

STATISTICAL AND ENUMERATIVE
ANALYSES OF STYLE FOR
THE CLASSROOM

By

ROBERT JOSEPH HASENFRATZ

Bachelor of Arts

Oklahoma State University

Stillwater, Oklahoma

1979

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
The Degree of
MASTER OF ARTS
May, 1983



STATISTICAL AND ENUMERATIVE
ANALYSES OF STYLE FOR
THE CLASSROOM

Thesis Approved:

William H. Pixton

Thesis Adviser

Paul J. Kays

Eric Southard

Norman N. Durham

Dean of Graduate College

PREFACE

Every profession has its own select terminology, its jargon, and freshman composition is no exception. We have developed numerous and subtle means to label grammatical errors, to designate the components of an essay, to explain the methods of paragraph development. Strangely, though, most composition instructors are at a loss to describe a student's style in non-metaphorical terms: the style is either choppy, dull, vigorous, awkward, elegant, jerky, or smooth. In the following study, I have attempted a remedy for this lack. By this scheme, if a student (or instructor) asserts that the style of a particular piece of writing is "elegant," he is forced first to define the term elegance specifically and then to identify and tabulate all occurrences of it in the given text. Statistics possess no magic. In fact, counting the number of concrete nouns per 1000 words proves nothing except that the occurrence of such nouns is high or low, but through such a procedure, students discover one of the contributors to concrete language. Thus, the value of the technique is not that it provides "objective" proof for stylistic pronouncements; rather, it helps to confirm and illustrate valid statements about style or indicate those

which are not square with the facts.

For his advice and moral support, I wish to thank my major advisor, Dr. William H. Pixton, without whom this work might not have reached completion. I express my appreciation to Dr. Paul Klemp for his close reading of the manuscript and his timely stylistic advice, and to Dr. Bruce Southard, whose objectivity helped temper the graduate student's fondness for convolution. For the original impetus for this work, I am indebted to Dr. John Milstead.

A note of thanks is in order for Mike and Debi McDonough, Domenic Bruni, George Wittmer, and Sue Denman, none of whom provided the slightest bit of help or moral support, but who were great company during the long months of composition. I also wish to extend my gratitude to Mrs. Louise Thomas, who more than any one else has helped to rid this work of its more glaring typographical errors and to shape it into the proper format. Finally, I reserve my warmest appreciation for my grandparents, especially Beulah Coyle, and my parents, Wayne and Nancy, for their kind prodding, their financial support, and mostly their confidence in me.

TABLE OF CONTENTS

Chapter	Page
I. INTRODUCTION	1
II. SURVEY OF ENUMERATIVE AND STATISTICAL ANALYSES OF STYLE	9
Attributive Studies	10
Literary Studies	14
Pedagogical Studies	26
Developmental Studies	41
III. THEORY OF ENUMERATIVE ANALYSIS	51
Objectivity	51
Form Versus Content	59
Rhetorical Considerations	62
Application	71
IV. ENUMERATIVE TECHNIQUES FOR THE CLASSROOM	76
Close Imitation	77
Attribution	83
Definition	88
V. CONCLUSION	94
A SELECTED BIBLIOGRAPHY	96

LIST OF TABLES

Table	Page
I. Percentage of Initial Connectives in 2000- Sentence Samples of Addison, Johnson, McCaulay, and Swift	24
II. The Style Machine Criteria for Measuring Style	33
III. Classification of Parallelism Samples	41
IV. The Basic Purpose of Composition	66

CHAPTER I

INTRODUCTION

In most composition text books, stylistic pronouncements are vague, sometimes mystical. Consider for example Sheridan Baker's advice about diction in his acclaimed text, The Complete Stylist:

A good diction takes work. It exploits the natural, but does not come naturally. It demands a wary eye for the way meanings sprout, and the courage to prune. It has the warmth of human concern. It is a cut above speech, yet within easy reach. Clarity is the first aim; economy, the second; grace, the third; dignity, the fourth.¹

This advice is couched in an almost exclusively metaphorical language which neatly avoids the specifics of practical application: stylistic eyes are wary; meanings sprout like beans and are pruned; words hover a few inches above the colloquial but still within grasp; like a blue sky, good diction is clear; like a Scotsman, it is parsimonious; like Fred Astaire's dancing, it has grace; and like a king, it has dignity--and for all we know--regality as well. Metaphors of this kind will not do. But Baker does offer more advice: "You can choose the high word [the Latinate word], or you can get tough with Anglo-Saxon specifics."² To illustrate this point, Baker produces three lines of

Shakespeare and an incomplete sentence by Faulkner--sketchy proof--but conventional wisdom requires little evidence.

Unfortunately, the student is not provided with specific methods to perform his task. Is the student to buy an etymological dictionary and search about for "tough" Anglo-Saxon words or "high" Latinate ones? Baker does not say.

Sadly, this sort of advice is to be found even in the most respected texts. In Sylvan Barnet's and Marie Stubb's Practical Guide to Writing, the following instructions are offered to student writers who wish to avoid "Instant Prose," prose in which every word is not made to "count":

Trust yourself, Writing Instant Prose is not only a habit; it's a form of alienation. . . . Distrust your first draft. Learn to recognize Instant Prose Additives when they crop up in your writing. . . . Acquire two things: a new habit, Revising for Conciseness [the chapter title] and . . . a wastebasket.³

Here, Barnet and Stubb offer considerable moral support, but no specific definition of Instant Prose nor possible remedies for it. Essentially, the student is being asked to write honestly and to revise his work. These pronouncements, of course, do not offer any hitherto unknown keys to success in writing; they are common sense notions and vague notions. If only the word could become flesh; if only the dictum "write honestly" could produce honest writing--then textbooks and English composition courses would be unnecessary.

Unfortunately, composition courses are necessary, and

instructors must provide their students with methods by which writing can be improved. The syntactical units in composition texts appear at first to provide these specific methods. In the end, however, a few sentence types are paraded before the student, who is expected to duplicate the patterns to the best of his ability. The process takes place in a vacuum where a few isolated sentences, not a succession of sentences as found in actual texts, receive exclusive attention.

Commonly repeated stylistic advice seems to rely more on the force of authority than on any firm grounding in the realities of particular texts. In Milfred Stone and J. G. Bell's Prose Style: A Handbook for Writers, the student is informed that

Most good writers agree on five basic preferences:

1. Prefer verbs to nouns.
2. Prefer the active to the passive.
3. Prefer the concrete to the abstract.
4. Prefer the personal to the impersonal.
5. Prefer the shorter version to the longer.⁴

Who are these good writers? In which novels, articles, essays, or poems are these preferences embodied? The authors do not comment. Apparently, single-sentence examples (more than likely composed specifically for the occasion) are made to suffice. For the third preference, the following explanation is offered:

Writing that runs heavily to abstract nouns is hard to read, partly because such nouns tend to be long and lifeless, partly

because they take the tamer sort of verb (abstractions never kick or ogle or rever each other; they cause or refer to or consist of each other), but above all because they require the reader to invest time and effort in translating the writer's generalities into particulars.⁵

Again, the description here relies upon metaphor; abstract nouns are "lifeless" and take verbs which are "tame." The word representation is, by this scheme, dead while tomato or hog boils with life. However, it is difficult to imagine a piece of exposition dealing with poetry, nuclear power, or English composition which makes reference mainly to hogs, tomatoes, blood, and turnip tops. Any examination of good writing will reveal that--depending on the nature of the subject matter, audience, and writer's ethos and purpose--the frequency of concrete nouns increases or decreases. The following paragraph from the Stone and Bell text illustrates this point:

Paragraphs are not just hunks of prose marked by indentations; they are basic units of thought out of which an essay is composed. They are building stones, parts of a larger whole. Though we shall necessarily in this chapter discuss paragraphs without reference to their content, they are in fact inseparable from that context. To put this another way, the problem is not so much to write an effective paragraph, let alone a dazzling paragraph, as to write your paragraphs in such a way as⁶ to make an effective--and integrated--essay.

Of the twenty-one nouns in this excerpt, two concrete nouns of the hog variety appear: hunks and building stones. The others are either abstract (problem, content, thought, etc.) or ambiguously concrete (chapter, paragraph, etc.).

The implication is clear: the student should not at all times "prefer the concrete to the abstract," but only when the situation warrants, when there is a story to tell, an object to describe, or an abstraction to illustrate. Stone and Bell's preferences are not necessarily incorrect; however, unfounded as they are by any close investigation of sample texts, they can be misleading.

Many such common conceptions (and misconceptions) contained in style handbooks can be traced to one source: book three of Aristotle's Rhetoric. The insistence on concrete language comes from chapter eleven, where "actualization" or "putting things before the eyes" is recommended. The popularity of the periodic sentence may be traced to chapter nine. Baker's clarity and dignity have their parallels in chapters two and six respectively.⁷

Whether derived from Aristotle or his followers (Cicero and Quintilian), most stylistic pronouncements have gone unchallenged; in fact, they have appeared in generations of composition texts, often with only the slightest illustration. Though the scientific community has scrutinized and rejected many of Aristotle's themes, a great many rhetoricians and composition instructors remain content to let him rule in the classroom. But any successful procedure must be based on facts, not merely on an ancient but untested tradition.

Fifteen years ago, Francis Christensen challenged the commonly held notion of sentence opener variety (that

monotony of style ensues if a majority of the sentences in a text begin with the subject). By simply counting the occurrences of sentence openers in samples from ten professional writers and classifying them into four categories, Christensen discovered that in discursive prose seventy-five percent of the sentences begin with the subject.⁸ Of course, Christensen's enumerative methods can be called into question, but at least an avenue for discussion of the texts is opened. "Prefer the concrete to the abstract" and similar statements, expressed as they are in the imperative, preclude discussion. In fact, Louis Tonko Milic, whose work figures prominently in this present study, makes a case for objectivity in the description of styles:

A description of style, when it is not quantitative, can only be figurative. Such a description tends to rely on comparisons, analogies and similarly crude approximations. . . . A feature of style, whether it be a favored area of the vocabulary, a preference in imagery, a rhetorical habit or tendency to have recourse to certain syntactical patterns must be described in concrete and verifiable terms,⁹ which finally means, in quantitative terms.

This last assertion is the subject of this study. For intuition is a useful, indeed indispensable tool, but an intuitive judgment that is contradicted by the facts is of dubious value. Aristotle and his fellow rhetors arrived no doubt at many truths--of logic, of rhetoric--through an acute and sensitive intuition. These truths, if they prove to be truths, should not be abandoned.

Enumerative research, as applied to the study and teaching of style, shows some promise of substantiating or refuting the intuitive advice brought into the classroom. In fact, such an approach could provide a framework for student analysis of style. To explore the enumerative approach more fully in this study I shall provide a survey of enumerative and statistical methods, discuss the theoretical underpinnings of such studies, and finally provide classroom applications.

ENDNOTES

¹ Sheridan Baker, The Complete Stylist and Handbook (New York: Harper & Row, Publishers, 1980), p. 198.

² Ibid.

³ Sylvan Barnet and Marcia Stubbs, Practical Guide to Writing (Boston: Little, Brown and Company, 1980), p. 1.

⁴ Wilfred Stone and J. G. Bell, Prose Style: A Handbook for Writers (New York: McGraw-Hill Book Company, 1972), pp. 72-73.

⁵ Ibid., p. 77.

⁶ Ibid., p. 54.

⁷ Aristotle, "Art" of Rhetoric, trans. J. H. Freese in The Loeb Classical Library (Cambridge, Massachusetts: Harvard University Press, 1975), pp. 405, 392, 353, 375.

⁸ Francis Christensen, "Sentence Openers," in Notes Toward A New Rhetoric: Six Essays for Teachers (New York: Harper & Row, Publishers, 1967), p. 46.

⁹ Louis Tonko Milic, "Metaphysics in the Criticism of Style," College Composition and Communication, 17 (1966), 125, 126. Hereafter, reference to this article will be parenthetically cited under the short title "Metaphysics."

CHAPTER II

SURVEY OF ENUMERATIVE AND STATISTICAL ANALYSES OF STYLE

I shall distinguish between enumerative and statistical methods for the evaluation of style. Simply stated, enumerative techniques involve frequency counts with lower-level calculations of scientifically developed gradients such as the Characteristic (discussed briefly below). Since complex statistical analyses do not readily lend themselves to adaptations for student use, they will receive less attention in this study than the simpler, easily calculated enumerative techniques. (Of course, areas of overlap do occur.) Average sentence length, for example, is relatively easy to calculate, but the results can be submitted to a battery of highly complex statistical procedures.

Because the goal of this study is to develop classroom applications of objective techniques--not to produce scientifically acceptable procedures--the minute details of sample sizes, control and experimental groups, as well as random selection of samples will be de-emphasized. Here follows summaries of research under four headings: attributive, literary, pedagogical, and developmental

studies. (Both enumerative and statistical studies appear under these headings.)

Attributive Studies

Most often by complicated means, these attempt to attribute work of unknown origin to a particular author and are by far the most scientifically rigorous studies to be considered in this chapter. More often than not, the calculations involve probability quotients impossible to adapt for classroom use.

Late in the nineteenth century, during the height of the Shakespeare-Bacon debate, T. C. Mendenhall, a geophysicist, developed one of the first mathematical measures of style. In an article for The Popular Science Monthly, "A Mechanical Solution of a Literary Problem," Mendenhall postulated that the characteristic word length in a piece of writing could be just as definitive a test to determine the disputed authorship of a literary work as metallurgical tests determine the composition of ores. Other studies followed which either disputed the Mendenhall hypothesis or offered other measures (average sentence lengths, for example).¹ A few years after the opening volleys, Mendenhall published an exhaustively researched study of characteristic word length in both Shakespeare's and Bacon's works and came to the conclusion that Bacon could not have written the plays attributed to Shakespeare. Word length, tabulated by the syllable, is easy to

calculate and figures in some modern enumerative techniques.

Some forty years later in 1941, G. Undy Yule attempted to settle the disputed authorship of The Imitation of Christ, traditionally ascribed to Thomas A. Kempis, but thought by some to have been written by Jean Charlier de Gerson. Yule predicted that vocabulary richness (the extent to which an author uses new words or repeats old ones) could be calculated and could provide a valuable test for authorship problems. He employed probability quotients in his work, noting that the appearances of different word classes followed "accident distributions." A complicated formula was developed, the Characteristic, which--in simple terms--measures the extent to which certain words (in this case, nouns) are repeated.

Using this method, Yule calculated that the number of nouns that appear only once represent 82.4% of the total in De Imitatione Christi and only 35.0% in Gerson's theological works.² Eventually, he compared this relative vocabulary "richness" in different authors. Assuming that the vocabulary distributions remain constant throughout one author's corpus, the Characteristic (calculated from 10,000 word samples) could be used to distinguish between the works of different authors. Yule also calculated the average sentence length in both author's works and found--paralleling his later results--that A. Kempis seemed a much more likely candidate than Gerson: Gerson writes

longer sentences than those found in A. Kempis' writings or in *De Imitatione*.³

Gustav Herdan, in his important study, Language as Choice and Chance, modified Yule's Characteristic and developed the theory that certain language phenomena, such as the appearance of the letter b, can be measured statistically as chance happenings. His Entropy quotient, though unintelligible to the non-statistician, attempts to map the probability of such "chance" occurrences and to utilize the results for attribution problems.⁴

Six years later in 1963, Alvar Ellegard, a Swedish researcher, performed important research in order to settle the disputed authorship of The Junius Letters. Ellegard pointed out the flaws in Yule's work, especially that genre and topic influence stylistic choices in a way which Yule had not taken into account.⁵ Unlike his predecessors, Ellegard began with a purely subjective judgment: in A Statistical Method for Determining Authorship: The Junius Letters, he isolated individual words which seemed to him distinctive of The Junius Letters (these words are labelled "plus words"). He then developed a distinctiveness formula (D):

$$D = \frac{\text{relative frequency of a plus word in Junius}}{\text{relative frequency of the word in a million word sample of non-Junian writings}}$$

Words with high D-values (those that occur frequently in Junius but infrequently in other writings) were considered good discriminators. The Junian D-values for

certain plus words and expressions were compared with those in the writings of the other likely candidates for authorship. After all the counts were tallied, Sir Phillip Francis appeared to be the clear "winner." Ellegard's insistence that intuition is valuable makes his work important, but students could hardly hope to count million-word samples.

