From: *Library Trends* 30(1) (Summer 1981): 53-64

Empirical and Theoretical Bases of Zipf's Law

Ronald E. Wyllys
Associate Professor, Graduate School of Library and Information Science, The University of Texas at Austin

Introduction

One of the most puzzling phenomena in bibliometrics—and, more broadly, in quantitative linguistics—is Zipf's law. As one commentator, the statistician Gustav Herdan, has put it: "Mathematicians believe in [Zipf's law] because they think that linguists have established it to be a linguistic law, and linguists believe in it because they, on their part, think that mathematicians have established it to be a mathematical law."[1]

Let us start by considering a basic form of Zipf's law. Suppose one has a natural-language corpus, e.g., a book written in English. Next, suppose one makes a frequency count of the words in the corpus, i.e., counts the number of occurrences of *the, and, of*, etc. Finally, suppose one arranges the words in decreasing order of frequency so that the most frequent word has rank 1; the next most frequent, rank 2; and so on.

For example, a frequency count of the 75 word-types (i.e., dictionary entries) represented by the 112 word-tokens (i.e., distinct occurrences) in the two preceding paragraphs yields the partial results shown in table 1. This set of rank-ordered frequency counts, though quite small for the purpose, serves moderately well as an illustration of the fact that rank and frequency have a surprisingly constrained relationship in natural-language corpora. The values of the products of rank r and frequency f fall in the relatively limited range 27-30 in the middle of table 1, and we may note that there was no a priori reason for us to expect that the middle products rf would fall within so limited a range.

TABLE 1

Word-Type	***Rank r***	***Frequency f***	***Product rf***
the	1	9	9.0
in, of	2-3, mean = 2.5	7	17.5
a, one	4-5, mean = 4.5	6	27.0
Law	6	5	30.0
and, it	7-8, mean = 7.5	4	30.0
suppose, that, Zipf's	9-11, mean = 10.0	3	30.0
(21 words)	12-32, mean = 22.0	2	44.0
(43 words)	33-75, mean = 54.0	1	54.0

The constrained relationship between the frequency of a word in a corpus and its rank gained wide attention in the 1930s and 1940s through the work of George Kingsley Zipf (1902-1950), a professor of philology at Harvard University. The name "Zipf's law" has been given to the following approximation of the rank-frequency relationship:

$$rf = c \qquad (1)$$

where r is the rank of a word-type, f is the frequency of occurrence of the word-type, and c is a constant, dependent on the corpus (often around one-tenth of the total size of [i.e., number of word-tokens in] the corpus).

When stated algebraically, Zipf's law is usually given in the form of equation (1), but the law is probably most familiar in the graphic representation of a mathematically equivalent form:

