Zipf (1949) Human Behavior + the Principle of Least Effort 98-108

With the above considerations in mind, we shall now inspect the dynamics of a phonetic system under the four headings: (*a*) the economy of permuting phonetic entities, (*b*) four fundamental principles of a phonetic system, and (*c*) miscellaneous considerations.

a. The economy of permuting phonetic entities. Although the permutation of phonetic entities into words is highly economical of human speech efforts, we must remember that such permutations are by no means necessary. Thus if we wanted a vocabulary of 10,000 different words, we could physiologically produce those 10,000 different words by means of 10,000 different unpermuted "word sounds." For example our American vowels

could be so varied in length that vast quantities of different vowels could be produced to serve as word sounds.* If, in the midst of stretching a phoneme's length to make a word sound, it became necessary to take a breath, the speaker could introduce a stereotyped grace note to signal: "After I take my breath I shall continue with the same invariant word sound and not begin a new one." In addition to differences in length we could use differences in pitch and amplitude, from whispering to shouting (within the limits of the principle referred to under the "Weber-Fechner Law"). In short, a vocabulary of thousands upon thousands of distinguishably different word sounds is physiologically and acoustically possible.

Nevertheless, once we envisage a stream of speech that consists of unpermuted word sounds of the above kind, we comprehend its impracticability. Thus the "meanings" of Lincoln's *Gettysburg Address* which were delivered in a very short time by using words that were permutations of phonemes might well have lasted for hours or even days if word sounds had been used. Since in this case the extent in time is directly correlatable to the amount of physical work both for the speaker and for the auditor, it is evident that *the permutation of speech sounds* (*or phonetic entities*) *into words saves work*. Hence we find that our Principle of Economical Permutation, which applies to holophrases, words, and morphemes, may be expected to apply to the fundamental phonetic entities of speech (no matter how defined).

This consideration of the economy of permutations leads us to various interesting questions about the most economical number, frequencies, and kinds of phonetic entities to be permuted. Four of these questions we shall now inspect both theoretically and, where possible, empirically.

b. Four fundamental principles of a phonetic system. There are four characteristics of phonetic systems that point to the existence of four fundamental principles of economy. The *first* is that the actual number of different phonemes and variphones in a given phonetic system is generally between 20 and 60 and not in the hundreds, or thousands, or millions. A *second* is that out of the vast range of possible phonetic types, certain particular vowels and consonants (e.g., *a, n, m, s,* etc.) seem to be found quite consistently in widely diverse languages; hence in respect of some of the particular phonetic-types employed, languages seem to agree to an extent that is by no means justified on the grounds of "chance." *Third,* when different languages happen to use approximately the same phonetic entities, they tend to use them with approximately the same frequency. And *fourth,* whenever in a given language a particular phonetic entity changes its form under particular conditions in a given word, it tends to undergo the same change in every other word in which the same conditions prevail (this fourth well-known fact was first brilliantly elaborated by Karl Brugmann in the 19th Century).

---

* Sheer difference in duration can distinguish different phonemes, as with German short ă and long ā. Witness German *kan* (spelled *kann*) "is able," and *kan* (spelled *Kahn*) "boat"; or *man* (Mann) "man," and *man* (mahn!) "warm."

Now let us attempt briefly to explore the dynamics governing the above four characteristics, starting with the question why phoneme systems have about the same number of different phonetic entities.

First, we may say that adult human beings possess by and large the same vocal apparatus with the same range of variation, regardless of the particular ethnic or racial groups to which they belong. For example, negroes and whites have approximately the same speech physiology, with variations from the norm in the one being more or less duplicated by variations in the other. This fact of an approximate agreement in physiological givens leads us to an interesting consideration. For if the members of the different ethnic or racial groups have approximately the same vocabulary needs —say, from 10,000 to 20,000 different words—and if they all use the same Principle of Economical Permutation upon approximately the same physiological givens, the result will be an approximate agreement in the number of different phonetic entities in the various phonetic systems. (The problem is related to the one previously discussed in reference to the Z-value of the different usages that can be performed by n different tools.) Therefore it is by no means surprising to find that phonetic systems do not vary widely in the number of their different entities, so far as we know. And even if the vocabulary needs of different speech-groups should happen to vary quite widely, the variation in the size of the phonetic system would not need to be commensurately large because of the logarithmic nature of the Principle of Permutation.

The *second* principle of a phonetic system refers to the specific phonetic types of actual phonetic systems. The easiest approach to an understanding of this principle is to ask the following question: which particular phonetic types will the various languages select for adoption from the vast number of possible types? And the answer to this question, according to the Principle of Least Effort, is that each language will tend to pick those phonetic types which are easiest both to articulate orally and to discriminate aurally. Insofar as the physiological givens of the vocal apparatus are the same, the easiest phonetic types will be the same, and the various languages will tend to agree as to the phonetic types selected. If there are physiological differences in the speech apparatus of various groups, then there may be corresponding differences in the phonetic types of those groups. Moreover, if there happens to be a number of different phonetic types of approximately equal difficulty, then one language may select certain ones, and other languages may select others. Yet within these qualifying restrictions there will tend to be a substantial agreement in the phonetic types selected. As to an empiric confirmation of this agreement, we know of no actual quantitative investigation of the topic, although it lends itself to quantification. Thus one could select at random 100 different languages whose phonetic systems are reliably established. Then one could take the International Phonetic Alphabet and for each symbol tabulate the number of different languages whose phonetic systems contained the symbolized phonetic entity. And finally one would determine whether the distribution of entities throughout the 100 different languages was that of chance. This investigation is

open. Yet even without the results of an actual investigation, we know that certain common vowels, diphthongs, nasals, and fricatives in Western European languages are quite general. Without the presence of this second principle it would be difficult to comprehend the data of the third principle, which now emerges from the one we have just discussed.

The nature of the *third* principle becomes clear if we ask the question: what are the most economical relative frequencies with which the different phonetic entities should be used? In reply, according to the Principle of Least Effort, we can only submit (1) that the easier phonetic entities will be the more frequently used, and (2) that if our preceding two principles are valid we may expect to find that similar phonetic entities in different languages will have similar percentage-frequencies. Both of these points (1) and (2) are empirically demonstrable.

To confirm this third principle empirically we shall proceed as follows. Our first step will be to select a set of corresponding pairs of phonemes in a given language where there is no doubt in the minds of competent phoneticists as to which is the easier member to pronounce; then we shall note if there is any positive correlation between greater ease of articulation and greater frequency of use. The phonemes selected for this purpose are the six pairs of voiceless aspirated stopped consonants, together with their nonaspirated counterparts in the Peiping dialect of Chinese as presented in Table 3–2.[24] The aspirated stopped consonant, which is indicated by a superscript *h,* is the more difficult of the pair because *first,* it has a tense or *fortis,* pronunciation whereas its mate has a nontense or *lenis* one, and *second,* its explosion is followed by a marked puff of air (*h*) that is lacking to its lenis unaspirated counterpart. Inspecting the percentages of Table 3–2 which refer to the phonemes in samples of 20,000 running Chinese syllables, we note that in all cases the easier unaspirated stop is almost twice as frequent as the more difficult aspirated stop. This confirms our hypothesis.

TABLE 3–2

Voiceless aspirated fortes and voiceless unaspirated lenes stops in present-day Peipingese.

(Percentages in reference to occurrences of all speech sounds in 20,000 running syllables)

	t^h/t	p^h/p	k^h/k	cc_c^h/cc_c	$tʃ^h/tʃ$	ts^h/ts
Aspirated Fortes Stops	2.56%	.56%	1.02%	1.04%	1.23%	1.40%
Unaspirated Lenes Stops.....	6.18%	2.37%	2.58%	2.69%	2.44%	2.63%

Data from other languages on aspirated and nonaspirated stops are in accordance with the above and have been published elsewhere with a detailed description of the phonetics involved.[25]

Turning now to the question of a possible correspondence in percentage-frequencies between similar phoneme types in different languages, we present data in Table 3–3 for the voiced and voiceless pairs of stops, *t/d*, *p/b*, and *k/g* in 17 different languages as previously published elsewhere. An inspection of the table discloses a rough correspondence between the magnitudes of each column. Thus the percentages for *t* are about 6%, those for *d* about 3%, and so on. In view of the differences in the sizes of the samples, the kinds of materials examined, and the methods of analysis employed by the various transcribers of texts, the presentation of mean-values and standard deviations may seem to be statistically dubious; nevertheless it is easier to include them than to explain their absence, nor is any harm done if they are not taken too seriously.[26]

Even more striking in Table 3–3 is the fact that the frequencies of the voiceless stops, *t*, *p*, and *k*, are with negligible exception greater than those of their corresponding voiced stops, *d*, *b*, and *g* in the 17 samples.* This is not to be anticipated from the law of probabilities. Indeed, the probability

TABLE 3–3

Percentage of occurrences of voiced-voiceless stops.