In 1963, Claude S. Brinegar, using the measure developed earlier by T. C. Mendenhall (word length distributions, that is, the percentage of one-syllable, two-syllable, three-syllable words, and so on) explored the probability that Mark Twain wrote "The Quintus Curtius Snodgrass Letters."⁶

Later that same year, Frederick Mosteller and David Wallace began examining the disputed authorship of twelve of The Federalist Papers.⁷ Mosteller and Wallace, aware of the influences of context and genre on stylistic choice, sought "context-free" variables which would remain at constant levels no matter what the subject matter. For this reason they chose to count the frequencies of function words (prepositions, determiners, and so on) which they considered to be subconsciously determined and thus not subject to changes in context. Study revealed that Madison's frequent use of the word "to" distinguished his from those of Hamilton, which contain relatively frequent occurrences of the word "by." Using these and other words as discriminators ("calibrating sets"), the

researchers determined that Madison is the likely author of the twelve disputed papers.

Louis Tonko Milic has also developed a technique whereby a case of disputed authorship may be decided. For his study, A Quantitative Approach to the Style of Jonathan Swift, Milic encoded a number of texts, some by Swift, some by other authors of the period, replacing each word with a code representing its part of speech. He then subjected the data to computer analysis which revealed, among other things, the percentage of nouns, verbs, verbals, adjectives, determiners, and so on. By comparing the differing percentages in the various authors' samplings, Milic developed a stylistic profile of Swift's writings and compared these, along with the profiles of the control authors, to a work of disputed authorship-- "A Letter to a Young Poet." This Discriminator Profile contains such measures as the percentage of verbals and introductory connectives.⁸ The results are not clear, but Milic feels confident enough to attribute the work to Swift, at least tentatively.

Literary Studies

Such studies seek the same ends as traditional stylistic studies of literature: to describe different styles and account for their significance. Most begin with a denunciation of impressionistic terminology and seek to describe style objectively; the conclusions which they

draw from these data, however, are often ambiguous.

The effort to make more scientific the exploration of literary style was begun in large part by Professor L. A. Sherman in the 1880's. Sherman's thesis was that the length of the English sentence was growing shorter and shorter. In his article "On Certain Facts and Principles in the Development of Form in Literature," Sherman expressed his confidence in scientific techniques:

The right way and the only to learn the facts and principles of English prose development was plainly to study the literature objectively, with scalpel and microscope in hand.¹⁰

Accordingly, Sherman counted and graphed the sentence lengths in Chaucer, Aecham, Lyly, Spenser, DeFoe, De Quincey, Macauley, Channing, Emerson, and many others. Surprisingly, Sherman found that an author's average sentence length remained relatively constant from section to section and from work to work--despite a wide range of sentence lengths. The results led Sherman to proclaim, perhaps overconfidently, that

. . . the evidence seemed to indicate the operation of some kind of sentence-sense, some conception or ideal of form which, if it could have its will, would reduce all sentences to procrustean regularity.¹¹

The implication here is that an author's characteristic sentence-length average is a good measure of overall style, but what exactly sentence-length tabulation measures is unclear. To Sherman, at least, the results were obvious: the movement in English prose has been a movement away

from long, complex, and subordinated sentences to shorter, coordinated, "oral" ones. On the basis of these findings, Sherman suggested that composition instructors, instead of endorsing complex "elephantine" sentences, should encourage students to write short, plain sentences.¹² Such a recommendation corresponds to Aristotle's call for clarity (perspicuity) in style.

The advantage to Sherman's enumerative approach is that sentence length is very easy to calculate. However, his sampling technique was called into question by R. E. Moritz, who found Sherman's failure to take differing genre and subject matter into account significant: "the sentence-constants varied not only when a comparison was made between drama and history, or essays, but in other forms of composition as well."¹³ This objection is an important one and will be discussed in the theoretical section below.

Perhaps because of the enormous task of tabulation necessary for analyzing entire periods of literary history, researchers began to take an interest in the styles of individual authors, W. E. Wimsatt, in his important work of 1941, The Prose Style of Samuel Johnson, supplemented his conclusions about Johnson's style with enumerative data. Proper definition of the stylistic feature in question was of key importance to Wimsatt:

Where a certain quality is recognized as a part of style, statistics may give a numerical ratio between the frequency of the quality in one writing and that in another. But the process of making statistics is one of gathering

items under a head, and only according to a definition may the items be gathered. Only by the definition have they any relevance. It is the formulating of the definition, not the counting after that, which is the work of studying style.¹⁴

Wimsatt defines two such qualities of syntax as characteristic of Johnson's style: parallelism and antithesis. Johnson, as borne out by Wimsatt's research, used parallelism (phrasal and sentential) more frequently than his contemporaries Addison and Hazlitt, though he employed certain species of it (pairs of single words, triplets, etc.) less frequently.

In his discussion of Johnson's diction, Wimsatt did no counting himself, but he did identify and define word classes of particular importance to Johnson's style: particular words, specific words, general words, concrete words, abstract words, sensory words, and non-sensory words.¹⁵ Relying on Zilpha Chandler's study, An Analysis of the Stylistic Technique of Addison, Johnson, Hazlitt, and Pater, Wimsatt passes along the relevant statistics: 19.7% of the first thousand words in The Life of Pope are concrete, whereas the figure for a comparable sample from Hazlitt is 32.3% and from Addison 38.3% (Johnson's Prose Style, p. 56, note 20). The enumerative evidence seems to confirm the long-held view that Johnson's style is more contemplative than illustrative--if one can judge this quality on the basis of concrete noun counts. Wimsatt considers the counting of specific and sensory words

counter-productive because, in these categories, the researcher "must proceed, not by statistics, but by examining the function of such words as may be securely called Johnsonian" (Johnson's Prose Style, p. 61). The author's assumption of the organic relationship between style and meaning, explicit throughout the study, exerts its influence here. Indeed, the enumerative evidence takes a secondary position to the discussion of meaning, the statistics being related exclusively in footnotes.

Though Wimsatt's study caused many reverberations, it was more than twenty years before another scholar investigated statistically the prose style of a single author. In 1962, Richard M. Ohmann published his study of George Bernard Shaw, Shaw: The Style and the Man.¹⁶ Ohmann's approach is based on the assumption that "Stylistic preferences reflect cognitive preferences" (Shaw, p. 25); that is, that certain stylistic elements in Shaw's works can be traced directly to peculiarly Shavian ideas. Several of these ideas are identified, and grammatical features are associated with them. Ohmann developed the idea four years earlier in an important essay, "Prolegomena to the Analysis of Prose Style."¹⁷ Here, the writer is perceived as the recipient of sensations and stimuli from the external world. To impose order on these sensations, the reader blocks out many and re-arranges others as the mind and by implication grammatical form impose order. According to Ohmann, the process of writing shapes experience. The writer chooses

from among the various structural possibilities, and the tendency to prefer one form to another affects and is affected by personality and thought:

If the critic is able to isolate and examine the most primitive choices which lie behind a work of prose, they can reveal to him the very roots of a writer's epistemology, the way in which he breaks up for manipulation the refractory surge of sensations which challenges all writers and perceivers. (Prolegomena, p. 9).

This process of arranging Ohmann calls "epistemic" choice because, by means of such a choice, the perceiver/writer comes to know. Style, the theory goes, consists of all the epistemic choices made to produce a piece of writing.

The technique is not strictly statistical--admittedly--but Ohmann does include the results of his frequency counts in an appendix in order to avoid a lapse into impressionism" (Shaw, p. xiii). He says

Few readers will want to pore over the tables, but they are there to give assurance when necessary that the linguistic patterns I discover in Shaw's work are not equally the stock in trade of every writer (Shaw, p. xiii).

Ohmann identifies several habits of mind (constitutional epistemic choices) which manifest themselves in various grammatical structures. Shaw's tendency to lump everything into strict categories (such as "socialism" and "capitalism") is exhibited in the relatively high number of degree words (12 per 1,000 words as compared to 8.3 for Wilde, 4.2 for Chesterton, 6.2 for Yeats, and so on) (Shaw, p. 175). The frequent occurrence of all-or-nothing

determiners" also signals this habit of thought. There is no middle ground for Shaw; everyone is a scoundrel; socialism is the only legitimate system, though nobody is intelligent enough to realize it. These are all-or-nothing determiners. (For all-or-nothing determiners, 16.7 occur per 1,000 words as compared to 11.4 for Russell, 19.2 for Chesterton, and 11.6 for Yeats) (Shaw, p. 174).

In the chapter entitled "The Uses of Discontinuity," Ohmann discusses, among other things, Shaw's love of interruption and surprise effects, which are illustrated in his use of paradox ("Direction shifts"--a broad category involving ironic turns of phrase and contradicting appositions (4.3 per 1,000 words as compared to the 2.2 average for a control group) (Shaw, p. 176). In the next chapter, "The Posture of Opposition," Ohmann explores Shaw's need to contradict, measured by the frequency of negatives (20.6 per 1,000 words as compared to the control group's 13.7 per 1,000 words) (Shaw, p. 177). These counts, which represent Ohmann's informal research, attempt to prove that the grammatical constructions under discussion are distinctively Shavian, not the property of his age. In conjunction with other Shavian turns of mind, Ohmann examines statistically the frequency of proper names, grammatical subjects that are person words, abstract nouns, infinitives, adjectives, "mental causations," that clauses, dependent clauses, degree words, comparative and superlative forms of the adjective and adverb, and so on.

Unlike Wimsatt and Ohmann, Louis Tonko Milic undertook a study, the main focus of which was statistical. Instead of relegating his data to footnotes or appendixes, Milic made them the subject of the work, as the title of his book indicates: A Quantitative Approach to the Style of Jonathan Swift. A year before this important work appeared, Milic had already formulated the ideas upon which it was based. In "Metaphysics and the Criticism of Style," he deplores such typical descriptions of style as the following: crisp, jaunty, sedate, wry, witty, and elegant. These terms are metaphysical--too vague to reveal anything specific about the text. Milic notes that most of the criticism on Swift's style is metaphysical in this way. Critics were fond of contending--often with scant reference to Swift's writings--that his style possesses clarity, propriety, and simplicity (Quantitative Approach, p. 21). Milic proceeds in the next chapter, "The Problem of Style," to outline the basic issues and approaches to the subject, and concludes--not surprisingly--that the statistical approach will "inescapably" be applied with increasing frequency to "certain types of literary work" (Quantitative Approach, p. 72).

To Milic's three central assumptions--"(1) that style reflects personality; (2) that this is an unconscious process; and (3) that in mature writers the process is consistent" (Quantitative Approach, p. 77)--may be added the assumption that style is ultimately measurable. Other less

central presuppositions include the idea that syntax is more "expressive" of unconscious thought than diction and is thus the worthier recipient of scholarly attention, that certain unconsciously determined features of style appear consistently--regardless of the mode of discourse, the audience, or the occasion (Quantitative Approach, p. 79).

In the next two sections--perhaps the most useful ones for the purposes of this study--Milic analyzes carefully Swift's use of seriation (lists) and connection (conjunctions). He explains,

I collected examples of these features of his style and then rendered the procedure objective by an actual count in the work of Swift and various other authors. This mode of proceeding, though it begins with an intuition, ends with concrete data in a form which may be verified (Quantitative Approach, p. 83).

It begins with intuition because, of course, the researcher must decide which feature of style among innumerable others is significant. In 10,000 word selections from their works, Dryden wrote eight series; Defoe, one; Steele, ten; Addison, nine; Goldsmith, nine; Johnson, four; and Swift, twenty-nine (Quantitative Approach, p. 89). The enumeration supports Milic's intuitive impression that seriation is a significant feature of Swift's style. Further, Milic notes that Swift's lists are rarely precisely parallel--most contain slight irregularities of one sort or another (Quantitative Approach, p. 93). He classifies the function of the "continuators" (tags at the ends of

lists such as "and a Thousand other Things," "besides many others needless to mention," and "with many other wild and impossible Chimeras"), many of which serve a satiric function, lumping together the dignified and the base (Quantitative Approach, pp. 97+98). Strangely, no distinct ordering principle can be found in the lists, not alliteration, not rank, not grammatical structure, not formal balance. And this seeming chaos leads Milic to the conclusion--the necessary conclusion, from the point of view of his assumptions--that Swift composed his lists unconsciously, not with a view toward rhetorical effect. An arguable proposition. But after his detailed analysis, accompanied by copious quotations, Milic reaches the conclusion that

The copiousness of imagination which can visualize the reality it conceives of under a legion of aspects in plausible and telling detail, the energy and passion which insist that only through cumulation can its fierce disquiet be expressed--these are the progenitors of his impressive cataloguing of experience (Quantitative Approach, p. 120).

The first assumption about style is confirmed, at least for Milic: style reflects personality. Swift's copious lists reflect his copious mind.

The other stylistic feature which struck Milic's attention was Swift's use of coordinating conjunctions at the beginning of sentences. A count of these initial connectives revealed the presence of the following: coordinating conjunctions (C), subordinating conjunctions (S), and "sentence connectors" (conjunctive adverbs) (SC).

TABLE I
 PERCENTAGE OF INITIAL CONNECTIVES IN 2000-SENTENCE
 SAMPLES OF ADDISON, JOHNSON, MACAULAY, AND SWIFT

Connective	Addison	Johnson	Macaulay	Swift
C	5.5	5.8	7.4	20.2
S	7.1	6.2	4.1	5.4
SC	<u>3.3</u>	<u>1.4</u>	<u>1.5</u>	<u>8.3</u>
Total	15.9	13.4	13.0	33.9

(Quantitative Approach, p. 125)

This count confirms Milic's intuition once more: Swift's use of initial coordinating conjunctions is distinctive. Further, many of the coordinating conjunctions--with additional transitional phrases following--tend to lose their "notional" function. That is, "and" does not signal an addition; "but" does not signal a contradiction; and "for" does not signal a reason. Rather, asserts Milic, Swift seems to use the coordinating conjunction "as a kind of neutral connective, that is a word which shows only that one sentence is connected with another without reference to the nature of the connection" (Quantitative Approach, p. 127). Here is an example from Tale of a Tub:

I shall not enlarge farther upon this Particular. But, another discovery for which he was much renowned, was his famous Universal Pickle. And having remark'd how your Common

Pickle in use among Huswives, was of no farther Benefit than to preserve dead flesh . . . Peter, with great Cost as well as Art had contrived a Pickle . . . (Quantitative Approach, pp. 130-131; italics mine).

The emphasized conjunctions display the non-notional function which Milic describes: neither "but" nor "for" retains the conventional meanings "in contrast to" and "because of this." Through this idiomatic use of the coordinating conjunction, Swift provides his arguments with "a semblance of inevitability" which connects the discourse in a persuasive way, yet does not offer clues to logical sequence (Quantitative Approach, p. 137). At the end of the chapter, Milic reiterates his assertion that Swift uses connectives unconsciously and persistently.

