TM-1908/300/00

Stylistic Analysis

Report on the Third Year of Research

(Final Report under this Contract)

1 March 1967

The research reported herein was conducted under SDC's independent research program and Contract Nonr-4427(00), Office of Naval Research, U. S. Navy.

SDC Document

TM-1908/300/00

TECHNICAL MEMORANDUM

(TM Series)

The research reported herein was conducted under SDC's independent research program and Contract Nonr-4427(00), Office of Naval Research, U. S. Navy, Task No. NR 348-005.

Stylistic Analysis

Report on the Third Year of Research
(Final Report under this Contract)

by

Sally Yeates Sedelow

1 March 1967

SYSTEM
DEVELOPMENT
CORPORATION
2500 COLORADO AVE.
SANTA MONICA
CALIFORNIA
90406

(page 2 blank)

Stylistic Analysis

Report on the Third Year of Research

by

Sally Yeates Sedelow

ABSTRACT

This report describes current research associated with the computer-based stylistic analysis project. Work directed toward the complete automation of thesaurus construction for the VIA (content analysis) program is reported in some detail; the focus of this work has been upon comparative VIA runs using Webster's Dictionary of Synonyms, Roget's International Thesaurus, and Roget's University Thesaurus. As a result of these runs, conclusions are drawn regarding preferable sources for the computer thesaurus as well as regarding some technical aspects of VIA's operation. The power of MAPTEXT for quickly revealing stylistic patterns to the information analyst or intelligence officer is illustrated.

PREFACE

This is the third annual report on research in stylistic analysis. For these three years, the research has been carried out at the System Development Corporation under a contract with the Office of Naval Research, Information Systems Branch, Contract Nonr-4427(00). I am very grateful to the Office of Naval Research for its generous financial assistance; I wish also to thank the members of the Information Systems Branch for their helpful suggestions and advice during the course of this project. The System Development Corporation has contributed administrative support and financial aid as well as many other kinds of research support; space prevents naming the many people at SDC who have in some way contributed to this project, but this general thanks is intended for each of them. Acknowledgment should also be made to Saint Louis University and, currently, the University of North Carolina at Chapel Hill, for teaching loads that provide time for research. In addition, I appreciate the helpful comments and suggestions of friends and colleagues in the wider professional and academic community; I am especially grateful to my husband, Walter A. Sedelow, Jr., who has generously given time to this project.

Once again, very special thanks are due Terry Ruggles of the System Development Corporation for major contributions to this research effort. He has been the sole programmer, as well as the coordinator for many of the day-to-day details associated with the project. In addition, he has helped with project documentation by contributing flow charts and extensive commentary regarding the descriptions of the programs' operations. He has also seen to

providing program listings, flow charts, etc., in response to a considerable number of requests for these materials from other research workers in the general area of computer-aided language analysis. In short, his supportive efforts for this project deserve generous thanks.

(page 6 blank)

TABLE OF CONTENTS

I. Introduction

The reasons for research on stylistic analysis have been discussed in some detail in a number of earlier publications related to this research project.[1] It seems sufficient to note here, therefore, that the ability to define the style of any verbal artifact, whether written or spoken, implies a definitive knowledge of the choice of words and the choice of arrangement and punctuation of those words within the set of verbal material being examined. Knowledge of word choice is of immediate importance to the information, or content, analyst who is looking for verbal documentation of suspected trends or attitudes or who is searching for new directions of attitude or thought in the verbal expressions of leaders or political groups in other countries. Knowledge of word arrangement and punctuation choices is of real importance to the information analyst who is attempting subtle distinctions among several different writers or speakers in an effort to reveal foci of disagreements or of shifts in view upon, for example, matters of diplomatic or military policy. To make these subtle distinctions, many different stylistic parameters may be required.

---

[1] Sally Y. Sedelow and Walter A. Sedelow, Jr., "A Preface to Computational Stylistics," in J. Leed (Ed.), The Computer and Literary Style: Introductory Essays and Studies, Kent State University Press, 1966, pp. 1-13 (also available as SDC document SP-1534, System Development Corporation, Santa Monica, California, 17 February 1964); S. Y. Sedelow, W. A. Sedelow, Jr., and T. L. Ruggles, "Some Parameters for Computational Stylistics," Proceedings, IBM Literary Data Processing Conference, IBM, Yorktown Heights, September, 1964; S. Y. Sedelow, "Stylistic Analysis: Report on the First Year of Research," SDC Document TM-1908/100/00, System Development Corporation, Santa Monica, California, 1 March 1965; S. Y. Sedelow, "Stylistic Analysis: Report on the Second Year of Research," SDC document TM-1908/200/00, System Development Corporation, Santa Monica, California, 1 March 1966.