$$\log r + \log f = \log c \qquad (2)$$

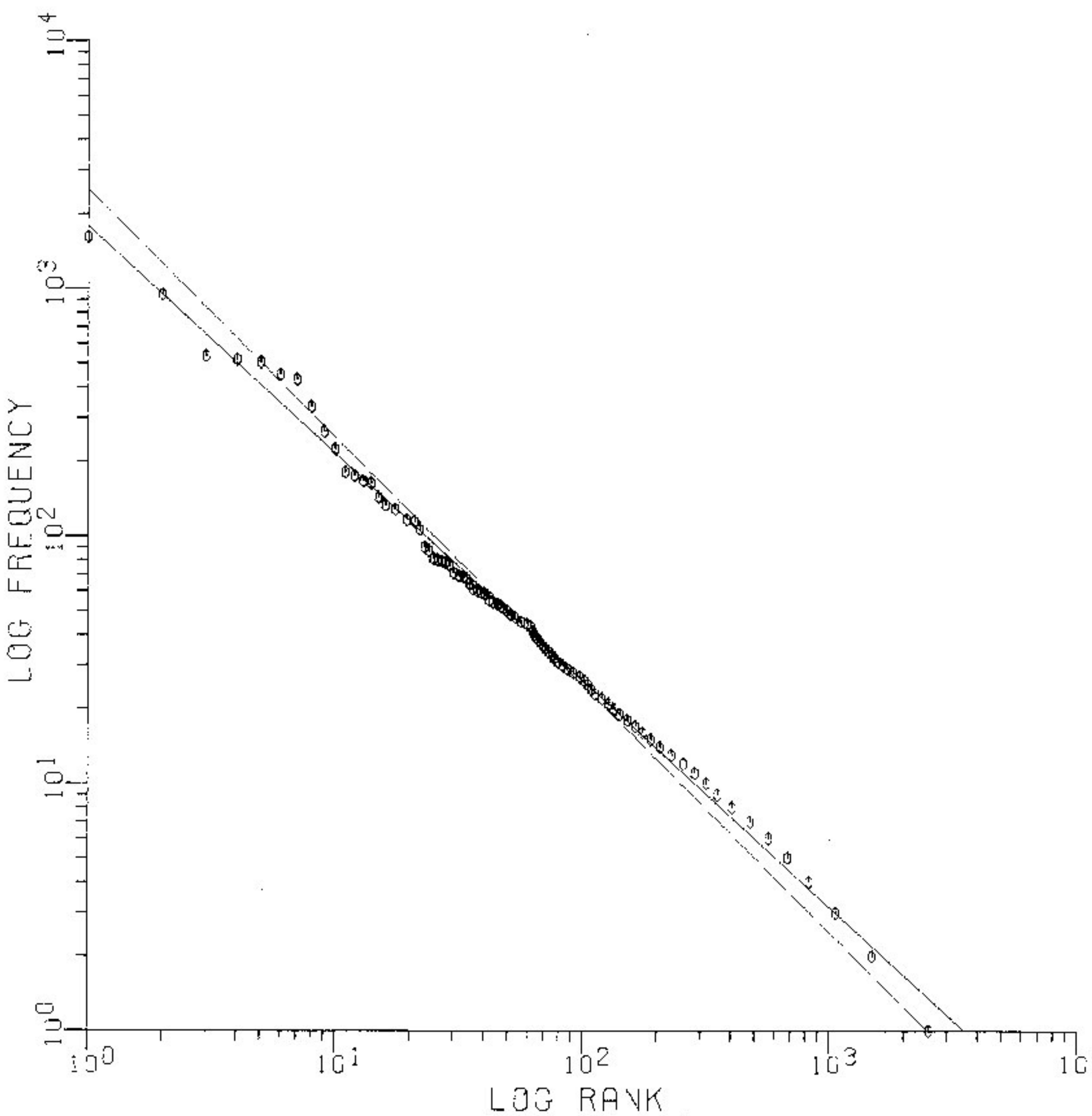

Figure 1. Observed Rank-Frequency Pairs for a Corpus of 21,354 Words
The solid line is the regression line for the data and has slope -0.92; the dashed line has slope -1.0.
Source: Wyllys, Ronald E. *The Measurement of Jargon Standardization in Scientific Writing Using Rank-Frequency ('Zipf') Curves*. Ph.D. dissertation, University of Wisconsin—Madison, 1974.

The dashed line in figure 1 illustrates what an idealized display of Zipf's law in the form of equation (2) might be, More generally, analytic geometry tells us that the equation of an arbitrary line whose slope is -B can be written as:

$$B(\log r) + \log f = \log c \qquad (3)$$

One such line is pictured by the solid line in figure 1, which has a slope of -0.92. (The relationship of this line to the data points will be discussed later.) If we write equation (3) in a form like that of equation (1), we have:

$$r^B f = c \qquad (4)$$

Note that if B takes on the particular value 1, then equation (4) becomes identical with equation (1). Thus, equation (4) is a generalization of Zipf's law, and we shall refer to it as the "generalized Zipf's law."

It should be noted that Zipf's law only approximates the relationship between rank r and frequency f for any actual corpus. Zipf's work[2] shows that the approximation is much better for the middle ranks than for the very lowest and the very highest ranks, and his work with samples of various sizes[3] suggests that the corpus should consist of at least 5000 words in order for the product rf to be reasonably constant, even in the middle ranks.

If one performs a frequency count on an actual corpus, arranges the words in decreasing order of frequency, and draws the resulting pairs of points by plotting the logarithm of rank on the horizontal axis and the logarithm of frequency on the vertical axis, the resulting points will form a slightly curved line. Such plots are known as "Zipf curves." An example of a Zipf curve is shown in figure 1.