In the next chapter, "Words without Meaning" (reviewed in part above), Milic moves from a consideration of intuitively selected items to a "microscopic" study of style, in which every word is replaced with a two-digit code, representing its word class.¹⁸ The encoded texts were fed into a computer which was instructed to count the occurrences of the individual word classes and combinations of word classes. Milic encoded not only a selection of Swift's works, but also a selection of those of his contemporaries and near contemporaries: Macaulay, Addison, Gibbon, and Johnson. For the purposes of this study, such an analysis is of little use; indeed, one of the sole applications of such a method would be the solution of disputed author problems, for it is extremely difficult to determine the

literary implications of a 6.7% occurrence of the definite article. Similarly, a fingerprint may help to identify a criminal, but the fingerprint itself cannot reveal that person's crimes. Milic does make some attempt, however, to link a high percentage of nouns (the nominal style) with "formality" and "impersonality" and a high percentage of participles (the verbal style) with chatty informalism (Quantitative Approach, pp. 195-200).¹⁹ Obviously, an author's use of the infinitive (one of Milic's word classes), whether the percentage of its occurrence is high or low, can help to fingerprint the work of that author--if the usage is indeed context free--but it cannot inform the student of literary style. Without meaning, the words are easy to count, but impossible to interpret.

Pedagogical Studies

In the following overview, research especially designed to facilitate the teaching of style will receive attention.

One of the champions of objective methodology for the classroom is Rudolph Flesch, author of the widely circulated book Why Johnny Can't Read. In a series of books, Flesch proposed the concept of "readability," that the ease or difficulty with which a piece of writing is read could be calculated mathematically. Flesch states his case in The Art of Readable Writing:

I am sure you realize by now that this book is not dealing with what usually goes by the names of grammar, usage, composition, or rhetoric. On the contrary. If you want to learn how to write, you need exact information--data about the psychological effects of different styles. And handy, usable facts and figures about common types of words, sentences, and paragraphs. And knowledge of the results achieved by various writing techniques. In short, you need a modern scientific rhetoric that you can apply to your own writing. That's what I tried to put into this book.²⁰

In order to calculate the "reading ease score" of a piece of writing, the student counts a 100-word sample and calculates the average sentence length and the number of syllables. The average sentence length is multiplied by 1.015 and added to the number of syllables, which is multiplied by .846. The resultant sum is subtracted from 206.835 yielding a readability score between 0 and 100, 0 being the most difficult and 100 the easiest (Readable Writing, p. 216). The multiplications and subtraction merely convert the score to a hundred-point scale. Armed with this analytic technique, the student is expected to make his writing more readable. In fact, Flesch would have the student abandon stuffy "bookish language" and take up a more personable, readable language. By implication, sentences should be shorter and words smaller. To test the degree of "personability," students count both personal words and personal sentences. Personal words consist of (a) all first-, second-, and third-person pronouns referring to people; (b) words that have natural gender (John, Mary, father, actress, etc.); (c) the group words "people" and

"folks." The student is to count as a personal sentence (a) spoken sentences, signaled by quotation marks; (b) questions, commands, requests, and other sentences directly addressing the reader; (c) exclamations; and (d) grammatically incomplete sentences. To figure this "human interest score," the student performs the following calculation:

$$\begin{array}{r} \text{No. of personal words} \times 3.635 \\ + \text{No. of personal sentences} \times .314 \end{array}$$

The sum, on a scale of 0-100, is the so-called "human interest" score. Apparently, after Flesch had analyzed a number of samples in this way, he determined subjectively that if the score is near zero, the writing is "dull," between 10 and 20 "mildly interesting," between 20 and 40 "interesting," between 40 and 60 "dramatic."²¹ A personable style may be created if the writer uses a relatively high proportion of personal pronouns, questions, exclamations, quotations, and incomplete sentences.

Mr. Flesch does not reveal much of the process whereby he evolved this method, but implicit throughout the work is the notion that writing can and ought to be directly and objectively observed. He implies that his readers should use the enumerative information to their advantage, but does not mention any methods whereby writers can achieve shorter sentences or more personal words. And although one is left with the impression that he prefers the more breezy, less formal style (as evidenced in his

labelling of Human Interest Score levels as "dull" at worst and "dramatic" at best), Flesch avoids endorsing any set style; instead, he emphasizes that different styles reach different audiences. The following chart sorts Reading Ease Scores by genre:

90-100	Comics
80-90	Pulp fiction
70-80	Slick fiction
60-70	Digests
50-60	Quality (<u>New Yorker</u>)
30-50	Academic
0-30	Scientific

Finally, perhaps revealing his own stylistic preferences, Flesch makes an appeal for clear, natural writing:

With all this wonderful opportunity, why do we speak and write the way we do? Why aren't our books and letters and speeches full of racy, colloquial, rhythmical, personal language? (Readable Writing, p. 207).

Flesch's answer: conventional rhetoric, based on assumption rather than observation, has led many to believe that stilted, difficult language is a sign of high social status.

In his next book, A New Way to Better English, Flesch states his case more plainly:

The rules given in this book add up to a simple recipe for better English: be relaxed and informal, stick to the first person singular, go into specific details, quote dialogue, use plenty of anecdotes.²²

While the two previously described scales (Reading Ease and Human Interest) measure the presence of personal pronouns, questions, and--to a degree--dialogue, Flesch requires a new method to determine the degree to which specific detail is present. This time, the scoring is much

easier. The student is to tabulate one point for the occurrence of any of the following in the tested piece of writing (100 words minimum): (1) any word with a capital letter in it; (2) any word that is underlined or italicized; (3) all numbers (unless spelled out); (4) all punctuation marks except commas, hyphens, and abbreviation points; (5) all other symbols, such as #, \$, *, &, and %; and (6) one extra point each for the beginning and ending of a paragraph. As measured by this scale, the presence of proper nouns, foreign words, book titles, specific numbers like 98.6, and the symbols which often accompany them supposedly indicate the presence of concrete detail. An ample number of periods, question marks, exclamation points, and dashes--on the other hand--reveals a tendency to use short sentences in a colloquial fashion, perhaps in dialogue. Thus, Flesch may offer a useful teaching method to the instructor who harps at puzzled students about the necessity of concrete detail.²³

Mr. Flesch's procedures are attractive because they require little linguistic or mathematical expertise and they yield interpretable results. One begins to approach them with some reservation, however, when it becomes clear that Mr. Flesch is the unreserved ambassador of a dapper, wholly colloquial style which clearly has some virtue and utility, but which is not always appropriate--especially in writing intended for a serious audience.

Although he includes stylistic advice and specific

pedagogical procedures in only one chapter and a lengthy appendix, it is not a mistake to place Walker Gibson's study Tough, Sweet, and Stuffy in this section on pedagogical approaches.²⁴ His approach is rhetorical. It contributes to a general understanding of the motives and purposes of different writing styles; thus, through a lengthy discussion of the rhetorical situation, the stylistic advice--though briefly stated--receives the proper grounding.

To Gibson, style is "self-dramatization"; that is, the writer, deprived of gesture, of smile, of grimace, must "dramatize" his character in his style (Gibson, p. x). Three of these styles--named in the title--are described in the book. "Touch talk" is the language, mainly, of the hard-bitten narrators of prose fiction. Advertisers croon "Sweet talk," and bureaucrats and committee members prefer "Stuffy talk." Gibson examines rigorously the psychology of each style--who the speaker is, who the audience is, who the audience is supposed to be. Not surprisingly (given the nature of his work), Gibson's three styles correspond to three points of view, roughly equal to the three personal pronouns and to the three major elements of the rhetorical situation: writer, subject, and audience:

The Tough Talker . . . is a man dramatized as centrally concerned with himself--his style is I-talk. The Sweet Talker goes out of his way to be nice to us--his style is you-talk. The Stuffy Talker expresses no concern either for himself or his reader--his style is it-talk (Gibson, p. x).

Based on these fundamental insights plus the more meticulous analysis presented in the body of the work, Gibson's "Style Machine" appears in a lengthy appendix. The "machine" is printed in Table II in full.

Essentially, the Tough Talker--a Hemingway narrator, or Augie March--uses short, gutsy words in short sentences (see #1, #2, and #11 above). Because the Tough Talker has little regard for the audience, he avoids subordinate clauses which explain logical relationships (see #12 and #15 in Table II) and employs the definite article more frequently (see #14), thus assuming gruffly an improbable familiarity in the reader. (Such is the case in the following sentence: "The river is wide, green." Which river? Where? The definite article implies that the reader knows, or should know these things.) The Sweet Talker, the advertiser, uses longer, more playful words (See #1 and #2 in Table II) such as "undeniably delectable" and a few more subordinate clauses. Sentences become shorter, modifiers occur more frequently (for example, "it's delightfully different; Bang gibes you more cleaning power"), and the reader is addressed directly ("You may dislike magazine ads, but you haven't read this one"). The language tends toward informality with a number of contractions and fragments (See #15). Stuffy Talkers, unlike the Tough and Sweet Talkers, seek to avoid responsibility for their decisions: Stuffy Talk is the talk of reports and studies. It avoids the use of personal nouns and pronouns (see #3

TABLE II
THE STYLE MACHINE CRITERIA FOR MEASURING STYLE

	Tough	Sweet	Stuffy
1. Monosyllables	over 70%	61-70%	60% or less
2. Words of 3 syllables and more	under 10%	10-19%	20% or more
3. 1st and 2nd person pronoun	1 <u>I</u> or <u>we</u> per 100 wds	2 you per 100 wds	no 1st or 2nd person pronoun
4. Subjects: neuters vs. people	1/2 or more people	1/2 or more people	2/3 or more neuters
5. Finite verbs	over 10%	over 10%	under 10%
6. To be forms as finite verbs	over 1/3 of verbs	under 1/4	under 1/4
7. Passives	less than 1 in 20 verbs	none	more than 1 in 4 verbs
8. True adjectives	under 10%	over 10%	over 8%
9. Adjectives modified	fewer than 1 per 100 wds	1 or more	fewer than 1
10. Noun adjuncts	under 2%	2% or more	45 or more
11. Average length of clauses	10 words or less	10 words or less	more than 10 words
12. Subordinate clauses, proportion of total wds	1/4 or less	1/3 or less	over 50%
13. "Embedded" words	less than 1/2 S/V combinations	less than half	more than twice
14. "The"	8% or more	under 6%	6-7%
15. Contractions, fragments	1 or more per 100 wds	2 or more	none
16. Parentheses & other punctuation	none	2 or more per 100 wds	none

(Gibson, p. 136)

and #4), relying instead on the passive voice (see #7). Thus, a Stuffy Talker never writes, "I have determined that" but "It has been determined. . . ." In addition, Stuffy Talkers use many subordinated clauses, and the subject and verb are often separated by "embedded" words which may make comprehension temporarily difficult (see #12 and #13).

What does this research mean for the writer? First of all, Gibson avoids endorsing any three of the styles which he has discussed, all of which have their dangers, but notes that Stuffiness is the greatest fault in modern prose (Gibson, p. 107). Stuffy Talk is to be avoided. To make this task easier, Gibson compiles a list of ten specific recommendations for avoiding this unbalanced language and improving prose style. These include suggestions to keep two-thirds of one's vocabulary monosyllabic, to make the subject a person where possible (not concepts or neuter pronouns), to avoid using the passive voice excessively, to reduce interruptions between subject and verb, to lighten the tone with question marks, italics, and other marks of punctuation (excluding the comma and semicolon), and so on (Gibson, pp. 108-109). Gibson wants students to learn a way "of becoming a person worth listening to" (Gibson, p. 110), not just a way to conform mechanically to a style machine.

Edward P. J. Corbett offers several methods of enumerative analysis in his composition text Classical Rhetoric for

the Modern Student.²⁵ After mapping the possible boundaries of a stylistic study (which includes diction, sentence length, type, variety, figures of speech, and paragraphing), Corbett concludes that

There are a number of incalculable features of style about which we might never be able to secure general agreement, but if we are to develop any system for analyzing prose style we must start with those features that are objectively observable (Classical Rhetoric, p. 440).

Corbett offers up his techniques (in the form of four student reports) with little ceremony and with less theoretical discussion, including only occasionally a comment on the observed practices of professional writers. The first report deals with sentence and paragraph lengths: the students are to compare the computations derived from their own writing to those derived from the work of professional authors (this comparison is central to all four reports). Students in Corbett's classes found that, although their average sentence lengths were roughly equivalent to those of the model, they wrote far fewer extremely short or extremely long sentences, the implication being that professionals cultivate some variety in sentence structure. Similarly, the students found a discrepancy in paragraph length: the professionals were on the whole much shorter than their own.

The second report deals with sentence types (simple, compound, complex, and compound-complex) and their sequence in a text. Students are asked to count the occurrences of each type and figure its percentage in relation to the whole.

Corbett mentions that modern sentences tend to be extended, not by the addition of main clauses, but by the addition of subordinate structures, a finding which directly opposes the conclusions reached by L. A. Sherman years before.²⁶ In this report, students are also asked to chart the sequence of these grammatical types but receive instructions for putting this knowledge to use.

Report number three duplicates, to a large degree, Christensen's study of sentence openers mentioned in the introductory paragraphs of this study (Classical Rhetoric, p. 456). The student compares the percentages of different kinds of sentence openers (subject, expletive, coordinating conjunction, adverb, conjunctive phrase, prepositional phrase, verbal phrase, adjective phrase, absolute phrase, adverb clause, or inverted word order). Corbett's students generally confirmed the findings of Christensen: most professionally written sentences begin with the subject.

The fourth report, concerning diction, measures--among other things--the lengths and percentages of "substantive words" (nouns, pronouns, verbs, verbals, adjectives, and adverbs) in contrast to structure words, the percentage of concrete nouns, of linking verbs, passive verbs, and adjectives. Corbett does not report on his students' findings.

Accompanied by little theoretical discussion, Corbett's four and analytical schemes rest squarely on the shoulders of classical rhetoric--though the enumeration may appear modern. In a separate article, "The Theory and Practice of

Imitation in Classical Rhetoric," Corbett identifies the two key stages in the imitative practices of the ancients: analysis and genesis.²⁷ Clearly, the techniques discussed above allow greater precision in the analysis stage, providing--so Corbett hopes--a basis for successful student imitation.

Corbett develops his methodology further in an article entitled "A Method of Analyzing Prose Style with a Demonstration Analysis of Swift's A Modest Proposal."²⁸ Here enumerative analysis is seen as the first step in a larger study of rhetoric:

The "why" of any stylistic feature can be answered only in relation to something else--the subject-matter or the occasion or the genre or the author's purpose or the nature of the audience or the ethos of the writer ("Method," p. 338).

The addition of rhetorical considerations is a necessary, and for Corbett, a quite natural one. As a matter of fact, without discovering the reasons for objectively observed stylistic features, students engage in a largely meaningless drudgery, mapping out a region with no aim to visit it or explore its contours. (Rhetorical considerations, because of their extreme importance for understanding style, will be treated separately in the next chapter.) Corbett demonstrates his technique on the most admired, most analyzed essay in the language. Discussing the common appearance of periodic sentences in the piece, Corbett notes that

This tendency toward periodic structure is evidence not only of a deliberate written style but of a habit of the persona that suits Swift's rhetorical purpose . . . to create a character who will, as it were, "sneak up" on the reader ("Method," p. 344).

Corbett's "sneaking up" is, of course, the comedian's punch line, Mark Twain's "snapper," which in Swift's hands becomes a weapon of irony wielded at the last minute. Students, armed with such knowledge, are theoretically much more aware of the uses of periodic sentence structure; thus they are better able to employ it in their writing than those who have merely tabulated its presence.