The general goal of this research project is to produce a programming system comprising a number of dynamically interactive computer-based procedures for the analysis of style. Several research projects focused upon one or a few aspects of the style of a given author or group of authors have been successfully undertaken. McDonough has worked on metrics in the Iliad,[2] Miles and Selvin on the vocabulary of poetry in the seventeenth century,[3] and Ellegard on the word choices of possible authors of the Junius Letters.[4] The results of these efforts encourage the belief that aspects of style can be rigorously defined and that highly similar styles can be distinguished one from another. One unattractive aspect of computer-aided stylistic studies undertaken thus far is the heavy reliance upon detailed manual research by the investigator; often the computer is permitted to perform a few statistical tests only after the human investigator has laboriously sorted data into categories of use frequency, syntax, etc. This initially modest use of the computer is a result, in some cases, of insufficient money to purchase computer assistance and, in other cases, of insufficient knowledge about the properties of the aspect of style being investigated. The latter insufficiency means that it was not possible to furnish the computer with the recognition criteria that would enable it to sort data into the categories on which the discrimination of

---

[2] Proceedings, IBM Literary Data Processing Conference, pp. 25-36.

[3] Josephine Miles and Hanan Selvin, "A Factor Analysis of the Vocabulary of Poetry in the Seventeenth Century," in J. Leed (Ed.), op. cit., pp. 116-127.

[4] Alvar Ellegard, A Statistical Method for Determining Authorship: The Junius Letters, 1769-1772, Goteborg, 1962.

style was to be based. The study of metrics is a case in point. The perceptions of pitch, length, and amplitude are still too imperfectly understood for precise specification to the computer. Thus, the achievement of the ultimate goal for the stylistic analysis project--very fast analysis and discrimination of any style by the computer, without human intervention--will not be realized for some time. Nonetheless, it has been the practice of this project to define tasks so that the high-speed sorting and matching capacities of the computer are heavily utilized and so that the computer, whenever possible, may draw inferences from the work it has already completed. Thus, in VIA, the program that examines word choice in any given set of words, the computer program provides the links among all but the first level of associated words. MAPTEXT, the program that provides abstract representations of a text or texts, takes some of its input from VIA's output. As we enlarge the capacities of these programs and add other programs to the system, our goal will be to make the programs increasingly interactive so that one can "learn" from the efforts of another that has already operated. Only this kind of interdependence will produce the multipurpose, high-speed discrimination of style necessary for the fast study of verbal information and behavior on a real-time basis.

II. Research Procedures and Results

The efforts of the stylistic analysis project thus far have been concentrated upon word choice, with a lesser emphasis upon word arrangement. Last year's report upon this project reviewed in some detail the use of the VIA (Verbally-Indexed Associations) program to reveal content in texts representing a considerable spread in variety of word choice--two translations of Sokolovsky's Soviet Military Strategy, and Shakespeare's Hamlet.[5] The linked lists of words produced by VIA successfully showed organizing concepts within, for example, chapters from Soviet Military Strategy, as well as more subtle ramifications of those concepts. For instance, although there is emphasis upon both conventional and nuclear warfare in the first chapter, "General Concepts" (RAND translation), word frequencies as well as the number of different words on the linked lists indicate a greater emphasis upon conventional warfare--an emphasis that is confirmed by specific statements in the chapter. When tested upon Hamlet, a play that has been seemingly exhaustively analyzed for content

[5] See "Stylistic Analysis: Report on the Second Year of Research," pp. 14-57. Translations used were Marshal V. D. Sokolovsky (Ed.), Military Strategy: Soviet Doctrine and Concepts, Translation Services Branch, Foreign Technology Division, Wright-Patterson AFB, Ohio (Trans.), Frederick A. Praeger, Inc., New York, 1963. V. D. Sokolovskii (Ed.), Soviet Military Strategy, Dinerstein, Goure, and Wolfe (Trans.), RAND Corporation, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1963. (Because reference to the translations by translators' names or affiliation is unwieldy, throughout this report the translation published by Praeger will be called the Praeger translation, and that translated as a RAND Corporation research study [published by Prentice-Hall] will be called the RAND translation.) The edition of Hamlet used for this study was the Kittredge edition, Ginn and Company, 1939.

by scores of people expert in such analysis, VIA not only revealed the themes upon which noncomputer-aided analyses have focused but also picked out other significant concepts that had not been noticed in earlier "manual" investigations.