(Diphthongs counted as one unit.)

No.		t	d	p	b	k	g
1	Czechish	5.60%	3.73%	3.52%	1.86%	3.93%	.15%
2	Dutch	7.83	4.67	1.99	1.20	3.21*	.09*
3	English	7.13	4.31	2.04	1.81	2.71	.74
4	Hungarian	7.18	3.30	1.04	1.71	5.72	2.45
5	Lithuanian	5.76	2.61	3.71	1.35	4.61	1.36
6	North Russian	7.97	1.52	3.36	1.01	3.36	.67
7	South Russian	7.05	2.46	2.79	1.51	3.97	1.66
8	Wendish	6.26	3.02	2.55	1.56	3.29	2.41
9	East Ukrainian	3.83	3.24	2.82	2.11	4.11	
10	Bulgarian	7.54	3.55	2.82	1.32	2.98	1.46
11	Greek	7.58	2.87	3.38	.49	4.07	1.74
12	Sanskrit	6.65	2.85	2.46	.46	1.99	.82
13	Latin	8.66	3.12	2.54	1.32	4.34	.76
14	Italian	4.72	3.64	2.14	.52	3.38	.48
15	Spanish	4.46	1.56*	2.92	.46*	3.84	1.02
16	Portuguese	5.06	2.44*	2.68	.30*	3.44	.92
17	French	4.90	4.54	3.96	1.82	3.30	.36
	Average	6.36	3.14	2.75	1.22	3.66	1.07
	Standard Deviation	± 1.37	± .28	± .22	± .18	± .25	± .70

* Stops marked with an asterisk (*) are variphones (see text) and all others are phonemes.

that the voiceless stops will be more frequent than the voiced stops, as observed in the voiced-voiceless pairs in the above 17 languages, on the assumption of the null hypothesis that either kind of stop is equally likely

* A count of 5,000 Japanese phonemes in *Roomaziaki Tanpen Syoosetusyuu* (pp. 1–10) as romanized by Dr. N. Tanakadate revealed the following percentages: *p*, .26%; *b*, 1.52; *t*, 9.24; *d*, 2.86; *k*, 6.26; *g*, 2.20; *m*, 3.84; *n*, 5.92. Except for the *p*, these percentages are not far off. They are not included in the tables because I was not sure of the phonetic structure of some of the phonemes.

to occur, is about 5 chances out of a million [or, more precisely $P = (.5085)(10^{-5})$, according to Drs. Henry S. Dyer and John K. Dickinson, who very kindly calculated this probability]. Hence we are justified in asking whether, on the whole, the voiced stops are more difficult to produce than their corresponding voiceless ones. That such is indeed the case has been established by the brilliant experimental research of C. V. Hudgins and R. H. Stetson on the depression of the larynx in the voicing of consonants.[27]

The reader may wish to ask whether the correspondences of Table 3–3 are due to the *genetic relationship* of the different languages (except Hungarian). In reply, no. The Dutch and English stops, for example, are not historically the same as the others, if only because of the operation of Grimm's Law. The intervocalic *t* of Latin *vita* which is preserved in Italian *vita* is lost in French *vie* and has become a spirant (like *th* in English *thy*) in Spanish and Portuguese. Indeed in the course of the thousands of years that have elapsed since some of these languages could have constituted a common ethnic group, a fairly large shifting of phonetic forms has demonstrably taken place.

The reader may also wish to know whether data are available for other phonetic types. It should be pointed out that except for stops (i.e., explosive consonants) there is likely to be a considerable variation in the length and stress of utterance of speech sounds, as is notoriously the case for vowels, and somewhat the case for liquids, nasals, affricates, and aspirants. It so happens, however, that when short vowels on the one hand are compared with long vowels and diphthongs on the other, the short vowels are almost without exception markedly more frequent. The liquids (e.g., *r* and *l*) show a wide variation in frequencies among languages; and so too do the others, except for the nasals, *m* and *n*. As we note from the data for 22 different languages in Table 3–4, the *m* seems with negligible exception to be much less frequent than the *n* in the same language, and also to have approximately the same percentage-frequency in other languages widely different in region and time and, in some cases, not even remotely related.[28]

Although we have presented mean values and standard deviations for the data of Table 3–4, we again remind the reader that the underlying samples differed widely in size and technique of recording (e.g., there was no consistent treatment of what were vowels, diphthongs, triphthongs and this inconsistency would somewhat affect the size of the sample and hence the percentages); moreover, the percentages might well vary with styles of prose selected. If we had a dozen samples, each 10,000 phonemes long, for a given language, and each of these from different styles of speech phonemically transcribed with phonetic variations indicated, we might well obtain values of significance for the tongue in question which we could profitably compare with the results of similar undertakings with other tongues. Yet even then a certain amount of caution is in order. For after all, some phonemes in some languages may have excessive frequencies and be on the point of undergoing a corrective formal change (i.e., *phonetic change*); therefore a given above-average frequency may merely indicate the instability of the phoneme in question. This consideration leads us to the *fourth* principle of a phonetic system—*phonetic change*—which is intimately related to the other three.

TABLE 3-4

The frequencies (in percentages of the whole) of *m* and *n* in twenty-two languages.

Language	m (%)	n (%)	Language	m (%)	n (%)
Czechish	3.52	6.42	Burmese	4.72	4.15
Dutch	3.18	7.09	Swedish	3.28	7.32
English	2.78	7.24	Danish	3.18	5.70
Hungarian	3.35	5.74	Singhalese	3.12	7.40
Bulgarian	2.22	7.00	Old English	2.81	8.40
Russian	3.12	5.13	Old High German	2.91	10.85
Icelandic	4.37	7.77	Latin*	3.42	5.42
Greek (Attic)	3.19	8.55	Italian*	2.62	7.10
Sanskrit	4.34	7.04	Spanish*	2.98	5.62
Peipingese	2.18	10.18	Portuguese*	3.38	4.92
Cantonese	4.07	5.70	French*	3.42	3.04
			MEAN	3.28 ± .63	6.72 ± 1.78

* From F. M. Rogers' Analysis.

We approach the *fourth* principle by asking the obvious question: what happens to the percentages of frequency of phonemes in the dynamic process of evolution as old words are either abbreviated or eliminated while new ones are introduced? Obviously, unless all speech alterations are made within the severe restrictions of preserving the pre-existent phonetic percentages, these percentages will fluctuate—and even fluctuate quite widely. Differently expressed, unless some regulatory mechanism is present in the phonetic system to correct excessively high or low percentages, we should expect to find that the frequency of a given phonetic type will vary quite widely not only between different languages at a given time but also between different periods of the same language.

As a matter of fact a regulatory mechanism does exist in the form of *phonetic change*. Thus, for example, if a given long vowel becomes too frequent it may be shortened; a too frequent *d* may be weakened to a *t*. In general, whenever a phonetic entity undergoes an increase or decrease in frequency beyond the thresholds of toleration for its particular form, it may be expected to undergo a compensatory change in form.[29] Whenever such a phonetic entity undergoes a particular phonetic change—such as a *d* to a *t*, or the reverse—that change will occur in every word that contains the affected phonetic entity.* We shall call this the *orderliness of phonetic change*.

Examples of the *orderliness of phonetic change* are literally legion, as the reader can determine for himself by consulting historical treatises on any language. This *orderliness* which Karl Brugmann and his school first established with rigor, and which, according to Dr. Clyde Kluckhohn, was the first disclosure of a rigorous law of action in the entire biosocial field, has served as the major premise for the exhaustive work of the historical-descriptive field known variously as *comparative philology* and *linguistics*. From this enormous stock pile we shall present only a few arbitrary examples.