The methods discussed in the following paragraphs are not pedagogical in the sense that students are expected to use them with an aim toward imitation; scholars use them to support stylistic theories and the predictions rising from them. In "Sentence Openers," an article cited in the introduction of this study, Francis Christensen explored sentence openers in expository prose and found that, contrary to the traditional pronouncements, most sentences begin with a subject, a few with adverbial phrases and coordinating conjunctions, and a tiny minority with adjective phrases such as verbals.²⁹ More important to the theory of generative rhetoric, however, is his enumerative study of free modifiers in non-professional, semi-professional, and professional writing.³⁰ Here, Christensen--attacking the concept of T-unit as a measure of mature style--argues for short main clauses and a generous use of free modifiers, "modifiers

not of words but of constructions from which they are set off by junctures of punctuation" ("Problem," p. 370). Free modifiers are indicative of a mature style. Christensen further argues that professionals have learned to write "cumulative" sentences, sentences with terminal free modifiers. His enumerative study reveals that while non-professional, presumably inept writers usually place their free modifiers before the main clause, the professionals add them predominately at the end ("Problem," p. 577). This cumulative sentence is, of course, the backbone of the Generative Rhetoric program. Students are taught to "add" to an idea by "generating" more minute descriptions appended to the main clause. For the most part, Christensen endorses the noun and adjective clusters as well as absolute phrases as candidates for terminal free modifiers, these preferences supposedly based on his enumerative studies.

In spite of his wide acceptance, Christensen is not without his detractors. Sabina Thorne Johnson in "Some Tentative Strictures on Generative Rhetoric" calls Christensen's sampling technique into question.³¹ Johnson conducts her own study of authors like Forster, Isherwood, Cather, Orwell, Baldwin, and others and finds that none of them has nearly as many words in free modification as Christensen's "best" writer, Halberstam. Johnson concludes:

If we are to measure the degree of skill in a writer by the percentage of words he has in free modification, then we should rate Cather, Fitzgerald, Forster, Isherwood, Baldwin, Auden and Orwell less skillful than Halberstam, or

assume that my passages, chosen at random, are atypical (Johnson, p. 364).

The danger is clear: anyone who basis a stylistic theory on "objective" data must be prepared to wrangle about sampling and tabulation technique.

Mary P. Hiatt's study of parallelism in modern prose style, Artful Balance: The Parallel Structures of Style, is based on enumerative data, like Christensen's work and Sabina Johnson's objection to it, though Hiatt's methods are much more rigorous. This computer-assisted study is long and technical, but the pedagogical implications receive ample attention:

Since some kind of parallelism occurs in approximately 50 percent of our written sentences, its presence certainly forces us to examine carefully the range of its structures and the effectiveness of its use. And with such a high frequency of occurrence, it can be considered a major measure of style. In teaching students some facts about parallelism, therefore, we are teaching them some facts about style and about clear and effective writing. We only need to be sure of the facts.³²

The study reveals that faulty parallelism is by no means uncommon nor--as Hiatt implies--abhorrent. Another finding implies that rhetorical parallelism (schemes, as well as "strict" parallelism) is less likely to be found in imaginative prose, the reading fare of most students, than in informative prose. Perhaps, postulates Hiatt, students fail at formal, rhetorical writing because their models contain little formal or rhetorical parallelism (Hiatt, p. 119). She is able to draw this conclusion because she divides her

samples into genre categories under two main headings as shown in Table III.

TABLE III
CLASSIFICATION OF PARALLELISM SAMPLES

Informative	Imaginative
Press: Reportage	Fiction: General
Press: Editorial	Fiction: Mystery and Detective
Press: Reviews	Fiction: Science
Religion	Fiction: Adventure and Western
Skills and Hobbies	Fiction: Romance
Popular Lore	Humor
Belles Lettres	
Miscellaneous	
Learned and Scientific Writings	

Although such conceptual rigor would be difficult to carry out in the classroom, as would indeed the sheer enormity of the counting, a reduced or simplified schema could benefit student writers.

Developmental Studies

Research on syntactic "maturity" or "fluency" remains to be considered.³³ In Grammatical Structures Written at Three Grade Levels, Kellogg W. Hunt attempts to discover

objectively discernable traits which differentiate the writing of children in different developmental stages.³⁴ With such information, the writing instructor would be better able to develop exercises designed to accelerate the maturation of the students' syntax.

The approach is for the most part disciplined and objective (the format is that of a scientific report rather than that of an impassioned essay). Hunt explains the concept of syntactic maturity:

In this study the word "maturity" is intended to designate nothing more than the observed characteristics of writers in an older grade. It has nothing to do with whether older students write "better" in any general stylistic sense (Hunt, p. 5).

Just as the attributive works sought merely to fingerprint and identify the writings of a certain author, Hunt wishes to fingerprint the characteristics of students at certain writing levels--without reference to qualitative judgment.

Hunt, finding other enumerative techniques unsatisfactory, develops the concept of T-unit (terminal unit). Previous studies noted the increase of subordination and main clause length with maturation, but sentence length research seemed inadequate to measure these factors (one sentence might contain two or more main clauses). The concept of the T-unit was developed to avoid this shortcoming; it consists of a main clause with all its subordinate clauses. Thus, quirks of punctuation, common in the writing of younger students, and compound sentences do not throw

off the results (as they had perhaps in Sherman's work).

According to Hunt,

The length of such a unit might turn out to be a good index of maturity. It might turn out to be an even better index than the two subsidiary factors [subordination, main clause length] because of the fact that an individual who was high in subordination index but low in clause length (or the reverse) would have those opposite tendencies moderated by this combining index (Hunt, p. 20).

Theoretically, T-unit length should increase with age because both main clause and subordinate clause length increases.³⁵ Not surprisingly, Hunt's data bear out this prediction.

In 1966, another study (completed by Donald Bateman and Frank Zidonis) The Effect of a Study of Transformational Grammar on the Writing of Ninth and Tenth Graders further refined the concept of syntactic maturity.³⁶ Convinced that maturer writers use more transformations, Bateman and Zidonis devised a simple indicator: the Structural Complexity Score (SCS). The SCS is calculated as follows: each transformation (one of forty-six in the grammar specially designed for classroom use) is scored with one point, a point being scored for each sentence also. This number is then divided by the number of sentences. Bateman and Zidonis wanted to compare pre- and post-experimental scores to determine whether or not a study of transformational grammar would have an effect on student writing. Interestingly, a study of transformational grammar did not improve the SCS more than traditional methods (Bateman and Zidonis,

p. 35). The program did significantly affect two other measures, marginally important for this investigation: the Proportion of Well-formed Sentences (PWS) (that is, the number of well-formed sentences, as intuitively judged, divided by the total number of sentences) and the Error Change Score (ECS) (that is, the number of errors in a pre-experimental sample subtracted from that of a post-experimental sample) (Bateman and Zidonis, pp. 13, 14). Instruction in transformational grammar was supposed to yield a higher SCS, a higher PWS, and a negative ECS, but only the latter two measures were affected as predicted. Important, though, are the analytical methods, all three of which (with the possible exception of the intuitively judged PWS) are purely enumerative and easy to calculate, though students would need instruction in transformational grammar to be able to figure the SCS.

In 1969, John C. Mellon developed a much more complete measure of syntactic fluency based on the work of both Hunt and Bateman and Zidonis. For Mellon, as for Bateman and Zidonis, the increasing length of the T-unit is not nearly as important as the increase in the occurrence of transformations, but he recognizes the obvious value of using the T-unit rather than the sentence or the total number of words for the base statistic.³⁷ (That is, it is much more useful to know the number of relative clauses per T-unit than it is to know the number of relative clauses per sentence or per 100 words.) Mellon recalculates Hunt's results using

the T-unit as the base and determines that certain transformations (the nominal and relative transformations) are likely to increase with maturation while others remain constant. For the purposes of his experiment, Mellon developed twelve more factors of syntactic fluency:

1. mean T-unit length (in words)
 2. subordination-coordination ratio
 3. nominal clauses per 100 T-units
 4. nominal phrases per 100 T-units
 5. relative clauses per 100 T-units
 6. relative phrases per 100 T-units (reduction of relative clauses: prepositional phrases, participial phrases, and so on.)
 7. relative words per 100 T-units
 8. embedded kernel sentences per 100 T-units
 9. cluster frequency (the percentage of T-units in which there are two or more modifiers attached to a single noun)
 10. mean cluster size
 11. embedding frequency (the percentage of T-units which contain one or more embedded sentences)
 12. mean maximum depth level (refers to transformations within transformations--there are three possible levels, 3 representing an embedded sentence within an embedded sentence within an embedded sentence)
- (Mellon, pp. 46, 48-49)

Mellon set out to determine, among other things, which transformations, if any, could be encouraged by transformational sentence combining exercises. The data were subjected to complex statistical scrutiny, revealing, though ambiguously, that sentence-combining drills lead to increased syntactic fluency. Like the Bateman and Zidonis methodology, Mellon's requires some degree of expertise with transformational grammar, a demand unlikely to be met by most students; however, the insight suggesting that the T-unit might best serve as a base unit is valuable.

In summary, the statistical and enumerative procedures presented in this chapter come from many sources and were intended to serve diverse functions. The underlying concept is, however, the same: certain features of style may be objectively observed either to explain the tendencies of individual writers or a succession of them or to attribute works of unknown origin to a specific author, or to assist students in stylistic analysis, or finally, to determine what constitutes syntactic maturity in writing.

Of course, not even the most resourceful instructor could accept all or even most of the techniques presented here for classroom use, but some certainly can be adapted. And those which cannot may offer insights into what can or cannot be accomplished successfully in the classroom. It is difficult to judge, however, which procedures might be employed successfully because for the most part very few of the researchers have offered their theoretical assumptions. In the next chapter, therefore, this study will turn to a consideration of the underlying concepts involved and attempt to point up some of the shortcomings of the methods discussed in this chapter.

ENDNOTES

¹T. C. Mendenhall, "A Mechanical Solution to a Literary Problem," The Popular Science Monthly, 77 (1901), 97-105. For a more complete summary of early work in this field see Richard W. Bailey's essay "Statistics and Style: A Historical Survey," in Statistics and Style, ed. Lubomir Dolezel and Richard W. Bailey, Number Six in the series Mathematical Linguistics and Automatic Language Processing (New York: American Elsevier Publishing Company, 1969), pp. 217-236.

²G. Undy Yule, The Statistical Study of Literary Vocabulary (Cambridge, England: University Press, 1944), p. 238.

³G. Undy Yule, "On Sentence-Length as a Statistical Characteristic of Style in Prose, with Application to Two Cases of Disputed Authorship," Biometrika, 30 (1938), 363-390.

⁴G. Herdan, The Advanced Theory of Language as Choice and Chance, vol. 4 of Kommunikation und Kybernetik in Einzeldarstellungen (New York: Springer-Verlag, Inc., 1966).

⁵For a detailed discussion of genre and subject matter and their effects on style, see Chapter II, Section 2, above.

⁶Alvar Ellegard, A Statistical Method for Determining Authorship: The Junius Letters, 1769-1772, no. 13 in Gothenburg Studies in English (Gothenburg: University of Gothenburg, 1962), p. 15.

⁷Claude S. Brinegar, "Mark Twain and the Quintus Curtius Snodgrass Letters: A Statistical Test of Authorship," Journal of the American Statistical Association, 58 (1963), 85-06.

⁸Frederick Mosteller and David L. Wallace, "Inference in an Authorship Problem," Journal of the American Statistical Association, 53 (1963), 275-309.

⁹Louis Tanko Milic, A Quantitative Approach to the Style of Jonathan Swift, (The Hague: Mouton & Company,

1967). Hereafter, reference to this book will be cited parenthetically under the short title Quantitative Approach.

¹⁰L. A. Sherman, "Some Observations upon the Sentence-Length in English Prose," University of Nebraska Studies, 1 (1889), 349. See also Sherman's follow-up article, "On Certain Facts and Principles in the Development of Form in Literature," University of Nebraska Studies, (1892), 337-366; and his book Analytics of Literature: A Manual for the Objective Study of English Prose and Poetry (Boston, 1893).

¹¹Ibid., p. 353.

¹²Ibid., p. 367.

¹³Robert E. Moritz, "On the Variation and Functional Relation of Certain Sentence-Constants in Standard Literature," University of Nebraska Studies, 3 (1894), 237.

¹⁴W. K. Wimsatt, The Prose Style of Samuel Johnson, vol. 94 in Yale Studies in English (New Haven, Connecticut: Yale University Press, 1941), p. 24. Hereafter, reference to this book will be cited parenthetically under the author's name.

¹⁵Ibid., see Chapter III, "Diction," pp. 52-62.

¹⁶Richard M. Ohmann, Shaw: The Style and the Man (Middletown, Connecticut: Wesleyan University Press, 1962). Hereafter, reference to this book will be cited parenthetically under the short title Shaw.

¹⁷Richard M. Ohmann, "Prolegomena to the Analysis of Prose Style," in Style in Prose Fiction: English Institute Essays, ed. Harold C. Martin (New York: Columbia University Press, 1959), pp. 1-24. Hereafter, reference to this article will be cited parenthetically under the short title "Prolegomena."

¹⁸These word classes are divided into two categories: parts of speech (noun, main finite verb, descriptive adjective, descriptive adverb, infinitive, participle, gerund, and foreign words) and function words (pronouns, auxiliary, determiner, post-verb particle, intensifier, function adverb, coordinator, subordinator, relative, interrogative, correlative, preposition, pattern-marker, interjection, numeral, and sentence connector). See the inset flap at the end of Quantitative Approach.

¹⁹For a more complete description of the characteristics of nominal and verbal styles, see Rulon Wells' article "Nominal and Verbal Style," in Linguistics and Literary

Style, ed. Donald C. Freeman (New York: Holt, Rinehart and Winston, Inc., 1970), pp. 297-306.

²⁰Rudolph Flesch, The Art of Readable Writing (New York: Harper & Row, Publishers, 1949), p. 9. Hereafter, reference to this book will be cited parenthetically under the short title Readable Writing.

²¹See the inside cover at the beginning of Readable Writing for Flesch's graphic representation of these two scales.

²²Rudolph Flesch, A New Way to Better English (New York: Harper & Brothers, Publishers, 1958), p. 134.

²³The danger exists, of course, that students might begin to value the symbols for their own sakes, not for what they indicate.

²⁴Walker Gibson, Tough, Sweet, and Stuffy: An Essay on Modern American Styles (Bloomington: Indiana University Press, 1966). Hereafter, reference to this book will be cited parenthetically under the author's name.

²⁵Edward P. J. Corbett, Classical Rhetoric for the Modern Student (New York: Oxford University Press, 1971). Hereafter, reference to this book will be cited parenthetically under the short title Classical Rhetoric.

²⁶See Sherman's first article, note 19 above.

²⁷Edward P. J. Corbett, "The Theory and Practice of Imitation in Classical Rhetoric," College Composition and Communication, 22 (1971), 243-250.

²⁸Edward P. J. Corbett, "A Method of Analyzing Prose Style with a Demonstration Analysis of Swift's A Modest Proposal," in The Writing Teacher's Sourcebook, ed. Gary Tate and Edward P. J. Corbett (New York: Oxford University Press, 1981), p. 245. Hereafter, reference to this article will be cited parenthetically under the short title "Method."

²⁹"Sentence Openers," p. 46. See note 8 in Chapter 1.

³⁰Francis Christensen, "The Problem of Defining a Mature Style," English Journal, 57 (1968), 572-579. Hereafter, reference to this article will be cited parenthetically under the short title "Problem."

