the absence of a d action curve. These invariant properties afford us a quantitative description of individual differences in 'personality' as exhibited in the rates of acting and being silent in interaction."

Although in the present writer's opinion the above Chapple observation is of enormous theoretical value in understanding the economy of mentation, it should be pointed out that Dr. Chapple has also found his methods of great practical value in the entire field of the diagnosis of personality disorders (cf. E. D. Chapple and Erich Lindemann, "Clinical Implications of Measurements of Interaction Rates in Psychiatric Interviews," *Applied Anthropology*, I [1942], No. 2, 1–11).

In addition to the foregoing studies there are many others for specialized fields, such as that of Mary Shattuck Fisher for children's speech ("Language Patterns of Preschool Children," *Child Development Monograph*, No. 15, Teachers College, New York, 1934). The problems of reading and spelling are quite properly fields in the general psychology of language. These will be found abstracted in the *Psychological Abstracts*. For the entire field up to 1936, including also the largely subjective treatises, there is an excellent critical bibliography by Donald V. McGranahan ("The Psychology of Language," *Psych. Bulletin*, XXXIII [1936], 178–216). For

additional bibliography and discussion see J. Eisenson, *The Psychology of Speech*, New York, 1938.

V. Geographical and Socio-economic Aspects of Speech

In addition to the internal dynamics of speech and their peculiar manifestations in the speech of different persons, there are also geographical and socio-economic factors in speech that may be of considerable importance to social psychology.

A. *Dialect Geography.* A native of New Orleans in talking to a native of Chicago, Boston, or New York will unwittingly inform the latter of his place of origin, because of his peculiar "accent." This relationship between a person's accent and his regional origin has long been noted and during the past century has evoked the ever-growing interest of specialists in the field of dialect geography (for an excellent account see L. Bloomfield, *Language*, New York, 1933, Chap. 19; for a fairly complete annual bibliography of articles consult *Indogermanisches Jahrbuch*, Strassburg-Berlin, annually since 1914). The chief weakness of dialect geography to date, as far as dynamic social structure is concerned, is its tendency—unlike present-day geography—to ignore the research in the growing field of demography (for a bibliography of demography see the quarterly *Population Index* of the Population Association of America).

B. *Speech as Socio-economic Cues.* It has perhaps always been known that a person's speech tends to reveal a person's social-economic class-origin (cf. G. B. Shaw's *Pygmalion*). In this sense a person's speech may be viewed as consisting in part of what can appropriately be termed *socio-economic cues.* This problem has been approached with the broadest possible perspective in the penetrating study of trait-names by G. W. Allport and H. S. Odbert ("Trait-names," *Psych. Monographs*, No. 211, 1936), whose discussion (pp. 1–37) of the entire problem of the verbal labelling of the "other fellow's" traits of disposition may be viewed as a classic introduction to the field. But as to the narrower topic of correlating particular traits as cues to membership in particular socio-economic groups, investigation still lags.

George Kingsley Zipf,
Harvard University

LEARNING.—The problem of learning is one of the most important in the whole field of psychology, because all human behavior, on whatever level and under whatever circumstances, involves learning of one kind or another. It is clear from this statement that we are making the meaning of the term learning more or less equivalent to the meaning of such words as change, modification, growth, development and adjustment. To be sure, these words have, in addition to their general meaning, specific implications, but any organism, at any stage of its evolution, is not only in an environment but is being acted upon by that environment and reacting in turn to it. And all such action-reaction behavior involves changes and modifications of the organism as well as, in some instances, changes in the environment.

These statements hold true regardless of what position we take as regards endowment or heredity. Whether we conceive of the organism from the moment of birth or from the moment of conception, the general statements of the first paragraph hold true. We shall not, in this article, be concerned with the problem of heredity, but rather, as is clear from the title, with the general problem of learning. It is also evident that change, modification, or learning take place in the action and reaction of the organism and the environment, whether such change is intentional, deliberate, and controlled, or haphazard, uncontrolled, and without intent. Some writers are wont to use the term learning only in deliberately controlled situations, but the modifications and changes that take place without intent are often even more productive of major shifts in the organism, in its relation to the environment, than the specifically deliberate modifications.

The same general principles are involved when a student sits down to learn how to extract cube root or when that same student inadvertently and unwittingly acquires at some period in his life a fear of cats or of the dark. Furthermore, the same principles of learning are involved when a cat learns to get out of a box, a rat to run a maze, a child to tie or untie his shoe-laces, or an adult to give a public speech. In short, learning may involve modifications of a motor nature, a verbal nature, an emotional nature, or as is most frequently the