Thus in Old English the phoneme, *ū*, changed to *ou* (sometimes written *ow*). Because of this change, *mūs* became *mouse*, *hūs* became *house*, *lūs* became *louse*, *cū* became *cow*. Although this change obliterated *ū* from Old English, the obliteration was only temporary. For subsequently the phoneme, *ō*, changed to *ū*, as *gōs* became *gūs* (written *goose*), and *mōna* became *mūn* (written *moon*), etc. These examples are particularly interesting because they illustrate how a given phonetic type, like *ū*, because of its instability, may change to *ou*, only to make place for a new *ū* which results from a change in the erstwhile phonetic type, *ō*. And after this second change from *ō* to *ū* had occurred in Old English, the erstwhile phonetic type, *ā*, took over the abandoned phoneme type, *ō* (e.g., Old English *stān*, *rāp*, and *gāt* appear today with *ō* pronunciation in *stone*, *rope*, and *goat*). So much then for the fourth principle of a phonetic system: the *orderliness of phonetic change*.

And so we may say in summary that a phonetic system alters the form of its component elements within the four principles of the phoneme system, as just explained. Thus (1) the phonetic system limits the number of its different phonetic types to approximately 20–60; (2) the phonetic systems of the earth seem to favor the use of the easiest phonetic types; (3) the frequencies of phonetic types are inversely related to their comparative work coefficient, with the result that like phonetic types tend to have like percentage frequencies; and (4) alterations in phonetic form occur in an orderly fashion throughout the entire vocabulary of a language in the sense that when a given change occurs under a given set of conditions in one word the same change will occur in all other words where the same conditions are found. Although phonetic changes are constantly occurring, we must not overlook the great antiquity of some present-day phonetic manifestations; thus the *m* and *s* of *mouse* are probably thousands of years old in that very word.*

*c. **Miscellaneous considerations.*** Now that we have outlined the four principles of the phonetic system, let us return to the question of the *phonetic entity*, *phoneme*, and *variphone*. We shall try to demonstrate first of all why a phoneme is not necessarily the exclusive minimal unit of distinctive significance (as some phonologists have argued); and second why

* As has been discussed in detail with copious illustrations in a previous publication, if the affected phonetic entity changes only in a particular set of conditions (e.g., in accented syllables, or finally, or intervocalically), it will change in all words where the conditions are found.

* Further detailed discussions of the effect of accent, analogy, assimilation, dissimilation, haplology, etc., have been presented in previous publications.[30]

a variphone can be a minimal unit of distinctive significance (as most phonologists deny).

Let us begin our demonstration by selecting a language like English, in which *d* is a phoneme (e.g., *bad vs. bat*). Then let us assume that because of a phonetic change all final *d*'s of words change to a form, *ø*, which is new to the language (e.g., *bad* becomes *baø*). Since the new form will occur only at the end of a word, and since no *d* will occur at the end of a word, it follows that no phonological opposition will exist in the entire language to show the difference between *d* and *ø*. Therefore by definition the new form, *ø*, will not be a phoneme but a *variphone* of *d*. Yet how do we know that *ø* is a *variphone* of *d* instead of a variphone of some other phoneme, like *l, m, n, o,* or *p*? Obviously we know that *ø* is a variphone of *d* only because we know that historically it developed from *d;* if this historical information were lacking we should not know it. Furthermore, before *d* changed to *ø* in final positions, those *d*'s in final positions were still phonemes which by definition were "the minimal units of distinctive significance." Yet after the change, what is *ø*? For us it is still a unit of distinctive significance, for we note a difference between a hypothetical *colø* (*cold*) and *cōl* (written *coal*).

Although we have defined *phonetic entities* as minimal units of distinctive significance, we have yet to be shown that an actual phonological opposition is necessary before phonetic entities can become minimal units of distinctiveness. As far as we are concerned, it is the permutation of the phonetic entities that is important. The existence of *phonological oppositions* is fortuitous. It is striking that practically all languages have *phonological oppositions* to illustrate most of the stock of different phonetic entities. But we submit that that is a result of the Principles of Economical Permutation. It is generally economical to make the small permutations of phonetic entities before making the larger ones; therefore *in the long run* small permutations are likely to arise to serve as *phonological oppositions* for most of the distinctly different phonetic entities of the given language's phonetic system.*

If we inspect further the "theory of the phoneme" we note a frequent belief which we express in Dr. Leonard Bloomfield's words (p. 83):[32] "the phoneme is kept distinct from all other phonemes in its language." This statement is in error. We have already pointed out, for example, that in

* As a final disproof of the phoneme as *the minimal unit of distinctive significance* (a definition that has become an article of faith in some quarters of American linguistics), let us take the case of Gothic *hausei!* "hear" and Gothic *hēr* "here" with a clear difference between *au* and *ē*. In Old English the respective words were *hīer!* and *her*—a true phonological opposition. But today they are the homophones, *hear!* and *here*. Unless one makes the ridiculous assumption that all phonetic change is instantaneous, we must conclude that for a long or short period between Old English and today the vowels of *hear* and *here* were ambiguous in respect of the phonological technique (i.e., phonologists would not know whether they were still different phonemes, or already the same phoneme). By extension of the argument, there may always be a phonologically unclassifiable residue in the stream of speech that is nevertheless not without distinctive significance either in light of the moment, or in light of what has been or is about to be.[31]

German short *ă* and long *ā* are different phonemes. Drs. Eberhart and Kurt Zwirner in a brilliant series of phonometric studies have found in phonographic recordings of actual samples of speech that the frequency distributions of the length of utterances of different vowel phonemes follow a "normal curve"; the only difference between *ă* and *ā* is in the mode of their respective distributions, because there are some longish utterances of *ă* that are in fact longer than some shortish utterances of *ā*.[33] Hence on a pure acoustical basis without a knowledge of the remainder of the permutation (or of the mode of the distribution), the phoneme *ă* is not always kept distinct from all other phonemes in its language.*

Penetrating even further into the minutiae of a phonetic system, and particularly into the question of varying lengths, we call attention to the brilliant empiric study of W. P. Lehmann and R–M. S. Heffner who found that the utterance of a given vowel tends to be longer before a voiced final stop than before a voiceless one (e.g., longer before *d, b,* and *g* than before *t, p,* and *k*).[34] Since, as the experimenters pointed out, the voiced stops are the comparatively less frequent ones, there is really a negative correlation between the length of a vowel and the frequency of the following stop—a correlation which they find valid also in reference to the decreasing frequencies of *d, b,* and *g*. Thus the vowel before the more frequent *d* is shorter than before the less frequent *g* (the reader can note the progressive lengthening of the vowel in his pronunciation of the successive syllables of *tot, tod, tog,* or in *pit, pid, pig*). This inverse relationship between the length of a vowel and the frequency of its following final stop (which the above experimenters ascribe without specific definition to differing "skills" in pronouncing the final consonants) ties up with the empiric research of Drs. Hudgins and Numbers.[35]

Thus Drs. C. V. Hudgins and F. C. Numbers in their pioneer work in investigating the speech of the deaf have observed that the errors made by the deaf in pronouncing vowels and consonants increase as the relative frequency of the vowels and consonants decrease. In this inverse relationship between error and frequency of occurrence the above experimenters propound a *pragmatic scale of difficulty*. Tying this observation to the previous Lehmann-Heffner observation, we might suggest that as a speaker approaches a rarer and more difficult final stop, his stream of speech slows down by way of preparation, with the result of a corresponding increase in the length of the preceding vowel. The Bonn phoneticist, Dr. Menzerath, and Dr. de Lacerda of Portugal, among others, have demonstrated the influence of a following phonetic entity upon the pronunciation of preceding ones.[36]

In discussing these various detailed phonetic problems as illustrations of an economy of effort, we must not overlook the basic problem of the

* The emergence of Zwirner's normal curve is interesting dynamically. We suggest, among other reasons, that it is the *speaker's economy* which tends to shorten the utterance, and the *auditor's economy* which tends to lengthen it. The result of these two opposing "Forces" might well be a normal curve. But see the following discussion of varying lengths and of errors. In any event the minimal magnitude of any phonetic entity would seem to be determined by what can be heard in normal speech situations.