³¹Sabina Thorne Johnson, "Some Tentative Strictures on Generative Rhetoric," in Contemporary Rhetoric: A Conceptual Background with Readings, ed. W. Ross Winterowd (New York: Harcourt Brace Jovanovich, Inc., 1975), pp.

352-364. Hereafter, reference to this essay will be parenthetically cited under the author's name.

³²Mary P. Hiatt, Artful Balance: the Parallel Structures of Style, Vol. 2 in the New Humanistic Research Series (New York: Teachers College Press, 1975), p. 121. Hereafter, reference to this book will be parenthetically cited under the author's name.

³³Because these studies are not intended for student use, they are difficult to adapt for the classroom and are presented here because they offer some useful insights into statistical procedure.

³⁴Kellogg W. Hunt, Grammatical Structures Written at Three Grade Levels: NCTE Research Report No. 3 (Champaign, Illinois: National Council of Teachers of English, 1965). Hereafter, reference to this report will be cited parenthetically under the author's name.

³⁵Christensen objects strongly to the long main clause. In "Problem," he states, "The very hallmark of jargon is the long noun phrase--the long noun phrase as subject and the long noun phrase as complement, the two coupled by a minimal verb" (p. 575).

³⁶Donald R. Bateman and Frank J. Zidonis, The Effect of a Study of Transformational Grammar on the Writing of Ninth and Tenth Graders: NCTE Research Report No. 6 (Champaign, Illinois: National Council of Teachers of English, 1966). Hereafter, reference to this report will be parenthetically cited under the authors' names.

³⁷John C. Mellon, Transformational Sentence-Combining: A Method for Enhancing the Development of Syntactic Fluency in English Composition, NCTE Research Report No. 10 (Champaign, Illinois: National Council of Teachers of English, 1969), p. 18. Hereafter, reference to this report will be parenthetically cited under the author's name. Of course, Mellon is interested not so much in fixing "syntactic maturity" statistically as examining the usefulness of the insights which transformational-generative grammar provide. In this sense, Mellon and his fellow researchers (Bateman and Zidonis) probe far beyond the "surface structure" which enumerative analysis reveals into the "deep structure." The obvious difference between "John is eager to please" and "John is easy to please" cannot be divined by enumeration, which after all is not a system for grammatical analysis. Thus, any complete composition program must address fundamental questions of structure with additional instruction from grammars specifically designed to explain such deeper problems.

CHAPTER III

THEORY OF ENUMERATIVE ANALYSIS

For the most part, the researchers whose methods were presented in the previous chapter left the theoretical foundations of their statistical and enumerative analyses undiscovered. But anyone who has considered even briefly the possible theoretical problems posed by such analyses--which certainly involve the unending discussion of the nature of style itself--will find this omission odd. To fill this lack, this chapter will focus on four theoretical cruxes of the enumerative techniques: the problems of objectivity, form and content, rhetorical considerations, and application.

Objectivity

Those who assert that an enumerative analysis of style is objective do so overconfidently. Literary judgments made on the basis of objectively collected data are not necessarily more accurate or "objective" than intuitive judgments; in addition, any such analysis must begin with intuition.

While attributive studies attempt only quantitative judgments (sample X of unknown origin has 5.5% N per T-unit; sample A by Johnson has 9.0%, sample B by Swift has 2.0%,

sample C by Hazlitt has 6.0%. Therefore, the chances are 6.8 to 1 that Hazlitt is the author of sample X), literary analyses usually cannot forbear qualitative ones (because the percentage of concrete nouns is 34.5%, we must conclude that Shakespeare was a man of vigorous, wide-ranging intellect, able to absorb the many details of his environment). This latter conclusion, based supposedly on the statistic, is subjective, just as subjective in fact as the "metaphysical" pronouncements about style condemned by Louis T. Milic in "Metaphysics in the Criticism of Style":

Absolutes like pure and perfect describe nothing at all, but are merely assertions by the critic that the writer is without flaw. Such terms as muscular, nervous, sinewy drawing as they do from various parts of the writer's anatomy, reveal the critic's desire to move from the style to the man, or to use the man to describe the style ("Metaphysics," p. 125).

Given such a statement, it seems ironic that Milic, in his study of Swift's seriation, could assert that

The copiousness of imagination which can visualize the reality it conceives under a legion of aspects, the fertility of invention which can realize these aspects in plausible and telling detail, the energy and passion which insist that only through cumulation can its fierce disquiet be expressed--these are the progenitors of the impressive cataloguing of experience. In its redundancy his cataloguing derives from the urge to control meaning (Quantitative Approach, pp. 120-121).

How are these conclusions justified? By the statistics?

How can numbers reveal "energy and passion" or ferret out a

"fierce disquiet"? These conclusions are subjective and in

Milic's use "metaphysical." It would certainly be misleading

though, to imply that Milic's study of Swift is filled with such questionable statements; indeed, he produces several useful insights, revealing in one section the satiric function of some Swiftian lists (Quantitative Approach, p. 97). This insight is not based on knowledge gained from bare statistics, however, but rather from a close examination of the lists themselves.¹

Milic's doubtful statement arises, however, from a belief that style necessarily reflects the personality of the author. Ironically, this is the basis for the entire work:

The major concern is . . . the unconscious expression of the writer's personality in his writing. That this unconscious reflection of the writer's personality in his work is consistently diffused through all his writings is the major assumption of this study (Quantitative Approach, pp. 76-77).

Thus, Milic feels himself compelled to comment not merely on the function of certain stylistic features in Swift's writings but on Swift's personality and mental make up as well. He is determined to make such an analysis and is willing to go out of the way to get it. Jonathan Swift has been dead for some two hundred years now, and though he were resuscitated for the purpose, it is doubtful that others could confirm Milic's claim that he urges to "control meaning" or that his "fierce disquiet" forced him to compile lists. It may be good sport to perform psychoanalysis on the prose of the dead, but it is not fair sport.

In an important article, "What is Stylistics and Why

Are They Saying Such Terrible Things About It?" Stanley E. Fish recognizes the general corruptibility of statistical data:

While it is the program of stylistics to replace the subjectivity of literary studies with objective techniques of description and interpretation, its practitioners ignore what is objectively true--that meaning is not the property of a timeless formalism, but something acquired in the context of an activity--and therefore they are finally more subjective than the critics they would replace. For an open impressionism, they substitute the covert impressionism of anchorless statistics and self-referring categories. In the name of responsible procedures, they offer a methodized irresponsibility, and, as a result, they produce interpretations which are either circular--mechanical reshufflings of the data--or arbitrary readings of the data that are unconstrained by anything in their machinery.²

If stylists depend wholly on statistics--which have no lexical meaning, no context--they are in danger of exceeding the impressionism of those they attack, who at least must refer to a few bits of text. Of course, most literary critics do not rely on statistics exclusively--as Milic at times tends to do--relegating them instead to footnotes and appendixes, as Wimsatt and Ohmann do.

But the problem runs deeper. Ohmann also believes in equating style with personality. In fact, his widely circulated article, now the basis of the most current theory of style (style is choice)--"Prolagomena to the Analysis of Prose Style"--is based on this assumption. According to Ohmann, a writer chooses--consciously or unconsciously--among the countless variations of expression which his

language offers him:

I have been outlining a theory of style which describes choices that I have called epistemic. These choices are important, for they are the critic's key to a writer's mode of experience. They show what sort of place the world is for him, what parts of it are significant or trivial. They show how he thinks, how he comes to know, how he imposes order on the ephemeral pandemonium of experience ("Prolegomena," p. 19).

This claim is extravagant. In the same way that a physician cannot determine a patient's entire medical history, but merely his present physical condition, from the sound of a cough, neither can a critic divine the whole man from his syntax, though he may comment upon the effect of syntax on the immediate thought.

That this error can appear in analyses of style intended for the classroom is borne out by Edward P. J. Corbett. In "A Method of Analyzing Prose Style," the author comments upon Swift's remarkably long sentences: "In A Modest Proposal we are listening to a man who is so filled with his subject, so careful about qualifying his statements and computations, so infatuated with the sound of his own words, that he rambles on at inordinate length" ("Methods," p. 341). Is the Swift persona "infatuated" with his own rhetoric? Obviously, the sentence length calculations cannot bear out such an assertion. Whether or not the persona qualifies his statement is, on the other hand, quite verifiable.

The moment in which a critic attempts to enter the mind

of his subject, to predict a certain emotion or intention on the basis of a statistic--or even on the basis of an explicit statement--the project can go awry. Human beings do not possess the ability to read minds. Such attempts to do so are inferences from the effect of the writing on the reader. Corbett makes similarly questionable statements about Swift's use of terminal absolute phrases: "These trailing-off phrases create the effect of a thought suddenly remembered and desperately thrown in" ("Method," p. 342). Is the Swift persona "desperate"? Perhaps, perhaps not. Every statement here is arguable--as it should be--but any judgment which requires the critic to refer to the writer's mental state is dangerous and highly debatable. Corbett's essay is by no means dominated by such pronouncements and, of those discussed above, only one was traceable to a statistic. The fallacious assumption that style is personality is bad enough when the analysis refers to specific texts; it is much worse when the only proof is a string of numbers. Such abuses must be avoided.³ That is not to say that such statements about a writer's personality cannot be valuable--they can. But they require a leap of faith, and that leap is entirely too easy to make. Thus, though the occurrence of any stylistic feature can be objectively determined, the interpretation of its function or relevance to an author's personality is just as subjective as such a statement without the statistics.

Statistical analysis is subjective in another way.

Unless the researcher counts all that may be counted, intuition is required at the beginning of the process to determine what is significant enough to enumerate. None of the scholars whose work is represented here denies that fact, and a few state it plainly. Milic recounts the process by which he developed his procedure:

I began by reading the works of Swift with careful attention to such peculiarities as I might observe to be present in Swift and absent from the work of his contemporaries. . . . I collected examples of these features of his style and then rendered the procedure objective by an actual count in the work of Swift and various other authors. This mode of proceeding, though it begins with an intuition, ends with the concrete data in a form which may be verified (Quantitative Approach, pp. 82-83).

Let there be no misunderstanding: the tabulation of the data can be verified, not the conclusions drawn from the data.

Wimsatt, careful in theoretical matters, displays more insight when he states,

When a critic is conscious of quality X in a writing, no accumulation of statistics will increase his consciousness of it. But if he simply announces that the writing has X, he may be challenged. If he says that it has X because he has found X in fifteen examples of fifteen hundred words each, he is less likely to be challenged; if he adds that the average is a hundred occurrences in each example, even less likely. This however, is not proof, but something more like persuasion, for logically the whole matter rests on the definition with which he began, and statistical details are taken, no less than a blanket statement, on faith (Wimsatt, p. 24).

There is nothing magical about statistics. As a matter of fact, if Wimsatt is correct, the value of any enumerative technique is likely to be determined by its design and the

definition of categories, both of which are intuitively determined. If one is to count the presence of "sensory" words, an exact definition of what is meant by "sensory" is much more important than the mere fact of tabulation. In other words, since the process begins with an intuition, those who design the methods must be especially sensitive to what they deem "significant" features of style.⁴ For, after all, the initial intuitive judgment is the pivotal one.

The intuitive nature of statistical analysis implies that critics with different tastes will devise different methods; thus, it is just as impossible as before to arrive at the set of instructions for the student that follows: In expository prose, the percentage of prepositions is not to exceed X%; concrete nouns should total X%, and so on. Such judgments can be made, of course, but no amount of counting will make them objective. It seems more likely that statistical analysis would be advantageous as a springboard to imitation, as Corbett suggests. Qualitative judgment is still involved, for the instructor must choose models worthy of emulation, but the misconception of a universal good style is avoided.

For the time being, literary analysis remains a somewhat subjective undertaking. This fact should not be upsetting though. The fact that a statistical analysis of style begins, and often ends, with an intuitive evaluation should be heartening. Most instructors are not mystics; they believe that an intuitive judgment is supportable.

Thus, the challenge of thinking, tracing, and feeling carefully confronts every instructor who wishes to attempt some enumerative technique.

Form Versus Content

Statistical compilations have no logical meaning. No reader can be informed, moved, amazed, amused, or convinced by a list of numbers. But whether style can be separated from meaning, form from content, is a highly controversial question which has received vigorous critical attention in this century.

Benedetto Croce, the Italian philosopher, developed in his book Aesthetic the organic or Crocean theory.⁵ Simply stated, this theory disallows any attempt to separate form from content. J. Middleton Murry has formulated this view perhaps more succinctly than any of the Croceans: "Style is not an insoluble quality of writing; it is the writing itself."⁶ Style evaporates. To the Croceans, every shift in wording or syntax means a change--no matter how slight--in meaning. Of course, this view precludes the study of style altogether. One cannot learn how to state a certain idea; one must learn what meaning is most effective, the implication being that exercises in logic would be more appropriate than exercises in rhetoric. Milic recognizes the radical effect this view has on the teaching of writing:

The consequences of the disappearance of style which results is that discussion of the student's writing must consist almost exclusively

of the philosophy, so to speak. The emphasis which this theory forces on us is the dominance of the subject. For if there is no form, we cannot discuss, much less improve, the student's means of expression. . . . The monistic view of style, therefore, cannot be allowed to infect the teaching of our subject, for it vitiates all the available pedagogical resources of rhetoric.⁷

Milic makes a valid point: what is the writing instructor to teach if he is allowed to discuss only the meaning, not the mode of expression? The inconvenience of adopting a theory, however, does not prove its falsehood. The question remains: Is style meaning? Admittedly, it is not within the scope of this study to answer this question, but two insights may make the matter easier to cope with.

First, there may be a mode of expression in which meaning differs from the form: irony and satire. A sentence from the much-analyzed Swift essay illustrates the point: "I shall now therefore humbly propose my own thoughts which I hope will not be liable to the least objection." Anyone familiar with Swift's proposal cannot accept that statement at face value. Does Swift expect no objection? Obviously not. How is it then that this sentence and ultimately the entire tract come to mean something quite different from the literal, semantic content? It is difficult to say. In speech, though, listeners are often guided to the ultimate "meaning" of a statement by the speaker's tone of voice. The statement, "My, don't you look lovely today," stated in a normal tone of voice is a compliment: however, the moment that eyebrows rise or the intonation becomes exaggerated,

"My don't you look LOVELY today," it may become an insult. It is possible that different sentence structures and different wordings simulate the visual and audio signals on which listeners depend to interpret a message. Here are cases, perhaps isolated ones, which reveal a separation of form and content: the statements mean the opposite of what they say. Of course, if one includes "emotion," "intent," or "emphasis" in one's definition of meaning, the issue becomes clouded. Thus, definition of the term meaning is crucial to the discussion.

Second, a more precise definition of the word meaning can simplify matters. I. A. Richards' four kinds of meaning, for example, ease the discussion of form and content considerably. According to Richards in Practical Criticism,

It is plain that most human utterances and nearly all articulate speech can be profitably regarded from four points of view. Four aspects [of meaning] can be easily distinguished. Let us call them Sense, Feeling, Tone, and Intention."⁸

Sense corresponds to lexical meaning, feeling to the author's attitude toward the material, tone to the attitude toward the listener, intention to what the author is attempting to accomplish. If the organicists maintain that there is no separation between sense, in Richards' usage, and style, their notions are suspect. Language is not all logic. As Richards admits, the author can sometimes "purpose no more than to state his thoughts" (Richards, p. 176). Such "honest" writing, however, does not by any means make up the

bulk of written discourse. If the organicist maintains that there is no separation between form (style) and meaning (feeling, tone, intention, and sense), there is little to argue about. Unfortunately, the Croceans were not specific enough in their definition of "meaning," but one thing remains clear: statistical analysis of style empties writing of sense, feeling, tone, and intent. Thus, the relevant question is not whether form can be separated from content--that has been done--but, should it be done? The next section will consider this important question.