One can speak of the "slope" of a Zipf curve by finding a straight. line that closely approximates the points of the curve and then taking that straight line's slope as the slope of the curve. Apparently Zipf himself fitted straight lines to his data by visual judgment only. Finding their slopes to be ordinarily close to -1, he appears to have assumed that the "true" slope of such curves was -1 and, hence, that equations (1) and (2)—rather than the more general equations (3) and (4)—were correct. This assumption is questionable, as will be discussed later.

The study of Zipf's law can be broken into three areas: (i) the initial discovery that equation (1) does approximate the relationship between rank and frequency, (ii) investigation of whether a better approximation exists, and (iii) attempts to provide a satisfactory rationale for the close relationship of rank and frequency.

The Discovery of Zipf's Law

The work that led to Zipf's law started when Zipf was a graduate student at Harvard in the 1920s. Studying phonetic changes in languages, he became interested in the frequency of use of phonemes as a factor in their tendency to change phonetically over long periods of time. From the relative frequencies of phonemes, he moved to studies of the relative frequencies of words, and in 1932 published a book, *Selected Studies of the Principle of Relative Frequency in Language*.[4] Of the approximately 125 pages in this book, over 100 are either diagrams or lists of words and their frequencies. About 22 pages are devoted to prose, which includes this passage of justification:

> Some have taken exception to the *Principle of Relative Frequency* simply because it is statistical. For statistics are hateful to the human mind; they are painfully definite for the group without being particularly definite for the individual. Undoubtedly, a primary law which knows no fluctuation within itself is pleasanter. If nature had consulted man in the latter, we should all have suggested primary laws. . . .But nature did not consult us . . . and has seen fit to let the laws of chance govern vast portions of the basic order of the physical universe, as well as no small amount of the biological.[5]

It is interesting to note that, unfortunately, the critics of quantitative analysis are still very much with us nearly fifty years later.

In his next book, *The Psycho-Biology of Language*,[6] published in 1935, Zipf called attention for the first time to the phenomenon that has come to bear his name. This book contained Zipf's first diagram of the *log(frequency)*-vs.-*log(rank)* relationship, a Zipf curve for his count of words in the Latin writings of Plautus.

Zipf's last book, *Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology*,[7] appeared in 1949. As its title indicates, this work is an exposition of what Zipf considered the fundamental reason for much of human behavior: the striving to minimize effort. The diversity of phenomena to which Zipf was able to apply his mathematical models, equations (1) and (2), is impressive.

Despite his strong defense of quantification, Zipf really did not argue in quantitative terms. It is true that he performed counts of linguistic phenomena, tabulated the counts, and displayed them. But his mathematics were weak, and his energies were spent in philosophizing about the implications of his principles. Support for this comment may be found in another passage from *Selected Studies*: "Before returning to linguistic considerations, let me say here for the sake of any mathematician who may plan to formulate the ensuing data more exactly, the ability of the highly intense positive to become the highly intense negative, in my opinion introduces the devil into the formula in the form of [the square root of -1]. And now to linguistics."[8]

Zipf appears to this writer to have been poorly trained for dealing with quantitative phenomena. His knowledge of mathematics was minimal; of statistics, apparently nonexistent. He never showed interest in exploring the quantitative nature of his data beyond noting that they came close to his model of the moment. This done, he would launch into lengthy speculations about hazily defined possible causes. It is a pity that he almost never collaborated with statisticians. On the other hand, he was an indefatigable worker, and pursued the rank-frequency phenomenon and related ideas for twenty years despite often harsh criticism. There can be little doubt that the ubiquity of these phenomena would be less well recognized were it not for his work.

Alternative Forms of Zipf's Law

In *Human Behavior and the Principle of Least Effort*, Zipf presented an interesting exception to his usual insistence that the slope of linguistic Zipf curves is -1, i.e., that only equation (1), and not equation (4), applies to linguistic data. He noted that frequency counts of the language of schizophrenics showed a different slope, commenting that "of all the rank-frequency data on words that have ever come to the attention of the present writer, only those of [two schizophrenics] have negative slopes . . . greater than unity."[9] Considering how poorly straight lines of slope -1 fit most of Zipf's other examples, one wonders why he found the departures of the schizophrenics' slopes from -1 to be remarkable.