selection of particular phonetic entities for particular permutations in order to save work for the auditor by making the succession of words of the stream of speech as different as possible, because obviously a too close juxtaposition of homophones, or even of "otherwise homophonous" words might confuse the auditor. Hence we may suspect the existence of a regulating principle that governs the distribution of phonetic entities among permutations so that the stream of speech will be richly variegated and completely unambiguous. As a result, successive words would be kept apart not by one but by many different phonetic entities in many different positions of permutation, so that the misarticulation of a phoneme would lead to little or no misunderstanding on the part of the auditor. This topic of the "phonetic variegation" of the stream of speech was kindly investigated by my then student, Mr. Frank Piano, who analyzed the 2,544 different words in a sample of 11,538 running words of R. C. Eldridge's list number 2. After reducing the 2,544 different words to the International Phonetic Alphabet, he found that 23 pairs of homophones fell together, leaving only 2,521 phonetically different words. With the removal of the phonetic difference between *l* and *r* (so that *call* and *core* would fall together) and also between *m* and *n* (affecting words like *some* and *son*), the number of different words dropped only to 2,481. With the further removal of all differences between short vowels (so that *bit, bet, bat, but, bought* would fall together) the number dropped only to 2,460. Then (in desperation) Mr. Piano let all long and short vowels and diphthongs fall together—with the result that without any vocalic differentiation whatsoever there were still 2,264 different phonetic forms. From this study it is apparent that, in actual samples of speech, words are kept apart not by one but by many differences, and that words that are phonetically very similar are not likely to appear often in the same context for obvious reasons of economy from the auditor's point of view. This problem seems to merit further investigation.[37]

Zipf (1935) Psycho-biology 109-121

The average Frenchman today in pronouncing *rive* (from Latin *ripa*) and *fève* (from Latin *faba*) and *vie* (from Latin *uita*) does not in any way reveal in the articulation of the three *v*'s the different consonants from which they originated. The shift of a phoneme from one form to another in the phonemic pattern is such a slow process of approximation, often extending over considerable time, that there is often likely to be an interval in the metamorphosis where classification is impossible.* Between the norm for *t* and the norm for *d* in English there is an interval in which a speech-articulation may occur as a deviation from either norm. Historical study can often indicate the origin of the phoneme in question, yet only its future behavior can determine the category to which it pertains.

Although it is at times difficult to categorize a given speech articulation at the time of its occurrence, nevertheless the dynamic forces behind the phenomenon are not difficult to detect, if the present theory of relative frequency and equilibrium be true. For, in the terms of our theory, every *assimilation points to a weakening or instability of the assimilated sound, and this weakening or instability is caused primarily by the excessive relative frequency of the assimilated sound.*

f. Frequency Thresholds of Toleration of Phonemes

Until this point in our investigation of the phoneme our method of analysis has been primarily inductive rather than deductive, and by this method we have found a state of equilibrium in the phonemic system of languages which seems to be maintained by various types of phonetic change. These same conclusions, as we shall now see, may be ob-

* Hence, a dynamic philologist who is making a frequency count of phonemes in a living language like modern English must be careful to indicate the doubtful phonemic norm of phonemes which seem to be in the process of gradual metamorphosis. It should be borne in mind, however, that doubtful phonemes of this type may be statistically insignificant.

tained by an *a priori* approach, which has the virtue of shedding additional light on the nature of the phoneme, by establishing the probable existence of *thresholds of toleration* of phonemes. These thresholds of toleration are nothing more than limits to the relative frequency of a phoneme, above which a phoneme will tend to weaken (abbreviatory change, page 92f.) and below which it will tend to strengthen (augmentative change, see page 113 f.).

In commencing an analysis of these thresholds, we must remember a very obvious fact: every language must possess a sufficient variety of discernibly different vowels, consonants, and other phonemic units, so that permutations of the same, together with other resources (such as accent, tones, syntax), can adequately express its body of concepts. Upon this statement, which is axiomatic, follows a corollary equally self-evident: no one phoneme in a phonemic system can have an unlimited relative frequency up to 100 per cent, for the simple reason that a 100 per cent relative frequency of any phoneme would preclude the existence of any other phoneme in the phonemic system. Since there is clearly an ultimate limit to the relative frequency of any one phoneme, the question arises whether this ultimate limit is the same as the presumed threshold of tolerable frequency for a phoneme. To make the discussion of this problem more tangible, let us refer it to some hypothetical language.

For instance, let us assume for the sake of argument that a language existed in which every word began with a *d*. Is it not difficult to believe that an idiosyncrasy of this sort could persist indefinitely? The ever-present initial *d* of this hypothetical language would cease to be a signally characteristic part of any word, inasmuch as every word possessed it. The *d*, being in no way peculiarly characteristic of any one word, would be completely unessential to a perfectly adequate conveyance of the meaning of any word. In short it would have no cogent reason for persisting as a symbol. The speaker might pronounce it out of sheer habit, but not from the exigency of distinguishing different meanings. Similarly the auditor would have no cogent reason for insisting urgently upon normal care in articulating this initial *d*. The particular permutation of phonemes which distinguished one word from another, would not include the initial *d* as an essential, or indeed even as an unsuperfluous element. It would be quite comprehensible if the articulation of this *d* were constantly neglected in the stream of speech and became weakened (an abbreviatory change, partial or complete, see page 92 f.). But would not the same neglect and weakening of *d* occur if, instead of being everywhere in the initial position, it were everywhere in the final position, or in the middle position in every word? Indeed, it would scarcely be necessary that every spoken word in the stream of speech had a *d* initially, medially, or finally before it began to be sufficiently superfluous to suffer neglect. The phoneme *d* would seem superfluous if every word contained it somewhere. Such a frequent phoneme would be a far less important characteristic of a word than, say, a phoneme which occurred only once in ten thousand words. In the short and frequently occurring words especially, the ever-present *d* would seem much less significant than any of the remaining phonemes. However, let us go further and assume that one half or three fourths or some other equally large fraction of the consonants in use were *d*'s, scattered initially, medially, and finally. Or suppose that the three hundred most frequently used words contained at least one *d*. Surely with the proportion of *d*'s so high, the *d* would tend to weaken.

Hence, there must presumably be some percentage of relative frequency, or *upper threshold of toleration* in a language above which the *d* will tend to weaken. Although we do not know what this threshold is, its existence is quite probable. Furthermore, the above discussion of *d* applies correspondingly to every other phoneme in the phonemic system of a given language: every phoneme must presumably have an upper threshold of frequency above which it cannot pass without tending to weaken.

But when we say that any phoneme, for example *d*, will tend to weaken in a dialect when its relative frequency passes a certain upper threshold of toleration, the question arises as to the possible effect of this weakening on the subsequent form of the phoneme. The answer is again the abbreviation of articulatory sub-gestures. If the *d*, for example, is proportionately too frequent in the beginning of words, weakening may occur by a tardiness in voicing it. The initial voiced stops of many languages [1] exemplify this tendency toward tardiness in voicing, which is absent in other occurrences of the voiced stops. This tendency may reach such a point that the voicing is entirely deleted. For example, the German dialectal word *teutsch* 'German' resulted from excessive tardiness in voicing the initial *d* of *deutsch*. Similarly if the *d* is proportionately too frequent at the end of words, weakening may be shown in ceasing the voicing before articulation is completed. This neglect of voicing of final voiced stops may again reach such a point that it is deleted. Thus, in German, final *d* is always voiceless; the historical *d* of *Tod* 'death' and final *t* of *tot* 'dead' are today indistinguishable in standard German pronunciation. Naturally the positions and manner of weakening of a phoneme are numerous; by the abbreviation of respective articulatory sub-gestures, a *d* may assume a form which may be described as an *n*, an *ð*, a *t*. And these forms (e.g. *n*, *ð*, *t*), resulting from partial abbreviation, may rightly be viewed as weaker, less complex, and even more economical than the *d* from which they weakened.