Rhetorical Considerations

Any statistical analysis of style which makes no reference to content must in the end be meaningless. Stylisticians who attempt this, according to Stanley Fish, want to "specify the meaning of the moves in the game without taking into account the game itself" (Fish, p. 133). Such total separation is, however, extremely difficult. Even those researchers like Mary Hiatt, who is interested in one syntactical form, parallelism, make reference to the general kinds of writing from which their samples are taken.⁹ Such a classification necessarily makes reference to a text's meaning. Rudolph Flesch accomplishes much the same thing when he lists the sources of his samples.¹⁰ Such a procedure brings meaning back to the statistics because genres or modes and aims of discourse provide general indications of probable intent, feeling, tone, and sense, and these

qualities are closely associated with stylistic variations. The fact that a piece of writing contains a large percentage of passive verbs, for example, does not mean anything of itself, unless the statistic is attached to a meaningful category, for example, to the fact that passive verbs occur frequently in academic and scientific writing, where the information, not the person relating it, is important, but not in comic books, where the action of various characters is most important.

With this sort of information, one can begin to understand the probable motives of the author and the probable reaction of the audience.¹¹ In "The Rhetorical Stance," Wayne C. Booth provides a framework for such an analysis, though his comments do not refer specifically to style:

The common ingredient that I find in all of the writing I admire . . . is something that I shall reluctantly call the rhetorical stance, a stance which depends on discovering and maintaining in any writing situation a proper balance among the three elements that are at work in any communicative effort: the available arguments about the subject itself, the interests and peculiarities of the audience, and the voice, the implied character, or the speaker. I should like to suggest that it is this balance, this rhetorical stance, difficult as it is to describe,¹² that is our main goal as teachers of rhetoric.

According to Booth, aberrations occur when the balance is disturbed. He identifies three imbalances, the pedant's stance, the advertiser's stance, and the entertainer's stance, which place undue emphasis on one of the three elements of the rhetorical stance. The pedant cares nothing

for the audience, but focuses on the subject, while the advertiser is unduly concerned with the audience, or the persuasion of the audience, to the point that the message is often deemphasized; the entertainer emphasizes his role to the exclusion of subject and audience. Booth expends very little space discussing these imbalanced stances, but Walker Gibson, in Tough, Sweet, and Stuffy, duplicates many of Booth's original contentions and elaborates upon them.

Each of Gibson's adjectives corresponds to one of Booth's unbalanced stances: the pedant's stance is Stuffy Talk; the advertiser's stance is Sweet Talk; and the entertainer's stance is roughly equivalent to Tough Talk. Gibson explores thoroughly the ethos of these three voices and discovers in the process the tone, feeling, and intent which are transmitted through certain stylistic features.¹³ Here follows a description of the Sweet Talker's style and ethos:

A Sweet Talker is not at all a hard man who has been around [as the Tough Talker]. He addresses me directly ("you"), and when he says "you" he doesn't mean just anybody, he means me. He is not a passionate or self-centered man. . . . On the contrary, he goes out of his way to be nice to me. . . . He may use the rhetorical devices of informal speech (contractions, fragments, eccentric punctuation) to secure his intimacy with me (Gibson, p. 83).

Such an analysis is useful to the student because it gives insight into the effect of a certain stylistic structure. It does no good to tell a student that (since short sentences dominate in persuasive writing) he must use short sentences when he writes persuasive prose. The student

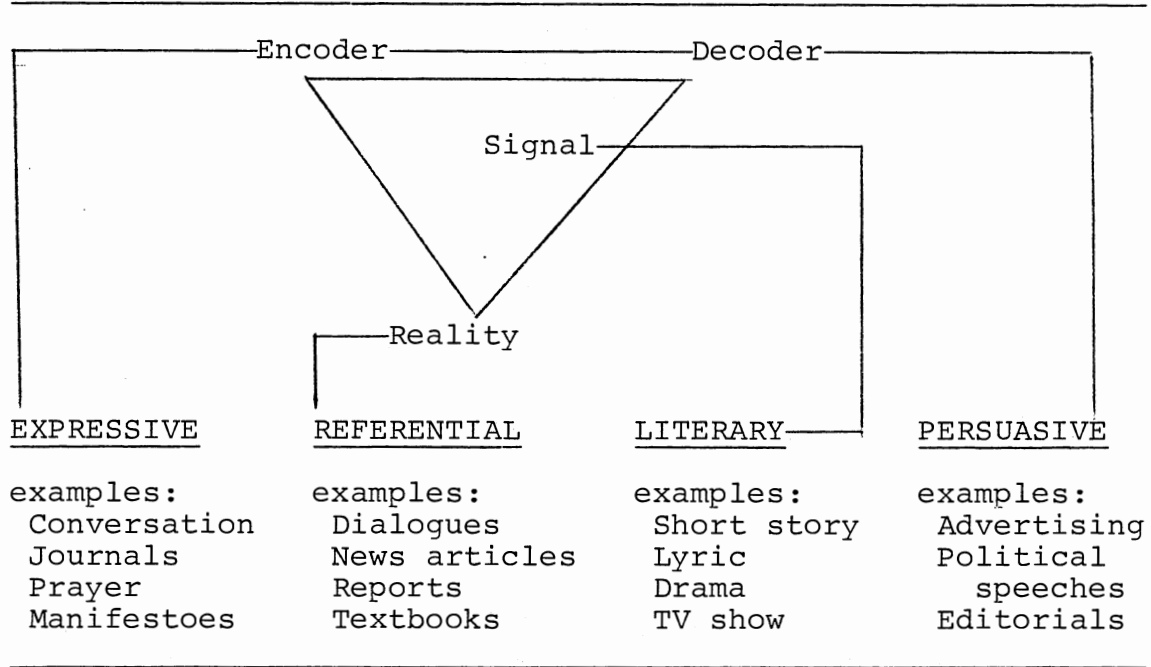
needs to know the reason behind the short sentence, that it often produces a colloquial effect, at least in informal writing. Without this analysis, the student may produce short sentences which have the opposite effect: "One determination has been made. This determination is highly significant. Smoking is not advisable." These sentences are short, but they are neither colloquial nor persuasive.

But the problem with the Booth-Gibson model, for all its insights into the functionings of style, is that it offers only three categories for study: tough entertainer talk, sweet advertiser talk, and stuffy pedant talk. Further, both Booth and Gibson see these three styles as exaggerations to be avoided. James L. Kinneavy presents an interesting variation in A Theory of Discourse.¹⁴ Here, the three elements (writer, audience, and subject) are present again under different headings: encoder, decoder, reality--to which a fourth heading, signal, is added. Table IV elaborates the system.

For Kinneavy, emphasis on the decoder (audience) does not produce Sweet Talk, but persuasive writing. In turn, Gibson's "I-language" is not necessarily tough. In Kinneavy's paradigm it is expressive. Stuffy Talk is referential, in that it refers to reality (subject matter). Thus Kinneavy does not focus on the harmful exaggeration of any one element (as Booth and Gibson do) but considers normal, healthy emphasis instead. Each type of discourse has a different nature. According to Kinneavy,

Each aim of discourse has its own logic, its own kind of references, its own communication framework, its own patterns of organization, and its own stylistic norms. Sometimes these logics and stylistic principles even contradict each other. Overlaps certainly occur but the ultimate conflation and confusion of any of the aims of discourse with any other is pedagogically disastrous (Kinneavy, p. 98).

TABLE IV
THE BASIC PURPOSES OF COMPOSITION



Each mode of discourse has "its own stylistic norms."
This is worthy of emphasis. Thus, the need for categorizing

becomes clear: different rhetorical orientations produce different styles, and the student cannot hope to produce effective persuasion if the models which he has been analyzing are expressive. In the same way, any analysis which mixes different types of discourse in a grab-bag of samplings may find the results highly ambiguous, if not misleading.

Kinneavy's discourse paradigm has several advantages for use in a statistical study. For one thing, other categories such as the traditional modes of discourse (narration, description, exposition, and persuasion) do not offer the insights into motive that Kinneavy's model does. Flesch's use of publication types could perhaps be useful with a more complete analysis of the ultimate goals of each type, but such a taxonomy is dangerous because articles written for wholly different reasons may be sandwiched between the covers of one magazine or book.

While the absence of meaningful categories makes interpretation of the results difficult, what one chooses to count within the confines of a particular category is of pivotal importance. What can be counted? Almost anything, but there are two broad, important divisions: morphological class and semantic class.

For the purposes of scientific study, morphological (grammatical) structures provide a clear cut advantage over such classes as "concrete" nouns or "sensory" adjectives. That is, morphological structures are identifiable by form

rather than content, enabling the researcher to make quick and unambiguous tabulations. Attributive studies, which need not arrive at an interpretation of a work, employ such categories. G. Undy Yule and Louis Milic both count word classes such as nouns, determiners, prepositions, and so on. Researchers interested in the effect of maturation on writing--not in the stylistic qualities of a text--count syntactic structures (coordination subordination ratios, occurrence of certain sentence patterns, and so on) with recent work focusing on transformations. Francis Christensen enumerates phrasal structures such as the absolute phrase, noun and adjective clusters, and verbal and adverbial phrases. Much earlier, T. C. Mendenhall calculated word-length distributions, while Sherman tabulated sentence lengths.

If the purpose of such counting is to fingerprint a piece of writing by its characteristic structures, such formal categories (morphological classes) are justifiable. If, however, one seeks to develop a meaningful literary or rhetorical interpretation of the results as Christensen, Milic, Corbett, and Sherman wish to do, their value is less defensible. For example, it is possible to determine the probable rhetorical significance of the terminal absolute modifier. That such a structure is present in a piece of writing may easily be determined, but how does that structure complement sense, feeling, tone, and intent? One is forced to judge the rhetorical effect of a certain

structure by gauging one's subjective reaction to it. This process is hardly scientific. However, one may, indeed one must, speculate as to the connection between form and meaning, but interpretation becomes much easier when the two are never separated.

Semantic classes make reference to the meaning contained in a certain structure, not to the structure itself. These could include emotive, concrete, metaphorical or sensory words, as well as asides, qualifying statements, ironic phrases, and so on. Obviously, these classes are much more difficult to identify than the morphological ones. For this reason, perhaps, those researchers who feel compelled to use semantic classes do so with some regret. Richard Ohmann, for example, comments on the design of his Shaw study:

[An] obstacle to my analysis is the difficulty of preserving a distinction between grammatical structure and meaning. I have in mind sins of this sort: treating classes like abstract and concrete words, evaluate [evaluative] words, causal words, and so on, as stylistic categories. It would clearly be preferable if the only categories used in the actual description of style were formally defined, grammatical--not semantic (Shaw, p. xiv).

Ohmann does use these "sinful" classes and is quick to express his discomfort at the lack of precedent: "the authority on which I associate semantic content with grammatical classes is mainly my own" (Shaw, p. xv). The counting of semantic classes need not be haphazard and subjective: they can be "formally defined" as Wimsatt notes.¹⁵ Such counting, though it is sanctioned wholeheartedly by none of the

researchers discussed so far and practiced by only a few, forces the student to pay close attention to the message of the text, not just to its outward form (sentence length, the number of syllables, the appearance of prepositions). Counting morphological classes often becomes tedious and literally mechanical. When students search for ironic turns of phrase, however, they must take in the full meaning (sense, feeling, tone, and intent) of the text. The advantage of this operation is clear.

Finally, one can maintain the connection between meaning and style by constantly referring to the texts, not just to the data generated from them. Any stylistic analysis--even enumerative analysis--seeks to make the text comprehensible. It is ironic then that the quest for this understanding could ever lead away from the text, away from the thing that is primary. The fact that it can is one of the dangers of enumerative analysis, and it is a trap into which Milic (in "Words without Meaning") and Sherman, among others, fall. Walker Gibson, W. K. Wimsatt, and Rudolph Flesch are quite careful, on the other hand, to discuss individual passages in depth with a view to explaining and clarifying the results of the more general statistical analyses. Their books are filled with excerpts, sentences, and bits of sentences. In fact, Rudolph Flesch's The Art of Readable Writing contains, almost page for page, one to three block quotations illustrating the pompous or the cluttered, the relaxed and the colloquial. Through the illustrations, the

reader acquires the sense, the meaning of the numbers. An analysis which omits such considerations is of little value and should be avoided.

In conclusion, though "meaningless" statistics (the percentage occurrence of the preposition "by" for example) are helpful in fingerprinting a piece of writing for possible identification purposes, they are of dubious value in a discussion of rhetorical or literary matters. Here, the link between form and content cannot remain severed: samples must be categorized; the items counted should where possible refer to semantic class; and the text itself should remain at the center of interpretation.¹⁶

Application

If an instructor has devised a good method which makes meaningful interpretation possible, one problem remains: analysis alone cannot enable the student to produce a good imitation. In this sense, the conclusions of a statistical analysis resemble the familiar dictums of English teachers: "If you wish to imitate informal writing sample O, use short sentences." Such commands have the disadvantage that they leave the students to find their own way to shorten sentences. Obviously, the instruction in analysis must be accompanied by a discussion of methods whereby the recommended stylistic features may be produced. Rudolph Flesch offers such a procedure in The Art of Readable Writing: wordy prepositions, conjunctions, and connectives may be

simply replaced, according to the following list:

Too Heavy Prepositions and Conjunctions

along the lines of: like
 as to: about (or leave out)
 for the purpose of: for
 for the reason that: since, because
 from the point of view of: for
 inasmuch as: since, because
 in favor of: for, to
 in order to: to
 in accordance with: by, under
 (Readable Writing, p. 131-134)

Beside these substitutions, only a few of which are quoted here, Flesch presents the student with lists of constructions conducive to word economy and relaxed expression. Such specific advice is necessary.

If the student is expected to write fewer prepositional phrases in order to successfully imitate a certain model, he should be shown steps to transform these prepositional phrases into other forms. In the following sentence, the prepositional phrase can be eliminated through the use of a possessive: "The failure of the legislation was the fault of the senator." Revised, the sentence looks like this: "The failure of the legislation was the senator's fault." Prepositions may also be eliminated through the use of a verb: "The senator failed to gain passage for the legislation." There is little doubt that the step-by-step assistance which transformational sentence-combining offers is in large part responsible for its apparent success.

In conclusion, enumerative analysis of style provides the base of knowledge which enables the student to imitate.

However, it provides no specific methods for achieving the styles which it may recommend. Such procedures should be provided, perhaps within the context of a writing program.

ENDNOTES

¹See Chapter III, Section 2.

²Stanley E. Fish, "What is Stylistics and Why are They Saying Such Terrible Things About it?" in Approaches to Poetics: Selected Papers from the English Institute, ed. Seymour Chatman (New York: Columbia University Press, 1973), p. 140. Hereafter, reference to this article will be parenthetically cited under the author's name.

³Buffon's famous statement "style c'est le homme meme" expresses the fallacy. Though there is perhaps some truth in this assertion, it is a distorted truth, placing primary interest on only one leg of the rhetorical stance. See note 59 above.

⁴The idea that style is a deviation from the normal mode of expression (see W. Ross Winterowd's introduction to Style in Contemporary Rhetoric, pp. 256-257) is closely linked with the concept of style as the man. That is, both look upon style as something distinctive and unique. This assertion, however, leads to the regrettable conclusion that ordinary writing has no style at all. Since this theory seems to be grounded on the principle that style is a noticeable feature, it is reasonable to assume that repetition, not deviation is the operating principle.