In fact, the slopes of Zipf curves, when measured more carefully than by Zipf's eye, turn out to be capable of considerable divergence from -1. An obvious way of fitting a straight line to a Zipf curve, i.e., to a set of pairs of observations of *log(frequency)* and *log(rank)* for a corpus, is by linear regression, with *log(rank)* playing the role of the independent variable. A study by the present writer using this technique found slopes ranging from -0.89 to -1.04 among only eight corpora.[10] Figure 1, taken from this study, shows a plot of *log(frequency)* vs. *log(rank)* for a corpus of 21,354 words from issues of the *Psychological Review* for 1969, together with the regression line of best fit to these points. The regression line, shown as a solid line, has a slope of -0.92; for comparison, figure 1 also shows a dashed line whose slope is -1.

In general, diagrams of the *log(frequency)*-vs.-*log(rank)* relationship for natural-language data typically show a downward concavity for the low ranks. The full set of products rf typically shows a fairly consistent slow rise in the values of rf as r increases, rather than any readily identifiable constant value. Thus, equation (2) seems to represent actual data less accurately than does the generalized Zipf's law, equation (4):

$$r^B f = c \qquad (4)$$

where $B < 1$. Note that if the product rf gradually increases with increasing r, the effect of giving r an exponent that is less than 1 will be to make r^B increase less rapidly than r, thus helping to keep the product $r^B f$

more nearly constant. This will tend to hold the left-hand side of equation (4) more or less in balance with the constant-valued right-hand side.

For the reasons just sketched, it seems clear that one should not expect equation (1) to be as satisfactory a description of Zipf curves for actual data as is equation (4) with B expected to differ from 1 ordinarily. Benoît Mandelbrot has published several studies of generalizations of Zipf's law, dealing both with the question of whether the slope is -1 and with the deeper problem of explaining why the rf products should be relatively constant (his work on this latter problem will be discussed later). Mandelbrot seized upon the idea that B could vary, and he related B to the diversity of a corpus (viz., the ratio of the number of word-types to the number of word-tokens in the corpus), holding that B tended to vary inversely with the diversity.[11]

Mandelbrot also developed a further refinement of Zipf's law:

$$(r + m)^B f = c \qquad (5)$$

where r is the rank of a word, f is its frequency, and m, B, and c are constants dependent on the corpus.[12] The key idea in this version is that m has its greatest effect when r is small, and that equation (5) therefore provides a better fit to typical data, especially to the low-rank, high-frequency words, than do equations (1) or (4).

An even more general formulation of the relationship of rank and frequency is due to H. P. Edmundson, whose "3-parameter rank distribution"[13] is:

$$f(r; c, b, a) = c(r + a)^{-b} \quad c > 0,\ b > 0,\ a \geq 0 \qquad (6)$$

where f is the frequency associated with rank r, and where a, b and c are constants. Equation (6) contains Zipf's and Mandelbrot's versions as special cases.

The Search for a Rationale for Zipf's Law

Why should there be such a surprisingly constrained relationship between rank and frequency for natural-language corpora? The problem is more complicated than this question suggests. There are many other phenomena that exhibit similar distributions; Abraham Bookstein has provided two unifying surveys of them.[14] Commenting on the ubiquity of such distributions, Herbert A. Simon has mentioned "distributions of scientists by number of papers published . . . of cities by population . . . of incomes by size, and . . . of biological genera by number of species."[15] He observed that "one is led to the conjecture that if these phenomena have any property in common it can only be a similarity in the structure of the underlying probability mechanisms."[16] At present, it is probably fair to say that there is not yet complete agreement about why these phenomena share similar distributions or why the distributions exhibit the behavior known as Zipf's law.

Zipf thought the reason lay in his Principle of Least Effort, which he defined as follows:

> The Principle of Least Effort means . . . that a person will strive to solve his problems in such a way as to minimize the *total work* that he must expend in solving *both* his immediate problems *and* his probable future problems. That in turn means that the person will strive to minimize the *probable average rate of his work-expenditure* (over time). And in so doing he will be minimizing his *effort,* by our definition of effort. Least effort, therefore, is a variant of least work.[17] (Italics in original.)

Unfortunately, Zipf never provided a clear logical development from this principle to equation (1).