It follows from the above that all phonemes need not, indeed cannot, have the same actual percentage-thresholds of toleration; for if a phoneme *x* weakens because of excessive relative frequency to *y*, then the new phoneme *y* must *ipso facto* be capable of sustaining a higher relative frequency than *x*. In other words, if *d* weakens to the voiceless stop *t* because of excessive relative frequency, then this *t* can tolerate a higher threshold of relative frequency than *d*. Although this particular *t* which has weakened from a *d* can tolerate a

greater relative frequency than a *d*, yet the *t* too must presumably have an upper threshold of frequency, and for precisely the same reason that *d* must have an upper threshold of frequency. If *t* surpassed its upper threshold, it too would tend to weaken, that is, tend to abbreviate one or more of its articulatory sub-gestures. For example, the occlusive gestures might be abbreviated in such a way that there would be no explosion; in this case the tongue would, say, not move so far, and the loss of explosion would represent a true weakening. The resultant weakened form would then be a spirant, probably similar to the English *þ* (written *th* in English *think*). By the same reasoning, there must also be an upper threshold of frequency for *þ*, which, like a vowel, may vary in duration. Indeed, if the frequency of any phoneme increases too much, it may weaken so far as to be completely dropped (i.e. deleted) from the stream of speech.*

We are led by the same manner of analysis to another conclusion: every phoneme must also have a lower threshold below which it cannot pass without strengthening. For we can reverse the argument and suppose that the relative frequency of a phoneme, say *t*, is abnormally low, so low that there are only a few *t*'s appearing in the stream of speech, so few indeed that *t* occurs only very rarely. The phoneme *t* would then become a distinctive and very characteristic part of every word in which it occurred, pronounced carefully by the speaker, heard distinctly by the auditor. It is quite conceivable that the speaker, in taking care to pronounce the *t* distinctly, would unconsciously add a following aspiration, or spirant, or some other element. Of course the speaker would not intentionally add an *h* or an *s* or some other element to the *t*. The speaker would merely unconsciously tend to articulate the *t* more carefully; the additional *s*, *h*, or other appended element would be a fortuitous excrescence, a kind of

* Especially weak phonemes, like *h*, are therefore especially susceptible to loss; e.g. the loss of *h* in Latin in developing into Romance, Latin *habēre* but French *avoir*.

accidental epenthesis — the result of an over-careful articulation. Since any tendency in this direction might easily become a consistent tendency, the infrequent *t* might develop into an aspirate t^h, or an affricate *ts*. Surely a *t* which has developed into a *ts* or t^h gives every indication that it was by no means neglected in the stream of speech.*

Hence a phoneme has not only an upper threshold but also a lower threshold. If its frequency surpasses the upper, the phoneme 'weakens'; if the lower, it 'strengthens.' It seems permissible, therefore, to infer likewise from the existence of lower thresholds that many phonemes may be viewed as potential strengthenings and weakenings of other phonemes.

Seemingly quite conclusive proof of the existence of these upper and lower frequency thresholds of toleration is offered by the percentages of voiced and voiceless stops for the twelve languages presented in the table on page 75. In each pair of corresponding voiced and voiceless stops, with the exception of the Spanish *d* and *t* and the Hungarian *b* and *p*, we find not only that each voiceless stop much outnumbers its voiced stop, but that the percentages of similar stops throughout the twelve languages are on the whole amazingly similar. For instance *t* is approximately 7 per cent, *d* approximately 3.5 per cent; the percentages for English and Bulgarian are especially close. Furthermore, a similar correspondence is evinced by the percentages for *m* and *n* on page 79. Of course, more extensive and refined phonemic analyses of many of these languages might well reveal an, even closer correspondence. However, we must be prepared to expect minor differences among the percentages of similar stops in

* An example of a general strengthening of *t* to *ts* is tendered by the Old-High-German sound-shift in which a Germanic *t* shifted to a *ts* in the majority of positions (though in many cases it went even further into *ss*, see page 119 infra). E.g. primitive Germanic *t*, still preserved in English, *two*, became *ts* (written z) in German *zwei*. Examples can be multiplied indefinitely, *ten, zehn; tug, zug; tooth, zahn*. That the probable cause of this was a decrease in relative frequency below the lower threshold, see page 120.

different languages if only because the phonemes in the different languages may vary slightly in their normal magnitudes of complexity. Furthermore, in remembering that this is but a statistical law which merely states probabilities of behavior we must, it appears, be prepared to find that an occasional phoneme continues to exist though its relative frequency be appreciably above its upper or below its lower frequency threshold of toleration. For no matter how very convenient it would be if we had absolute percentage-thresholds for every phoneme, above which or below which a phoneme would instantly change in form, the simple fact remains that at least so far as our present findings are concerned there is no absolute threshold.[1] On the basis of our findings we are justified in saying only this: as a given phoneme approaches a threshold, the chances favor the appearance of an instability in the form of the phoneme which will lead to change; and as the relative frequency of the phoneme surpasses a threshold by ever more and more, the chances of its change become ever greater and greater. Though the chances may be a hundred to one in favor of change, the eventuation of no change in one case out of a hundred is no disproof of our statistical law. And finally we must not be misled by a very striking and probably deeply significant relationship among our percentages. For instance, in the Peipingese aspirated and unaspirated stops, the ratio of the percentages of frequency of aspirated to unaspirated is on the whole approximately 1 to 3; with the Danish aspirated fortes and unaspirated lenes the ratio is on the whole approximately 2 to 3; with the voiced and voiceless stops of the twelve languages the ratio is on the whole approximately 1 to 2. These ratios are strikingly simple and seem in accord with Nature's frequent fondness for simple relationships. Nevertheless the simplicity of these ratios has not been accounted for in any way in our investigation.[2] We have shown only that the total magnitude of complexity of a phoneme bears some inverse relationship to the relative frequency of occurrence; this inverse relationship may of course be directly proportionate or it may be some non-linear mathematical function — the present investigation remains noncommittal on that point; whoever succeeds in measuring quantitatively the magnitude of complexity of phonemes without respect to the relative frequency of their occurrences will be able to give us more precise information on that point (see pages 58 ff.). In short, our thresholds of frequency are only approximate and indicate only probabilities of behavior.

Yet even approximate thresholds which indicate only probabilities of behavior may be highly useful [1] and we shall now see that with the help of our thresholds of toleration many apparent exceptions to the rule of relative frequency may be explained.

g. *The Apparent Exception of Spanish Dentals and Other Phonemes*

In the table of percentages for the relative frequencies of occurrence of the voiced and voiceless stops in twelve languages (page 75) there were only two pairs in which the relative frequency of the voiced stop was greater than that of its corresponding voiceless stop, i.e. the Spanish dentals and the Hungarian labials. Evidence will now be advanced to show that the Spanish dentals, though apparently exceptional, conform in all probability to what we may term the principle of relative frequency. Whether the Hungarian labials also substantiate this principle, or whether they are to remain the sole exception in the entire tabulation cannot be decided by one who is as unfamiliar with the historical development of Hungarian as is the present writer. Hence we shall restrict our immediate attention to the Spanish dentals.

When the statistics were first discussed (page 76), Spanish *d* with a percentage of relative frequency of 5.20 per cent and

Spanish *t* with a percentage of 4.27 per cent seemed to offer an exception to our general findings and evidence against our rule. Now we shall see that the excessive frequency of Spanish *d* offers in all probability a confirmation. If 5.20 per cent represents the frequency of *d* in Spanish, we may plausibly assume that it has crossed the upper threshold,[1] either because it is the most frequent *d* in the column, or because it is almost 25 per cent more frequent than the Spanish *t*. We should, therefore, expect that in Spanish the *d* would tend to lose some of its articulatory sub-gestures, as happens to be actually the case. According to the Spanish phoneticist T. Navarro Tomás,[2] only in the absolutely initial position, or when preceded by *n* or *l* is the written *d* pronounced as a voiced stop, and then only with weak articulation. Elsewhere it has lost its explosiveness, becoming a spirant (ð or þ). In many cases the spirant is so weak as to be neglected entirely in the vulgar pronunciation current in the greater part of Spain (see footnote, page 113). Hence the weakening of an excessively frequent phoneme seems to be illustrated by Spanish *d*.