⁵Benedetto Croce, Aesthetic. tr. D. Ainslie (New York: Noonday Press, 1966).

⁶J. Middleton Murry, The Problem of Style (London, 1936), p. 77.

⁷Louis Tonko Milic, "Theories of Style and Their Implications for the Teaching of Composition," College Composition and Communication, 16 (1965), 69, 126.

⁸I. A. Richards, Practical Criticism: A Study of Literary Judgment (New York: Harcourt, Brace and Company, 1929), p. 175.

⁹See Mary P. Hiatt's classification scheme on p. 41 above.

¹⁰See Rudolph Flesch's scheme on p. 29 above.

¹¹Analyzing the author's motives or his ethos is not the same thing as inferring the writer's personality from his writing. Ethos refers to the temporary character of the writer in one particular piece, and is quite correctly derived from the reader's reaction.

¹²Wayne C. Booth, "The Rhetorical Stance," in The Writing Teacher's Sourcebook, p. 111; originally published in College Composition and Communication, 14 (1967), 139-145.

¹³See "The Style Machine," p. 33 above.

¹⁴James L. Kinneavy, "The Basic Aims of Discourse," College Composition and Communication, 20 (1969), 302.

¹⁵Only a few of the examples are included in this table.

¹⁶See p. 16 above.

¹⁷John Gage--in "Philosophies of Style and Their Implications for Composition," College English, 41 (1980), 615-622--suggests that there may be a conflict between linguistic analysis of a text ("how many main verbs are there?") and a rhetorical analysis ("how does this style effect the audience?"). For Gage, the conflict arises from two distinct ideologies, one which sees style as a shaper of thought (the monistic view) and one which sees style as a tool to shape thought (the dualistic view). Gage writes, "As a linguistic concept, style may be defined in terms of grammatical norms and operations of variance. But such a definition can provide little help in settling typical stylistic questions raised from a rhetorical perspective, such as whether all stylistic features thus defined are chosen and, indeed, whether they will have an effect on an audience" *pp. 615-616). Fortunately, this gulf may be spanned if, according to Gage, instructors know which view they are drawing upon and keep it separate from the other view, for "what we actually teach is a combination of both theories" (p. 621).

CHAPTER IV
ENUMERATIVE TECHNIQUES FOR THE
CLASSROOM

I base the classroom techniques offered in this chapter on the preceding investigation of the available methods and the theories behind them. At least six implications arise from this previous discussion:

1. Because researchers must choose intuitively what stylistic features to count, different researchers analyzing various works will devise different methods.
2. Instructors should categorize writing samples intended for analysis by aim of discourse or by other available schemes.
3. Where possible, semantic classes should receive more attention than morphological classes.
4. After an instructor has selected an item to analyze, he should define it precisely in order to render tabulation easier.
5. One of the few legitimate applications of enumerative analysis--besides attribution studies--is to promote imitation.
6. Analysis alone cannot affect imitation. Specific procedures designed to produce the desired stylistic feature are also necessary.

These six conclusions apply to the three basic techniques which I shall present in this chapter: close imitation, Attribution, and definition.

Close Imitation

The practice of imitation is an ancient one. And although to examine the theory of imitation lies beyond the scope of this study, the words of Quintillian lay out immediately the two problems with which instructors will have to grapple:

The nicest judgment is required in the examination of everything connected with this department of study [imitation]. First we must consider what it is that we should set ourselves to imitate in the authors chosen.¹

As has been implied previously, selection of the model will depend on a number of things, not the least of which is mode of discourse. In most freshman composition programs, expository (Kinneavy's referential) prose receives the most emphasis. If the fallacy that style reflects personality (explored in the first section of the second chapter) is to be avoided, instructors should refrain from choosing as models the prose from the great masters of the English language. Such writing exudes personality (novels narrated in the first person, for example, exude personality because of the subject matter--the author's feelings, attitudes). According to Kinneavy's scheme, such writing is "expressive" because the emphasis is on the encoder. The work of acclaimed authors is not the ideal model for a freshman writer because any imitation should not only attempt to copy style, but subject matter, ethos, and audience as well. Students might find the cool disdain of George Orwell quite distant from their own personalities; his

subjects--the decline of the English language, for example, as in "Politics and the English Language"--removed from their concerns; his audience, too proud and mighty. How many students can approach the profound moral indignations of a Jonathan Swift frothing at that "odious vermin," man? Not many, one may suppose. For this reason, it is much better to provide as models everyday prose pieces such as those found in newspaper and magazine articles, essays, or non-fiction books. To imitate an article about cutbacks in education, students must not assume any extraordinary stance; they are not forced to don the mein of an accomplished expert, of a moral philosopher, or of a blazing literary star.

Quintillian's second consideration has been dealt with at some length in Chapter III: the instructor--or in certain cases, the student--must intuitively identify those components of the text (morphological or semantic) which are most significant. If the investigator finds no obviously important stylistic features immediately, he may employ Corbett's taxonomy of style (see Classical Rhetoric, pp. 450-458), but any consideration of style usually involves an investigation of lexicon (individual words), syntax (the structure of the individual sentence) and what W. Ross Winterowd calls "transitions" ("coherent relationships beyond the sentence").² Of course, a feature must occur at least two or three times to qualify as a distinguishable trait of style. A sample author may compose an

effective terminal participial phrase, but if such a construction appears nowhere else in a 1,500 word sample, it can hardly be considered notable enough to single out for imitation.

Once these steps are accomplished, the instructor can assign a writing exercise based on the subject matter, audience, and ethos of the model. For example, had the students analyzed a magazine article such as "DNA Coils: Link to Aging" from Science Digest, they should write on a similar topic ("Hydrogen Combustion as an Alternate Energy Source," for example), with the same "scientific" ethos, aiming at the same audience, which is interested in science but has no special background in advanced technology. Only when these components match is a close imitation of style likely to yield satisfactory results, for if a student were to incorporate the stylistic features of the scientific article (a high proportion of connectives, especially correlative conjunctions, short sentences with only one main clause, and a predominance of specific, concrete language) into a paragraph describing the ethereal beauty of Emily Dickenson's "I heard a Fly buzz--when I died," the production would be marked by a stark incongruity, in all probability.

A variation on this procedure is conceivable: the instructor could assign a writing exercise which mirrors the subject, audience, and ethos of a model which he will introduce only after the students have completed their

original assignment. For example, students would be asked to write about "Hydrogen Combustion as an Alternate Energy Source," they would receive a copy of "DNA Coils: Link to Aging" and carry out an analysis of the same, finally comparing their approach to that of the model.

Since differences in subject matter or approach can invalidate a comparison of stylistic techniques, even when the topics are generally related, the instructor may wish to employ a different procedure which eliminates such disparities. Here follows such a procedure. The instructor issues a set of instructions such as the following to students who are unaware of what will follow:

You are a reporter for The New York Times and consider yourself witty and literate--you like to turn a nice phrase. Most of your readers are well-educated and want to be informed in a pleasant way. Your writing assignment? Report on the filming of the last M*A*S*H episode. Your audience probably knows the show and its characters, but you will play it safe and get all the facts in. Although the break-up is obviously an emotional one for the cast, you report the events with a detached, slightly ironic style. Here are your notes:

Paragraph #1--M*A*S*H on for 11 years; on CBS; anti-war comedy; very successful; it ends today (Jan 14); many (hundreds) newspeople on set to watch.

Paragraph #2--[History of the Show] in first 10 yrs. 99 Emmy nominations, 14 Emmy awards; show ends not because bad ratings; good ratings; ('82 - '83 season): 60 Minutes (#1), Dallas (#2), M*A*S*H (#3); is 20th-Century Fox show.

Paragraph #3--[About Alan Alda, star] Alda quotations: "I've been here 11

years, one-fourth of my life."
"It's the right time to stop before we decline."

.... and so on

These instructions and paragraph outlines are derived from an article in The New York Times, and they are advantageous in that they eliminate invention and content problems which cloud a true comparison between the model, which is to be introduced next, and the student writing based on these instructions. The instructor should phrase these notes as tersely as possible in order to prevent the student from picking up stylistic hints from the original.

After the students have written their "articles," the actual New York Times piece is introduced. The instructor should encourage students to compare their writing with the model and then with each other's productions. What is the difference in average sentence length? in the occurrence of terminal participial phrases? of ironic words? of subordinated clauses? Of course, no two attempts will resemble one another exactly, and the instructor should stress that fact: the resources of the language are vast, and every person has a style.

After the students have completed their counts, enabling them to compare the percentage occurrence (per T-unit) of the particular features in their "articles" to that in the original piece, the instructor assigns a new topic closely related to the previous one, "Hill Street Blues: A Quality Series?" perhaps, for which paragraph

outlines similar to the ones provided above may be included. Students would attempt to match the occurrence of the desired stylistic features (the terminal participial phrase, for example) in this writing, allowing themselves reasonable divergences (no more than 10% more or less) from the enumerative profile of the professionally written M*A*S*H article.

More traditional methods can also prove successful. The student receives a copy of the model and is asked to discuss writer ethos, audience, subject in relation to style. With student help, the instructor identifies the most prominent stylistic features (for example, absolute terminal phrases, mildly ironic asides, concrete nouns, action verbs, and so on). These are defined and tabulated. The student may then compose an essay on a similar subject, attempting to duplicate not only the percentage occurrence of a particular feature but the rhetorical situation as well.

Besides analyzing and imitating single articles, students may be required to imitate the style of a particular magazine such as Time or Newsweek provided that the source, with its many contributors, has a recognizable style. Again, the samples must belong to the same category of discourse to insure that they are similar enough in intent and tone to render comparison possible. As before, the class identifies the salient stylistic features, defines them, relates them to the writer's purpose and the audience's need, tabulates them, and attempts to reproduce them in a practice assignment.

From personal experience in tabulating, I have found that for significant results, the student writing as well as the models should be at least five to six hundred words long, and if possible one should analyze the entire piece in order to eliminate disparities in style called forth by the different demands of introduction, assertion, support, and conclusion. If one breaks down a sample into one-hundred-word increments and subjects them separately to enumerative analysis, he discovers that only after five or six counts can one be certain that a seemingly significant stylistic feature actually occurs with enough frequency to be significant.

Armed with the three methods which I have presented in this section, students can attempt close imitations of sample models, but it is of vital importance that they imitate the tone, intent, feeling, and sense of the model, not merely its outward grammatical structures.

Attribution

Attributive exercises require students to analyze the style of three to five models and then to determine on the basis of that analysis the author of an unidentified sample (composed, of course, by one of the three to five authors). The attributive studies summarized in the first section of this study, because they aim at scientific proof and methodological integrity, emphasize morphological features as the best indicators of authorship. For the classroom,

however, such rigor is unnecessary. In fact, as frequently repeated, rhetorical considerations must accompany such stylistic analysis and figure prominently in any justification of a particular attribution.

Students could begin, as Ellegard did, by searching for "plus" words and structures which seem to them indicative of one sample but not of another. Students then compare the presence of these features in the known and unknown samples.

The following three paragraphs about Samuel Johnson will help to illustrate the technique:

Sample #1

He was born at four o'clock in the afternoon of Wednesday, September 18, 1709, in the town of Lichfield, Staffordshire, which then had a population of about three thousand. He was the first of two sons of a bookseller, Michael Johnson, and his wife, Sarah, who were both much older than parents usually are at the birth of their first child. Michael--a self-made man, large-framed and gaunt, conscientious, prone to melancholy--was by now fifty-two. Sarah, who prided herself that her family connections were socially superior to those of her husband, was forty. They had been married a little more than three years (June 19, 1706). Samuel's birth took place in the bedroom above Michael's shop, in their house, which still stands across from St. Mary's Church and overlooks the Market Square.³

Sample #3

Samuel Johnson was born at Lichfield in Staffordshire on the 18th of September, N.S., 1709; and his initiation into the Christian Church was not delayed; for his baptism is recorded, in the register of St. Mary's parish in that city, to have been performed on the day of his birth. His father is there stiled Gentleman, a circumstance of which an ignorant panegyrist has praised him for not being proud; when the truth

is, that the appellation of Gentleman, though now lost in the indiscriminate assumption of Esquire, was commonly taken by those who could not boast of gentility. His father was Michael Johnson, a native of Derbyshire, of obscure extraction, who settled in Lichfield as a bookseller and stationer. His mother was Sarah Ford, descended of an ancient race of substantial yeomanry in Warwickshire. They were well advanced in years when they married, and never had more than two children, both sons; Samuel, their first born, who lived to be the illustrious character whose various excellence I am to endeavour to record, and Nathanael, who died in his twenty-fifth year.⁴

Sample #3

Wednesday, 18 September 1709: the Market Square in Lichfield was quiet, for this was not a trading day. The occasional rattle of a cart, or the talk and laughter of a knot of citizens passing the time of day at a corner, would sound clear across the square, while at regular intervals the long swell of melody from the great cathedral bells came washing over the rooftops. All these sounds penetrated to the ears of Sarah Johnson, the bookseller's wife, as she lay in her bedroom in the handsome three-storied house that dominated the north-eastern end of the square. To Sarah, the day must have seemed a long one. At forty, she was giving birth to her first child, and the labour was prolonged and difficult. Her husband, Michael, had engaged the services of George Hector, the best man-midwife in the neighbourhood; beyond that there was nothing anyone could do but wait. Finally, at four o'clock in the afternoon, Hector was able to take up the strangely inert yet living body of the child and say encouragingly, 'Here is a brave boy!'⁵

Sample #1 comes from W. Jackson Bate's biography Samuel Johnson; sample #2 from James Boswell's Life of Johnson; sample #3 from Samuel Johnson, a modern biography by the English poet John Wain. Since the excerpts describe the same event, though from different points of view, there is little danger that differing aims of discourse will make

comparison difficult. Such a sampling technique also makes a comparison of rhetorical approaches possible.

Students begin by cataloging the distinctive stylistic features of each sample. Bate's style is simple: most sentences begin with the subject; most words do not exceed three syllables. It contains a number of interruptions set off by commas: these are mainly appositives and relative phrases. Boswell's style is more involved. His frequent use of the semicolon is distinctive, rendering most sentences quite long. His vocabulary tends to be more Latinate than Bate's (circumstance, indiscriminate, appellation), and he does not shy away from the first person pronoun. Wain's style, as is appropriate for a poet, is more expressive. His diction draws on the auditory senses (rattle, swell, quiet, laughter, penetrated) and is metaphorical (knot of citizens, swell of melody, bells came washing over the rooftops). Most of his sentences contain at least one subordinated clause, making the sentences relatively long.

Along with this sort of analysis, the instructor should provide pertinent facts about the careers and characters of the three authors. With these facts, the student may begin to make a rhetorical analysis: Bate wants to relay the facts objectively, yet elegantly; Boswell is a snob who is interested not only in Johnson, but in his own relationship to Johnson; Wain wants to write a passionate biography; he wants to discover the human qualities of

these great people.

After class discussion, the instructor introduces the fourth sample, written by one of the three authors, yet unidentified. (This sample, as well as the first three should be considerably longer for the purposes of reliable analysis; they should be at least 500 words long.)

Sample #4

Perhaps we may see some connection between those years of fierce, intermittent work and vibrant emotional response and the illness which overwhelmed him in 1766, the year after the Shakespeare was published. This illness was evidently something in the nature of a nervous collapse. Johnson's will deserted him; he lay on his bed for weeks on end, and gloomy forebodings gnawed incessantly at his mind.⁶

Students should search for the "plus" characteristics of any of the three initial samples in this fourth one then compare their counts. There is some use of semicolon, but it does not appear with the same frequency as in sample #3. The simple sentence structure is similar to that of Bate, yet the first sentence contains a long subordinated clause and the last sentence, three main clauses. The emotive diction points to Wain: fierce, vibrant, gloomy forebodings gnawed.