Intellectually much more satisfying than Zipf's principle is the approach of Mandelbrot, who used ideas from information theory to explain the rank-frequency phenomenon. The essence of Mandelbrot's contribution was his considering communication costs of words in terms of the letters that spell the words and the spaces that separate them. This cost increases with the number of letters in a word and, by extension, in a message. Mandelbrot showed that Zipf's law, equation (1), follows as a first approximation from the minimization of communication costs in terms of letters and spaces. Linguistically, this amounts to minimizing costs in terms of phonemes, which is why the phenomenon holds for both written and spoken language. Mandelbrot's more accurate second approximation has been shown in equation (5).

Many attempts have been made to provide other rationales for the Zipf phenomenon. Most of them are probabilistic in their approach; i.e., they consist of derivations, from various premises, of the probability that a word will occur with a certain frequency in an arbitrary corpus. The frequencies can, at least in concept, be ranked and thus be made to imply probabilities that a certain rank *r* will be associated with a certain frequency *f;* however, the implication may be difficult to make explicit. In the space available here, only the nature of these attempts can be sketched; the principal goal is to emphasize their variety and, hence, the inconclusive current state of explanations of Zipf's law.

One such attempt involved the combined efforts of Gustav Herdan, J. O. Irwin, and an eighteenth-century British mathematician, Edward Waring. Herdan[18] presents the model as:

$$p_f = \frac{x-a}{x} \qquad \text{for } f = 1 \qquad (7.1)$$

$$p_f = (x-a)\frac{a(a+1)(a+2)...(a+f+2)}{x(x+1)(x+2)...(x+f-1)} \qquad \text{for } f = 2,3,... \qquad (7.2)$$

where p_f is the probability that a word will appear with frequency f in a large corpus, and a and x are constants, dependent on the corpus, such that $0 < a < x$. The function is due to Irwin,[19] who discovered it in a search for distributions useful in biology, and who credited Waring with discovery of the basic inverse factorial expansion underlying the probability function. Since it was Herdan who recognized that Irwin's result had linguistic applications, the function has come to be known as the Waring-Herdan formula in linguistics. Several investigators have reported that it fits observed rank-frequency data well. Good fits to observed rank-frequency data by another model, the lognormal distribution, have been reported by V. Belevitch[20] and John B. Carroll.[21]

Bruce M. Hill[22] and Michael Woodroofe[23] have pursued the derivation of a probabilistic form of Zipf's law by applying Bose-Einstein and Maxwell-Boltzmann statistics to the classical occupancy problem. A similar derivation has been offered by Yuji Ijiri and H. A. Simon.[24] These papers employ various initial conditions to yield various of the Zipf, Bradford and other related distributions. The interrelatedness of these distributions has been shown by, *inter alios,* Bertram C. Brookes[25] and Robert A. Fairthorne.[26]

A different starting point has been suggested by H. S. Sichel. He assumes that "each word in . . . [an] author's vocabulary has a long-term probability of occurrence."[27] The mixing of thousands of such probabilities during the production of speech or writing can be expressed as a compound Poisson probability, of which "a number of known [distribution functions] such as the Poisson, negative binomial, geometric, Fisher's logarithmic . . . Yule, Good, Waring and Riemann distributions are . . . limiting forms."[28] Sichel reports very close fits of his model to some twenty published frequency counts. A related paper by B. C. Brookes[29] treats a model of "a very mixed Poisson process," and another article by Brookes and José-Marie Griffiths[30] derives from this process a "frequency-transfer coefficient" as a means of measuring the correlation of frequency and rank. Empirical tests of the theories are sufficiently rare that reports of such tests by Beth Krevitt and Belver C. Griffith[31] and by Anita Parunak[32] deserve mention.

The negative binomial distribution has been the starting point for other investigations, including one by B. M. Hill treating the number-of-species problem hut mentioning its relation to Zipf's law.[33] A major effort along these lines is that of Derek de Solla Price, who has developed a modification of the negative binomial that he calls the cumulative advantage distribution (CAD). In the CAD the conditions of the negative binomial are modified "so that success *increases* the chance of further success," but unlike in the negative binomial: "failure has no subsequent effect in changing probabilities. . . .Failure does not constitute an event as does success. Rather it must be accorded the status of a 'non-event'; thus lack of publication is a non-event and only publication becomes a markable event."[34] Rephrasing this for words rather than publications, we can say that if at a certain point in writing a corpus an author uses a given word, it seems plausible that the chance of his or her using that word again in the corpus is increased, whereas the author's failure to use some other word at that point says essentially nothing about the chance that this other word will be used later in the corpus. As a probability density function for the CAD, Price derives a modified Beta function. Further comments on the CAD have been made by Paul B. Kantor, Price, and I. K. Ravichandra Rao.[35] Closely related is the "contagious Poisson process" of Paul D. Allison.[36]

Conclusion

What is our present state of knowledge about Zipf's law? Its remarkable range of applicability to diverse phenomena continues to amaze us, but we have come far along the road toward an understanding of why it should exist and why it should be so widespread.