The frequency of Latin *m* with 5.82 per cent (see page 79) is another point in question. Not only does it have the highest relative frequency for *m* of any of the languages, but it is over twice as frequent as the least frequent, and nearly one third more frequent than the next most frequent (i.e. Burmese *m*, 4.72 per cent). Faithful to our expectations, Latin *m* subsequently weakened, particularly in the final position where it eventually vanished. The extent to which the total relative frequency of *m* in Latin was reduced by the loss of final *m* in all occurrences is indicated by the fact that 56 per cent of the total occurrences of *m* in the Latin analysis were final; * from this one change alone, all else being equal, the total relative frequency of Latin *m* would decline from 5.82 per cent to 2.55 per cent, coming well below any upper

* It is to be remembered that the maintenance of final *m* in Latin was often merely orthographic.

threshold. The case of *m* in Latin not only illustrates a weakening attendant upon crossing an upper threshold, but it offers a valuable example of the weakening of a phoneme in one position made especially vulnerable by a concentration of occurrences in that position. It need scarcely be pointed out that initially and medially Latin *m* remains remarkably stable (witness present-day French *me* from Latin *me*, French *ami* from Latin *amicus*).

The Cantonese velars (page 71) are another point in question, especially the unaspirated voiceless stop which with a percentage of 8.7 per cent is more frequent than the corresponding dental (cerebral) stops of that dialect (*t* is 6.14 per cent). It may be remarked that in final position in Cantonese, *k* is often replaced by the glottal stop.[1]

These three typical examples of weakening suffice to illustrate the effect upon the form of a phoneme when it transgresses the upper threshold. The general phenomenon of these three examples is clearly the same as the general phenomenon discussed from the point of view of abbreviatory phonetic changes (pages 92 ff.).

h. Lower Thresholds and Augmentative Phonetic Changes

On page 92 we classified determinable phonetic changes into abbreviatory and augmentative phonetic changes, depending upon whether the total magnitude of complexity was abbreviated or augmented by the change. The determinable changes discussed until now have all been examples of abbreviatory changes, a class which appears to be far more frequent in occurrence than the augmentative. Indeed, abbreviatory changes are so predominant in the histories of languages that some early scholars were again and again tempted to explain phonetic change solely on the basis of simplification or attrition. However, to the minds of later scholars, the instances of augmentative change were sufficiently abundant to lead to the belief that the abbreviatory changes were the exceptions and that the probable cause of phonetic change was the desire for increased complexity. And so convincingly has each side argued that the comparative philologists of today, insofar as they are concerned with the dynamic problems of the phoneme, may perhaps be said to be divided into three camps:[1] (1) the proponents of greater simplification, (2) the proponents of greater complexity, and (3) those who reserve judgment.

Augmentative phonetic changes may be expected, according to our theory of relative frequency and equilibrium, whenever a phoneme becomes so rare as to cross its lower threshold. The Slavic languages are said to provide many examples of augmentative changes. But since our immediate interest is more in illustration of principle than in marshalling examples, the single instance of the shift of voiceless stops to voiceless affricates in many of the dialects of Old High German will perhaps suffice.

In Old High German the Germanic *d* changed in form * until it became a phoneme similar to *t*; thus the Germanic *d* preserved in English *do* appears in High German as *t* in *tun* because of this change. Similarly the Germanic interdental spirant preserved and written today in English as *th* (e.g. *that* or *think*) became a *d* in Old High German, appearing in present-day German as *d* in the words *das* and *denken*. The Germanic *t* of English *to, eat, heat* changed according to its position in Old High German words and according to the particular Old High German dialect either to the affricate *ts*

* While it is convenient in exposition to say that Germanic *d* changed to *t* in Old High German, it is more accurate to say that Germanic *d* changed until its form was one of *t*. Though this qualification may perhaps seem pedantic, it is nevertheless a more accurate description of the phenomena involved in this, or any other phonetic change. For convenience only do we speak of 'phonemic types'; actually there is in a phonemic system a gradation of countless steps in the matter of magnitude of complexity, and hence of form. In strengthening or weakening to restore balance, a phoneme passes up or down this scale until it attains a point where equilibrium is restored, regardless of whether or not this point of equilibrium falls within a well-recognized 'phonemic type.'

(often written as *z* in present-day German), or *s*, *ss* (generally written *ss* today); e.g. German *zu* 'to,' *essen* 'to eat,' *Hitze* 'heat.' Because of similar changes in Old High German, the English voiceless stops *p* and *k* appear in German as *pf*, *ff*, and *k*, *ch*, respectively (e.g. English *pipe*, German *Pfeife*, Old English *cirice* 'church' and German *Kirche* 'church,' borrowed from the Greek *kuriakon*). These changes differed slightly in different Old High German dialects, and were more stable in some dialects than others. Our chief interest in them at present, however, is not one of dialect geography, but rather that they were all 'spontaneous' phonetic changes (see page 88), and all augmentative changes (see page 92) except the change of *d* to *t*.

Before turning to a statistical analysis of a sample of Old High German prose, let us formulate what we may theoretically expect to find statistically. In the first place, the new *t* from older *d* should have a percentage approximating 7 per cent, or roughly 3 per cent more than the upper threshold of *d* which we maintain was crossed in the change. The new *d* from older *þ* and *ð* should have a percentage approximating 4 per cent (the Spanish *d* which has crossed its upper threshold and is weakening to *þ* and *ð* has 5.20 per cent, see page 116). The affricates which developed from *t*, *p*, and *k* respectively should have only minimal percentages to justify such a severe augmentative change.

In an analysis of samplings from Tatian's *Gospel Harmony* (*Evangelienharmonie*),[1] totalling 50,000 phonemes in extent and reduced to a uniform orthography consistent with the phonemic data derived from our knowledge of the origin and subsequent development of Old High German, we find that the new *t* has an occurrence of 7.77 per cent; the new *d* has an occurrence of 5.38 per cent, the affricates from *t*, *p*, and *k* having the low frequencies of 1.87 per cent, .11 per cent, and .39 per cent respectively. In other words, we find actually what we have anticipated theoretically.

Expressed differently, it may be said that if the original Germanic *d* had undergone no change in Old High German, its percentage of 7.77 per cent would have been excessively above what appears to be approximately the usual upper threshold of *d*. On the other hand, the Germanic voiceless stops, had they remained unchanged, would have been well below their lower thresholds. In view of these considerations, which do not seem to be the results of random chance, the inference appears quite plausible that the abbreviatory and augmentative changes which occurred in Old High German were to restore equilibrium, in whatever way that equilibrium may have been originally disturbed.

i. Analogy as a Coercive and Restraining Factor in Phonetic Change; the First Germanic Sound-Shift

There is, however, another factor which may operate in phonetic change: analogy. Since analogy is often important in accentual changes where the peculiar nature of its behavior is perhaps most readily apparent (pages 159 ff.), we shall at present merely illustrate the manner in which analogy may coerce or restrain the behavior of a phoneme in respect to a phonetic change, thereby in itself disturbing what the normal course in preserving or restoring equilibrium would otherwise be. Since the influence of analogy upon phonetic change has never received, even from comparative philologists, the attention which it deserves, the territory is therefore practically virgin, and this brief discussion of it may be viewed merely as a beginning in the direction of exploration.

The clearest example of analogic phonetic change is the familiar First Germanic sound-shift (i.e. phonetic change), described under Grimm's Law, in which all the voiced, voiceless and aspirated stops of Indo-European were changed.

Zipf (1935). Psycho-biology
73-81

4. THE RELATIVE FREQUENCY OF VOICED AND VOICELESS STOPS

In turning to a consideration of the phonemic series of voiced and voiceless stops we have the task of discovering whether in the stream of speech of each of the languages possessing these phonemic series there is any preference for the voiced or voiceless stops in each pair. If we find a clear and unmistakable correlation between high relative frequency on the one hand and voicing or voicelessness on the other, we shall be justified, it seems, in attempting to reduce to a common denominator the phenomenon of voicing with the previously discussed (page 68 ff.) phenomenon of aspiration.