The instructor may, of course, introduce any degree of statistical sophistication deemed necessary or feasible, but, again, statistical data are not the only means of proof. In fact, students could write profitably about their conclusions, justifying them with a discussion of their statistical and rhetorical findings. (The writer of

sample #4 is interested in Johnson's emotional state, not his intellect or his relationship to the author; thus, Wain, the poet, is the likely author of the piece.) Such an essay topic incidentally provides the advantage that students have a clear purpose, audience, and subject matter.

Definition

This type of exercise is considerably simpler than close imitation. The object of definition exercises is to describe certain subjective terms by the stylistic manifestations associated with them. Walker Gibson did just that: he explained the words "tough," "sweet," and "stuffy" in terms of countable stylistic features. Several adjectives commonly employed to characterize style could be similarly analyzed: emotive - objective, formal - colloquial, awkward, indignant, metaphoric, concrete, and so on.

As always, intuition must guide the selection of features to be counted. The best way to begin this process is first to outline what the word colloquial, for example, suggests: such writing exhibits many of the characteristics of speech, with thrown-in asides, simple sentence structure, and informal vocabulary. Using appropriate sample texts, students and instructor could then identify at least three features (morphological or semantic) which they intuitively associate with the term: short average sentence length, a preponderance of monosyllabic

words, frequent use of you, of hyperbole, and so on.

For a discussion of the meaning of colloquial, Studs Terkel's oral history books can serve as illustrative models since they are transcriptions of actual interviews. In Working, for example, a receptionist explains her duties:

Oh sure, you have to lie for other people. That's another thing: having to make up stories for them if they don't want to talk to someone on the telephone. At first I'd feel embarrassed and I'd feel they knew I was lying. There was a sense of emptiness. There'd be a silence, and I'd feel guilty. At first I tried to think of a euphemism for "He's not here." It really bothered me. Then I got tired of doing it, so I just say, "He's not here." You're not looking at the person, you're talking to him over the instrument.⁷ So after a while it doesn't really matter.⁷

Several observations might be made about this excerpt. First of all, the average T-unit length (as calculated from a larger sample than printed here: 669 words, 69 T-units) is very short (9.69 words/T-unit). In the larger sample, the word you appeared frequently (45 times per 100 T-units). Contractions were also frequent (47 per 100 T-units).

Clearly, these figures mean very little unless they can be contrasted with those derived from formal prose, the opposite of the colloquial. The following excerpt from G. M. Trevelyan's A Shortened History of England provides an opportunity for such a contrast:

But the English East India Company, when driven from the Spice Islands, pushed its

trade on the Indian mainland. In James I's reign it founded a successful trading station at Surat, and in Charles I's reign built its Fort St. George, Madras, and set up other trading stations in Bengal. Such were the humble mercantile origins of British rule in India.⁸

In this excerpt, the T-units are substantially longer (20.5 words/T-unit, as calculated from a 636 word sample containing 31 T-units). There are no contractions or you's. A number of other counts might reveal significant differences between the two passages: percentage of monosyllabic words, percentage of nouns, and so on.

So far, though, only morphological classes have received attention. Using the two passages quoted above, an instructor could also illustrate a semantic classification: the general versus the particular. First, one must define the categories suitably. Since there are many different levels of generality, it is best to create three headings: general, particular, and very particular. A word (noun) is very particular when it can be replaced by two words, each more general than the other (the word neurosurgeon is very particular because it can be replaced by doctor, which can be replaced by person). Similarly, a particular word can be replaced by one more general word (the word anger is particular because it can be replaced by emotion). Finally, a word is general when nothing more general can be substituted for it. Verbs can also undergo this kind of analysis (to trot is very particular; to run is particular; and to go is general), while adverbs

and adjectives by their very nature make nouns and verbs more specific.

The colloquial sample from Working is relatively general (73 general, 54 particular, and 0 very particular verbs per 100 T-units; 100 general, 100 particular, and 9 very particular nouns per 100 T-units; 45 predicate, 9 attributive adjectives per 100 T-units; and 27 adverbs per 100 T-units). The Trevelyan passage is much more specific (33 general, 100 particular and 0 very particular verbs per 100 T-units; 33 general, 41 particular, and 400 very particular nouns per 100 T-units; 0 predicate, 260 attributive adjectives per 100 T-units; 22 adverbs per 100 T-units).

Of course, to make any generalizations about what constitutes particular and general writing, students and instructors must undertake many more counts. Categorizing of the samples is also necessary because different aims of discourse could throw off the results. The receptionist's monologue is basically expressive, whereas Trevelyan's history is referential--though both are essentially narrative. In historical narrative, proper nouns abound and account for most of the "very particular" score. In the receptionist's personal narrative, there is little opportunity for proper nouns, thus the low score. Assigning the samples to Kinnevan categories allows students to keep track of how the mode of discourse influences general or particular writing. Nevertheless students may write

generally because they tend toward colloquialism. Word counting is a good way to bring the matter to their attention.

After such analysis is complete, students can attempt a loose imitation of the writing under discussion, perhaps with the assistance of instructions similar to those offered in the last section.

In conclusion, many variations on these exercises are possible, to be sure. Instructors should guide students but should at the same time allow their charges to grapple with stylistic problems and think on their own. If past experience is any indication, students will find enumerative analysis invigorating--perhaps because it is new, perhaps because they are challenged with a problem and provided with specific procedures by which to solve that problem, or perhaps because they respect mathematics.

In any event, such methods as presented in this chapter (close imitation, definition, and attribution) are best employed as part of a writing program, not its main focus. Students tire quickly of tedious counting, and an overdose is likely to stir rebellion. Spaced at reasonable intervals in the curriculum, however, enumerative analysis can add substance to a study of style.

ENDNOTES

¹Quintillian, Institutio Oratoria, trans. H. E. Butler, vol. 4 in The Loeb Classical Library (Cambridge Massachusetts: Harvard University Press, 1922), p. 83. For more information about the theory of imitation, see Corbett's article (note 27, Chapter II) and Frank D'Angelo's "Imitation and Style," College Composition and Communication, 24 (1973), 283-290.

²W. Ross Winterowd, introduction to "The Grammar of Coherence," in Contemporary Rhetoric, p. 225.

³Studs Terkel, Working (New York: Avon Books, 1972), pp. 58-59.

⁴G. M. Trevelyan, A Shortened History of England (Harmondsworth, England: Penguin Books, Ltd., 1942), pp. 281-282.

⁵W. Jackson Bate, Samuel Johnson (New York: Harcourt Brace Jovanovich, 1975), p. 5.

⁶James Boswell, The Life of Samuel Johnson, with introduction by Frank Brady (New York: New American Library, 1968), pp. 34-35.

⁷John Wain, Samuel Johnson (New York: Viking Press, 1975), p. 17.

⁸Ibid., p. 261.

CHAPTER V

CONCLUSION

Enumerative analysis in the classroom can enhance students' writing ability if it takes place under three conditions. First, the design of the counts must be constructed with accurate insights: if students count insignificant structures, the exercise becomes useless or damaging. Second, such analysis must add emphasis to rhetorical considerations, relating form to content in a meaningful way. Finally, the technique can be valuable only when integrated into a writing program which provides instruction in invention and arrangement as well.

The value of the technique is not that it provides "objective" proof for stylistic pronouncements; rather, it helps to confirm and illustrate any statement about style, or indicate those which are not square with the facts. A second advantage is that the technique calls students' attention to style in a specific way and provides a terminology which avoids the subjective criticism that novice writers are accustomed to give--attention is focused not on the feelings of the student ("I don't like it," or "I really liked it") but on the text itself ("I found forty-five concrete nouns"). A third favorable effect of

enumerative analysis is that it allows students to compare certain stylistic features in their own writing to those of professional writers, and it can thus prepare these students for successful imitation exercises.

Finally, enumerative analysis does not discourage instructors from advising to "prefer the concrete to the abstract." On the contrary, such analysis forces the students to back up the assertion, to search for concrete evidence, and to qualify the pronouncement on the basis of that evidence. The problem with unsubstantiated dictums is that they often resemble helpful, parental warnings to "be good" or "be careful," the inevitable, obligatory reply to which is "I will." While students may ignore these kind warnings, they will certainly be better able to affect their writing if they can act on suggestions which give hints to the reason behind style.

A SELECTED BIBLIOGRAPHY

Aristotle. "Art" of Rhetoric. Translated by J. H. Freese in The Loeb Classical Library. Cambridge, Massachusetts: Harvard University Press, 1975.

Baker, Sheridan. The Complete Stylist and Handbook. New York: Harper & Row, Publishers, 1980.

Barnet, Sylvan and Marcia Stubbs. Practical Guide to Writing. Boston: Little, Brown and Company, 1980.

Bateman, Donald R. and Frank J. Zidonis. The Effect of a Study of Transformational Grammar on the Writing of Ninth and Tenth Graders: NCTE Research Report No. 6. Champaign, Illinois: National Council of Teachers of English, 1966.

Booth, Wayne C. "The Rhetorical Stance." College Composition and Communication, 14 (1963), 139-145.

Brinegar, Claude S. "Mark Twain and the Quintus Curtius Snodgrass Letters: A Statistical Test of Authorship." Journal of the American Statistical Association, 58 (1963), 85-96.

Christensen, Francis. Notes Toward a New Rhetoric: Six Essays for Teachers. New York: Harper & Row, Publishers, 1967.

----- . "The Problem of Defining a Mature Style." English Journal, 57 (1968), 572-579.

Corbett, Edward P. J. Classical Rhetoric for the Modern Student. New York: Oxford University Press, 1971.

----- . "The Theory and Practice of Imitation in Classical Rhetoric." College Composition and Communication, 22 (1971), 243-250.

----- . Edward P. J. "A Method of Analyzing Prose Style with a Demonstration Analysis of Swift's A Modest Proposal." The Writing Teacher's Sourcebook. Ed. Gary Tate and Edward P. J. Corbett. New York: Oxford University Press, 1981; pp. 333-352.

- Croce, Benedetto. Aesthetic as Science of Expression and General Linguistic. Translated by Douglas Ainslie. New York: Noonday Press, 1966.
- Dolezel, Lubomir and Richard W. Bailey. Statistics and Style. Number 6 of Mathematical Linguistics and Automatic Language Processing. New York: American Elsevier Publishing Company, 1969.
- Ellegard, Alvar. A Statistical Method for Determining Authorship: The Junius Letters, 1769-1772. Number 13 of Gothenburg Studies in English. Gothenburg, Sweden: University of Gothenburg, 1962.
- Fish, Stanley E. "What Is Stylistics and Why Are They Saying Such Terrible Things About It?" In Approaches to Poetics: Selected Papers from the English Institute. Ed. Seymour Chatman, New York: Columbia University Press, 1973.
- Flesch, Rudolph. The Art of Readable Writing. New York Harper and Row, Publishers, 1949.
- . A New Way to Better English. New York: Harper and Row, Publishers, 1958.
- Freeman, Donald C. Linguistics and Literary Style. New York: Holt, Rinehart and Winston, Inc., 1970.
- Gage, John. "Philosophies of Style and Their Implications for Composition." College English, 41 (1980), 615-622.
- Gibson, Walker. Tough, Sweet, and Stuff: An Essay on Modern American Styles. Bloomington, Indiana: Indiana University Press, 1966.
- Herdan, Gustav. The Advanced Theory of Language as Choice and Chance. Volume 4 of Kommunikation and Kybernetic Einzeldarstellungen. New York: Springer-Verlag, Inc., 1966.
- Hiatt, Mary P. Artful Balance: The Parallel Structures of Style. Volume 2 of the New Humanistic Research Series. New York: Teachers College Press, 1975.
- Hunt, Kellogg W. Grammatical Structures Written at Three Grade Levels: NCTE Research Report No. 3. Champaign, Illinois: National Council of Teachers of English, 1965.
- Johnson, Sabina Thorne. "Some Tentative Structures of Generative Rhetoric." In Contemporary Rhetoric: A Conceptual Background with Readings. Ed. W. Ross

- Winterowd. New York: Harcourt Brace Jovanovich, Inc., 1975.
- Kinneavy, James L. A Theory of Discourse: The Aims of Discourse. New York: W. W. Norton & Co., 1971.
- Mellon, John C. Transformational Sentence-Combining: A Method for Enhancing the Development of Syntactic Fluency in English Composition, NCTE Research Report No. 10. Champaign, Illinois: National Council of Teachers of English, 1969.
- Mendenhall, T. C. "A Mechanical Solution to a Literary Problem." The Popular Science Monthly, 77 (1901), 97-105.
- Milic, Louis Tonko. "Metaphysics in the Criticism of Style." College Composition and Communication, 17 (1966), 124-129.
- . A Quantitative Approach to the Style of Jonathan Swift. The Hague: Mouton & Company, 1967.
- . "Theories of Style and Their Implications for the Teaching of Composition." College Composition and Communication, 16 (1965), 66-69; 126.
- Moritz, Robert E. "On the Variation and Functional Relation of Certain Sentence-Constants in Standard Literature." University of Nebraska Studies, 3 (1894), 229-255.
- Mosteller, Frederick and David L. Wallace. "Inference in an Authorship Problem." Journal of the American Statistical Association, 53 (1963), 275-309.
- Murry, J. Middleton. The Problem of Style. London: Oxford University Press, 1936.
- Ohmann, Richard M. "Prolegomena to the Analysis of Prose Style." In Style in Prose Fiction: English Institute Essays. Ed. Harold C. Martin. New York: Columbia University Press, 1959.
- Quintillian. Institutio Oratoria. Translated by H. R. Butler. IV Volumes. In The Loeb Classical Library. Cambridge, Massachusetts: Harvard University Press, 1922.
- Richards, I. A. Practical Criticism: A Study of Literary Judgment. New York: Harcourt, Brace and Company, 1929.
- Sherman, L. A. "Some Observations upon the Sentence-Length

in English Prose." University of Nebraska Studies,
1 (1888), 119-130; 337-366.

Stone, Wilfred and J. G. Bell. Prose Style: A Handbook
for Writers. New York: McGraw-Hill Book Company,
1972.

Winterowd, W. Ross. Contemporary Rhetoric: A Conceptual
Background with Readings. New York: Harcourt Brace
Jovanovich, Inc., 1975.

Wimsatt, W. K. The Prose Style of Samuel Johnson. Volume
94 of Yale Studies in English. New Haven, Connecticut:
Yale University Press, 1941.

Yule, G. Undy. The Statistical Study of Literary Vocabulary.
Cambridge, England: University Press, 1944.

----- . "On Sentence-Length as a Statistical Character-
istic of Style in Prose, with Application to Two Cases
of Disputed Authorship." Biometrika, 30 (1938),
363-390.

VITA

Robert Joseph Hasenfratz

Candidate for the Degree of

Master of Arts

Thesis: STATISTICAL AND ENUMERATIVE ANALYSES OF STYLE
FOR THE CLASSROOM

Major Field: English

Biographical:

Personal Data: Born in Perry, Oklahoma, May 17, 1957,
son of Mr. and Mrs. W. L. Hasenfratz.

Education: Graduated from Kingfisher High School,
Kingfisher, Oklahoma, in May, 1975; received
Bachelor of Arts degree in English with a minor
in German from Oklahoma State University in May,
1979; completed requirements for the Master of
Arts degree at Oklahoma State University in
May, 1983.

Professional Experience: Graduate teaching assistant
Department of English, Oklahoma State University,
August, 1978-May, 1981; August-December, 1982.