It seems intuitively plausible that some kind of general Poisson process should underlie the pervasiveness of Zipf's law and its siblings, such as the Bradford and Lotka laws discussed elsewhere in this issue. After all, these laws deal with phenomena that we can characterize as consisting of the occurrence of events whose individual probabilities are ordinarily quite small and, hence, can be expected to behave in a Poisson-like fashion. Even Zipf's hazy Principle of Least Effort can be interpreted as a groping toward a Poisson process, in that the principle suggests that people find it easier to choose to use familiar, rather than unfamiliar, words and that the probabilities of occurrence of familiar words are therefore higher than those of less familiar ones.

On the other hand, it is clear that the process cannot be a pure Poisson process, since the choices of words are not independent, as the Poisson distribution requires. Already in 1955 Simon recognized this in employing a stochastic model "in which the probability that a particular word will be the next one written depends on what words have been written previously."[37]

Practically all the work on developing a rationale for Zipf's law has involved probabilistic models related to the Poisson in some fashion. Among these models is Price's cumulative advantage distribution, which the present writer finds very persuasive. Research on a rationale for Zipf's law has not yet achieved a consensus, but we are probably close to one.

What implications does Zipf's law have for the design of information systems? The honest answer has to be: few, if any. So far as vocabulary control is concerned, Zipf's law offers no useful information beyond what frequency-counts alone can easily supply. The present writer has suggested that different subject-fields may be characterized by different slopes of Zipf curves,[38] but again this possibility seems to have no practical applications at present in information system design. Perhaps such applications will develop in the future. Meanwhile, we can continue to surprise ourselves with the ubiquity of the Zipf phenomenon and to enjoy the intellectual challenge of achieving a full, rational understanding of it.

References

1. Herdan, Gustav. *The Advanced Theory of Language as Choice and Chance.* Berlin: Springer-Verlag, 1966, p. 33.

2. *See, for example,* Zipf, George K. *Human Behavior and the Principle of Least Effort.* Cambridge, Mass.: Addison-Wesley, 1949. Reprint ed., New York: Hafner, 1965.

3. Ibid., p. 291.

4. ________ , *Selected Studies of the Principle of Relative Frequency in Language.* Cambridge, Mass.: Harvard University Press, 1932.

5. Ibid., p. 9.

6. ________ , *The Psycho-Biology of Language*. Boston: Houghton Mifflin, 1935. Reprint ed., Cambridge, Mass.: MIT Press, 1965.

7. ________ , *Human Behavior*.

8. ________ , *Selected Studies*, p. 21.

9. ________ , *Human Behavior*, pp. 295-96.

10. Wyllys, Ronald E. *The Measurement of Jargon Standardization in Scientific Writing Using Rank-Frequency ('Zipf') Curves.* Ph.D. diss., University of Wisconsin-Madison, 1974.

11. Mandelbrot, Benoît. "Structure formelle des textes et communication." *Word* 10(1954):1-27, 424-25.

12. ________ , "An Informational 'Theory of the Statistical Structure of Language." In *Communication Theory: Papers Read at a Symposium on "Applications of Communication Theory,"* edited by Willis Jackson, pp: 486-502. London: Butterworths, 1953.

13. Edmundson, Harold P. "The Rank Hypothesis: A Statistical Relation between Rank and Frequency." Technical report TR-186. College Park: Computer Science Center, University of Maryland, 1972.

14. Bookstein, Abraham. "The Bibliometric Distributions." *Library Quarterly* 46(Oct. 1976):416-23; and ________ , "Explanations of the Bibliometric Laws." *Collection Management* 3(Summer/Fall 1979):151-62.