A note of caution must be sounded in respect to using the ensuing tabulation. Strictly speaking, the transcriptions of these different languages fall into three different classes. The percentages for the first three languages (Dutch,[1] Czechish,[2] and French [3]) are derived from *accurate phonemic* transcriptions of these languages; the next three languages (Italian, English, and Hungarian [4]) had their samplings transcribed accurately according to a *phonetic* system of transcription; the remaining six (Bulgarian, Russian, Spanish, Greek, Latin and Vedic Sanskrit [5]) represent transcriptions into the customary *alphabet* of each respective language. Except for Greek, Latin, and Sanskrit, all of these languages are still living and the frequency distribution of any of their phonemes may be analysed again if data in a different form are desired. This note of caution must be sounded because, strictly speaking, the complete results of a phonetic analysis, a phonemic analysis, and an analysis of an alphabetic transcription of a language will by no means be necessarily identical in all respects.*

* At this point, still another note of caution seems to be in order. Dynamic philologists must also beware the often exaggerated statements of phoneticists, especially of professional phonetic transcribers, who generally carry their refinements unnecessarily far from our point of view. For example, the actual number of occurrences of the six different stops in the respective samplings of the twelve languages below would in some cases not be appreciably different whether one selected the phoneme as a unit of transcription, or the variant form, or the alphabetic symbols conventional for each language. Thus, if a given sample of English text were analysed first phonemically and then phonetically and then alphabetically, the differences in actual occurrences of, say, *p*, as adduced by these three methods of transcription, would probably be insignificant. In instances, such as *k*, where the phoneticist would indicate the occurrences of variant forms of a phoneme by the consistent use of different symbols, one need but add these together to obtain the total occurrences of the entire phoneme irrespective of phonetic variant forms, a procedure which has been followed as far as possible in all of the phonetic investigations below. If the conventional alphabet is at times quite inconsistent, as in present-day English and Irish, it is not necessarily inconsistent in respect to all speech-elements (e.g. the English alphabet is quite consistent, on the whole, in its use of, say, *b*). Moreover, not all conventional alphabets are of necessity inconsistent; the

Let us now turn to a consideration of the statistics which are given for each language in terms of percentages of the whole of that language. The table is to be read horizontally and not vertically, since we are not yet comparing languages but only the pairs of corresponding stops in each language by itself. The phoneme at the head of each column designates the general phonetic type; we remember that the actual pronunciations may vary from language to language.

	t	*d*	*p*	*b*	*k*	*g*	
Czechish	5.60	3.73	3.52	1.86	3.93	.15	phonemic
Dutch	7.83	4.67	1.99	1.20	(3.21)*	(.09)*	phonemic
French	6.28	3.55	3.54	1.39	4.81	.76	phonemic
Italian	7.02	4.74	2.78	.89	3.63	.41	phonetic
English	7.13	4.31	2.04	1.81	2.71	.74	phonetic
Hungarian	7.18	3.30	1.04	1.71	5.72	2.45	phonetic
Bulgarian	7.54	3.55	2.82	1.32	2.98	1.46	alphabetic
Russian	7.49	3.42	2.19	1.76	3.49	1.10	alphabetic
Spanish	4.27	5.20	2.64	2.05	3.82	.07	alphabetic
Greek	7.58	2.87	3.38	.49	4.07	1.74	alphabetic
Latin	7.72	3.41	2.01	1.40	3.71	.96	alphabetic
Sanskrit	6.65	2.85	2.46	.46	1.99	.82	alphabetic

conventional Sanskrit alphabet is an amazingly accurate phonetic alphabet, practically *nulli secundum*; on the other hand the conventional Czechish alphabet is practically as accurate a phonemic alphabet as can be devised for Czechish. This note of warning, which was foreshadowed on pages 52 ff., is necessary lest we be persuaded into the erroneous belief that none of our otherwise highly valuable records of past and present speech can ever be utilized by Dynamic Philology for the almost absurd reason that their transcriptions were not made with the use of the most refined phonetic system of symbolization. In approaching these statistics these notes of caution must be borne in mind.

* The absence or presence of voicing in the Dutch speech-element indicated above by *k* and *g* is not phonemically significant in Dutch, the voiced form being but a variant of the voiceless.[1] The fact that these two speech-elements cannot be compared like the others has been indicated by percentages in parentheses.

The first three languages (Czechish, Dutch, and French), which afford the best material, unmistakably show that the voiceless stops are more frequent than their corresponding voiced stops in each language. The material of the remaining nine languages, though not as accurate as that of the above, points definitely in the same direction with the exception of the Spanish dentals and the Hungarian labials. We shall later find (page 116 f.) that the Spanish dentals, when viewed from the angle of dynamic development in Part II, cease to be an exception. The exceptional relationship of the Hungarian labials will probably not be found altered in an accurate phonemic transcription. Save for the Spanish dentals and the Hungarian labials, the unvoiced stops on the whole are appreciably more frequent than their corresponding voiced stops. Out of thirty-five opportunities for deviation from this general tendency, there are only two actual deviations. Hence this relationship as evidenced by these data from twelve languages is 94.3 per cent valid.

5. THE RELATIVE FREQUENCY OF OTHER PHONEMES WHOSE MAGNITUDES OF COMPLEXITY ARE IN PART COMPARABLE

The question naturally arises at this time as to the possibility of comparing other phonemes in respect to magnitude of complexity. For the stops which we have examined, though an important category numerically in the phonemic systems of languages in which they occur, represent neither a majority of available phonemes nor a majority of occurrences of phonemes.

In the case of monophthong-vowels there is frequently the opportunity in many languages of comparing corresponding long and short vowels. In some cases (e.g. ă and ā in Sanskrit or German) the phonemic difference of sheer length is a

sufficient indication of differences in magnitude of complexity. Thus in the case of German ā and ă, e.g. *kān*, (written *Kahn*) 'boat,' and *kăn* (written *kann*) 'can,' the ā may be said to represent a greater magnitude of complexity than ă, because ā represents everything that ă represents, *plus* added duration. In the case of many other pairs of long and short vowels in many other languages, however, the difference is more than that of mere duration. For example, in the English pair ī and ĭ (e.g. *machīne* and *ĭt*) the position of the vocal organs is by no means identical and hence the comparative magnitudes of the two are not determinable even though it might be found that the long vowel has a greater average duration than the short. Though we may and probably correctly do feel that the long vowel represents a greater magnitude of complexity than the short, the validity of this subjective feeling would be practically impossible to prove empirically except by assuming (which is not yet permissible) that relative frequency is indicative of comparative magnitude, and by showing, as will doubtless be found true, that the short vowels have on the whole an appreciably higher relative frequency of occurrence than their corresponding long vowels. At present there are but few reliable data available on the subject in any language. What we have on the subject of the phonemes ā and ă is from Vedic Sanskrit,[1] which show that the phoneme of greater magnitude of complexity is also of lower relative frequency (e.g. ā = 8.19%, ă = 19.78%).

It is also perhaps possible to compare monophthong-vowels with diphthongs which contain them, e.g. ă with ăi and ău, ŏ with ŏi and ŏu, etc. Whether it can be correctly said that the magnitude of complexity of the typical vowel ă is the same when occurring in the typical diphthong ăi as when occurring alone is doubtful. But on account of our general subjective feeling that the magnitude of complexity of a diphthong is greater than that of its component parts when occurring alone, it may be stated that the few available and reasonably reliable statistics that exist on the subject indicate clearly that the diphthong, which presumably is of greater magnitude of complexity than that of each of its parts when occurring alone, is also relatively less frequent (e.g. Sanskrit *ăŭ* = .18%, *ă* = 27.97%;[1] German diagraph *au* = 2.242%, German letter *a* = 13.147%,[2] the German percentages reckoned on the basis of all occurrences of vowels in Kaeding). This new and interesting subject[3] merits the attention of further and more accurate investigation.* The topic does not, however, justify our including in the text here the illuminating statistics for Modern Icelandic vowels and diphthongs.[4]

The problem of comparative magnitudes of complexity of spirants (e.g. *f*, *v*, *þ*, *ð*, *s*, *z*, etc. in English) is truly difficult, even in languages where the presence or absence of voicing is phonemically significant, because of the extremely variable factor of duration. Investigation of these phonemes may, therefore, not be particularly rewarding.

With the *trills* (e.g. the *apical trill* or 'rolled *r*' of Italian, or the *uvular trill* or 'uvular *r*' of Danish) as well as with the *laterals* (i.e. the *l*'s) of a language there is not only the variable factor of duration, but also frequently the absence of any other phoneme suitable to serve in comparison.