15. Simon, Herbert A. "On a Class of Skew Distribution Functions." *Biometrika* 42(Dec. 1955):425. Reprinted in *Models of Man: Social and Rational.* New York: Wiley, 1957, pp. 145-64. *See also* ________, "Some Further Notes on a Class of Skew Distribution Functions." *Information and Control* 3(March 1960):80-88.

16. Simon, "On a Class of Skew Distribution Functions," p. 425.

17. Zipf, *Human Behavior and the Principle of Least Effort,* p. 1.

18. Herdan, Gustav. *Quantitative Linguistics.* London; Buttcrworths, 1964, pp. 85-88.

19. Irwin, Joseph O. "The Place of Mathematics in Medical and Biological Statistics." *Journal of the Royal Statistical Society, Series 1* 126(1962):1-41.

20. Belevitch, V. "On the Statistical Laws of Linguistic Distributions." *Annales de la Société Scientifique de Bruxelles, Series 1* 73(18 Dec. 1959):310-26.

21. Carroll, John B. "On Sampling from a Lognormal Model of Word-Frequency Distributions." In *Computational Analysis of Present-Day American English,* edited by Henry Kučera and W. Nelson Francis, pp. 406-24. Providence: Brown University Press, 1967.

22. Hill, Bruce M. "Zipf's Law and Prior Distributions for the Composition of a Population." *Journal of the American Statistical Association* 65(Sept. 1970):1220-32; and

________ , "The Rank-Frequency Form of Zipf's Law." *Journal of the American Statistical Association* 69(Dec. 1974):1017-26.

23. ________ , and Woodroofe, Michael. "Stronger Forms of Zipf's Law." *Journal of the American Statistical Association* 70(March 1975):212-19.

24. Ijiri, Yuji, and Simon, Herbert A. "Some Distributions Associated with Bose-Einstein Statistics." *Proceedings of the National Academy of Sciences* 72(May 1975):1654-57.

25. Brookes, Bertram C. "The Derivation and Application of the Bradford-Zipf Distribution," *Journal of Documentation* 24(Dec. 1968).247-65.

26. Fairthorne, Robert A. "Empirical Hyperbolic Distributions (Bradford-Zipf-Mandelbrot) for Bibliometric Description and Prediction." *Journal of Documentation* 25(Dec. 1969):319-43.

27. Sichel, H. S. "On a Distribution Law for Word Frequencies." *Journal of the American Statistical Association* 70(Sept. 1975):543.

28. Ibid.

29. Brookes, Bertram C. "Theory of the Bradford Law." *Journal of Documentation* 33(Sept. 1977):180-209.

30. ________, and Griffiths, José-Marie "Frequency-Rank Distributions." Journal of the ASIS 29(Jan. 1978)5-13.

31. Krevitt, Beth, and Griffith, Belver C. "A Comparison of Several Zipf-Type Distributions in their Goodness of Fit to Language Data." *Journal of the ASIS* 23(May- June 1972):220-21.

32. Parunak, Anita. "Graphical Analysis of Ranked Counts (of Words)." *Journal of the American Statistical Association* 74(March 1979):25-30.

33. Hill, Bruce M. "Posterior Moments of the Number of Species in a Finite Population and the Posterior Probability of Finding a New Species." *Journal of the American Statistical Association* 74(Sept. 1979):668-73. (In this paper Hill mentions a relevant, unpublished work: Chen, Wen-Chen. *On Zipf's Law*. Ph.D. diss., University of Michigan, 1978.)

34. Price, Derek de Solla. "A General Theory of Bibliometric and Other Cumulative Advantage Processes." *Journal of the ASIS* 27(Sept.-Oct. 1976):293.

35. Kantor, Paul B. "A Note on Cumulative Advantage Distributions." *Journal of the ASIS* 29(July 1978):201-04; Price, Derek de Solla. "Cumulative Advantage Urn Games Explained: A Reply to Kantor." *Journal of the ASIS* 29(July 1978):204-06; and Rao, I. K. Ravichandra. "The Distribution of Scientific Productivity and Social Change." *Journal of the ASIS* 31(March 1980):111-22.

36. Allison, Paul D. "Estimation and Testing for a Markov Model of Reinforcement." *Sociological Methods & Research* 8(May 1980):434-53.

37. Simon, "On a Class of Skew Distribution Functions," p. 427.

38. Wyllys, *Measurement of Jargon Standardization.*

PDF version prepared: 2009 August 14