With the nasals, *m* and *n*, it might be taken as some evidence that *n* is the simpler of the two phonemes because of the observations[5] of comparative philology which indicate that quite often, when *m* disappears in any of its usages in a given language, it becomes (i.e. 'weakens' to) *n* before disappearing. Whether we may safely conclude from this fairly frequent phenomenon that *n* is therefore, on the whole, a phoneme of smaller magnitude of complexity than *m* is clearly a controversial matter. In any event available statistics[6] reveal that *n* is, on the whole, appreciably more frequent in occurrence than *m* (Burmese being the only exception):

* The magnitudes of complexity of *ĭ*, *ĕ*, *ă*, *ŏ*, *ŭ*, etc., when compared with one another are at present indeterminable; similarly with *ī*, *ē*, *ā*, *ō*, *ū*, etc.

	m	*n*		*m*	*n*
Czechish	3.52	6.42	Latin	5.82	6.47
Dutch	3.18	7.09	Sanskrit	4.34	7.04
French	2.56	3.19	Peipingese	2.18	10.18
Italian	3.11	6.25	Cantonese	4.07	5.70
English	2.78	7.24	Burmese	4.72	4.15
Hungarian	3.35	5.74	Swedish	3.28	7.32
Bulgarian	2.22	7.00	Danish	3.18	5.70
Russian	3.12	5.13	Singhalese	3.12	7.40
Spanish	2.29	6.08	Old English	2.81	8.40
Icelandic	4.37	7.77	Old High German	2.91	10.85
Greek	3.19	8.55			

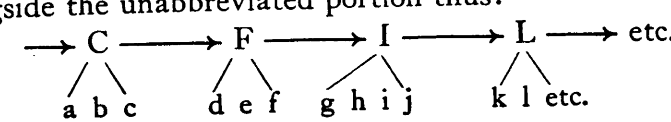
6. CONCLUSION: EQUILIBRIUM

If we pause now to reflect upon the statistics which have been presented in respect to the relative frequency of occurrence of many different phonemes in many different languages[1] that are but distantly related, if not altogether unrelated, and whose speakers in most cases belong to different quarters of the globe, or to different ages, or to different national cultures, we find two interrelated phenomena. First, it is clearly evident that the frequency distribution of phonemes in the stream of speech is by no means completely a matter of random chance but that the relative frequency of occurrence of a phoneme depends to a considerable extent upon its form. And second, wherever the comparative magnitudes of complexity of phonemes are determinable, the magnitude of complexity bears an inverse (not necessarily

Zipf (1935) psycho-biology of Lg
"Velocity of communication"
283-287

8. ABBREVIATION: THE BASIC DIMENSION OF SPEECH

The chief difference between abbreviated and unabbreviated speech is that the latter is more articulated in its meaning. We may represent an unabbreviated portion of the stream of speech by some such sequence as *abcdefghijklmnopqrs* in which each letter stands for a given word. Abbreviated, this same portion of the stream of speech may be represented, say, by *CFILOR*, — that is, six elements instead of nineteen. The abbreviated portion refers to whatever the unabbreviated portion refers to, only it is shorter and, if all else is equal, makes the reference more swiftly. The abbreviated portion might be represented alongside the unabbreviated portion thus:



That is, while the more fundamental meaning is the stream *abcdefghijkl, etc.*, the actually occurring stream of speech is *C F I L, etc.* in which each capital letter (e.g. *C*) refers to a configuration of small letters (e.g. *abc*).

Now, assuming that both the average magnitude of complexity and the average rate of utterance of each speech-element, whether abbreviated or unabbreviated, is the same, we may view the abbreviated stream as a short-cut through the unabbreviated stream, whether the abbreviation is one of truncation or substitution. Abbreviation is then actually a short-cut; and moreover, since the stream of speech knows no other arrangement than that of time, *an abbreviation of speech is a short-cut in time.*[1] Furthermore, since by means of abbreviation ground is covered more rapidly than in the absence of abbreviation, the belief is plausible that *the more extensive the abbreviation, the greater is the velocity of the stream of speech.* Though the actual minutes consumed may be the same both for 100 connected words representing abbreviations of more articulated language, and for 100 words of unabbreviated articulated language, the former is moving at a greater velocity than the latter from the point of view of effective communication. If the speeds of the two appear the same, they are nevertheless no more the same than those of an airship flying a mile high and a bird flying in the same direction at a lower level and at the same apparent speed; though the bird in its flight may remain in the line of sight between the observer and the airship, the actual speeds of the two are different.

For the phenomenon of abbreviation, convenient analogies can be found in the field of mathematics. When, for example, in the stream of speech $y\ c^2\ z$ serves as the abbreviation for *yabcdez*, one may conceive of *yabcdez* as an arc on the circumference of a circle, whose extremities, *y* and *z*, are connected by the straight line yc^2z. This manner of representation is descriptive of the isolated phenomenon though its employment would be difficult in describing subsequent

abbreviations. More adequate analogies could be drawn from spherical geometry.[1] Perhaps the most extreme use of abbreviation is in algebra, where a single letter may be used, at the arbitrary choice of the mathematician, to symbolize a whole paragraph of verbal description.

In saying that the stream of speech has a rate of velocity which varies according to the extent to which it is abbreviated (i.e. the more abbreviated, the greater the velocity) we do not mean that one necessarily utters words more rapidly in the stream of high velocity, but that one covers more ground in effective communication in less time. Furthermore, no matter how abbreviated the stream of speech may become, it can, theoretically at least, always be more abbreviated because any two or more events in sequence may be represented symbolically by a third. Conversely, no matter how great the degree of articulation and concomitant slowness of the stream of speech, it may, theoretically at least, be ever more articulated and of slower velocity. Manic and obsessive language (pages 217 ff.) represent unusual velocities, the first above average, the second below.

There seems to exist nothing in the observed nature of speech which permits one to deduce an absolute and constant velocity, in terms of which differing velocities may be measured. The velocity of one stream of speech is fast or slow only in comparison with that of another. The velocity of any stream of speech is then relative; and any language pattern or speech-element is, in terms of the velocity of its occurrence, relative.

Moreover, since the function of language in actual communication depends upon the representative use of symbols, the relative velocity attained through abbreviation is the basic dimension in language development. Our examples have mostly illustrated the abbreviation of a configuration of words; the same principle applies to the abbreviation of any other speech-configuration, no matter how large or how small, whether the components are articulatory sub-gestures, phonemes, morphemes, words, sentences, etc.[1] The most fundamental unit of speech may well prove to be a unit of velocity. However, that there may be no confusion, let us view in somewhat greater detail what is meant, first, by a dimension of speech, second, by velocity as a dimension of speech, and finally, by a unit of the dimension of velocity.

Now, a *dimension* is a measurable extent of any kind, as length, breadth, thickness, area, volume; in algebra it may also be each one of a number of unknown quantities contained as factors, say, in a product (e.g. the product *vwxyz* may be viewed as possessing five dimensions). The mathematical comparison of two automobiles, for example, may include not only the three conventional dimensions of space, but a fourth for weight, a fifth for speed, a sixth for power, and so on. Each dimension may have its unit (e.g. the foot a unit of the dimension of length, the horsepower a unit of the dimension of power, a mile per hour the unit of the dimension of speed). Similarly the stream of speech may have many dimensions (e.g. pitch, amplitude, etc.);[2] but it also has the one that we have mentioned — velocity.

That is, the stream of speech is representative of experience, but it consumes far less time as a symbolic representation of experience than occurred in actual experience. For example, Caesar's *ueni*, *uidi*, *uici*, 'I came, I saw, I conquered,' represents in a few seconds a portion of Caesar's experience of such duration that the few seconds of the symbolic representation are but an almost infinitesimally small fraction. And the more abbreviated speech is, the more is time telescoped or compressed or condensed; and it is this telescoping of time which is a fundamental characteristic of speech, and which is also probably the significant dimension of speech. The unit of speech is the ratio of the time of experience to the time of the gesture (i.e. configurations thereof) in conveying the experience adequately in communication. In other words, the basic unit of speech is a unit of time, and not one of meaning or phonetic articulation.

If this conception of a speech-unit in terms of speed is valueless as a practical measure, it seems nevertheless to be the only measure obtainable by empirical study and may be serviceable in exposing the fallacy of the easy belief that some particular speech-element (such as the phoneme, word, or sentence) is the basic unit.