As it currently operates, VIA is dependent upon human intervention at one stage of its operation. This stage occurs after words appearing in the set of words being examined have been grouped together by root, and a frequency count of word occurrences by root groups has been made. At this point, the human investigator examines the output and, on whatever basis he selects--usually high frequency of occurrence--designates words he would like to investigate further. Next, the human investigator searches thesauri, synonym dictionaries, and contexts of the designated words for other words that, if they occur in the text, should be linked to the designated words. These lists of possible-associated-words are then submitted to the computer, which, after searching for these first-level list words, produces, on its own, sublists down to any desired level. Since the computer produces sublists, and since all suggested associated words are saved by the computer for later searches on other sections of text, human intervention is not preclusively onerous. Nonetheless, a goal of this research is a system of programs that can operate with a minimum of human assistance. Thus, other than a number of VIA runs on remaining chapters of Soviet Military Strategy, the effort connected with VIA this year has concentrated on research directed toward automating the search for first-level associated words.

So far as work with MAPTEXT is concerned, further experiments with representations have been conducted. A discussion of the results of these experiments occupies the second part of this section.

Work on further programs could not be undertaken this year because the Philco 2000, the computer on which the stylistic analysis programs have been running, was phased out at midpoint of the contract year for the stylistic analysis project. VIA and MAPTEXT are in the process of being rewritten for the IBM S/360 series of computers. While rewriting the programs, some features that experience has shown desirable will be added. Description of these plans appears on pages 90-92, Plans for the Future, of this report.

A. VIA

1. Introductory Comments on Frequency Counts. For the purposes of this report, it would be redundant to demonstrate again, on the basis of VIA's operation on chapters other than those discussed in last year's report, that VIA's output is a reliable, often subtle guide to content. Comments already made in this report (as well as in the documents mentioned in footnote 1, page 7, of this report) provide references to extensive appraisals of VIA's output. Since the initial operation of VIA is based upon frequency counts--most often high-frequency counts--it may be useful, however, to provide an anecdote about an exercise involving the high-frequency lists produced by the first subsections (INDEX and SORT) of the VIA program. This anecdote is not intended to serve as commentary upon the output produced by the complete VIA program. It is concerned, rather, with the validity of using high-frequency words as guides to gross content; VIA then expands and refines that content so that any false leads provided by the original frequency counts are immediately apparent. The interesting point here is how safe an initial guide a frequency count can be for the moderately informed reader. Too often such admittedly

gross guides to content are dismissed, by both the casual reader and the research scientist who are appraising such a list from a work or intellectual discipline they know nothing about, as thoroughly uninformative, therefore profitless to produce. However, in many situations, the individual working with such a list is by no means totally uninformed. He probably has read part of the work rather carefully or, at the least, he probably knows the area with which the work is concerned rather well. An intelligence analyst, for instance, might look at a small subsection of data carefully and then search for some quick guide to the gross content of the rest of the data. This anecdote, which involves an experience of my own with high-frequency word lists, is intended to relate to that kind of situation.

The lists of words with which I dealt were those in root groups that had more than 50 occurrences for each chapter in the RAND translation of Soviet Military Strategy. The task I set myself was to take the list for each chapter and then infer a probable title for the chapter on the basis of the words on the list. I had read the first two chapters in the book and, two years before this anecdotal experience, glanced at the table of contents of the Praeger, but not the RAND, translation. Given my experience with the book, I assumed that the postulated titles for the first two chapters would be a good deal closer to the target than those for the other chapters. Actually, this was not entirely the case. Figure 1 shows the root groups on which my guesses were based. In order to fit the words into the chart, etc. is sometimes used. Whenever "etc." is used, the root group has appeared in an earlier chapter and all the forms used were listed at that point.

1	2	3	4	5	6	7	8
aim aims aimed	air aircraft airfields	air, etc.	all	all	against	arms, etc.	act acts acted
all	America American Americans American's	area areas	arms, etc.	arms, etc	aggression aggressive aggressor aggressors aggressors' aggressor's	country, etc.	action actions active activity activities
arms armed army army's armies armament armaments	arms, etc.	arms, etc.	capital capitalism capitalist capitalistic capitalism's capita	combat combating combative	aim aims aimed	force, etc.	actually actual
country country's countries countries'	Britain Britain's British	attack	country, etc.	develop, etc.	air, etc.	important importance	agency agencies agent
develop develops developed developing development	country, etc.	combat conduct defend defending defense defenses defensive	destroy destroyed destroying destruction destructive	enemy enemy's enemies	all	industry industries industrial	all
economy economies economic economics economists economically	defense	destroy	develop, etc.	equip equipment equipped	America, etc.	military militarily	arms, etc.
force forces forced forceful forcibly	develop, etc.	develop	force, etc.	force forces forceful	arms, etc.	mobilize mobilizes mobilizing mobilized mobilization	central centralized centralization centers
general generals generally generalized generalizations	economic	enemy enemy's enemies	imperialist imperialists imperialists' imperialism imperialistic	military militaristic	attack attacks attacked attacking attacker's	operate operates operating operation operational operations operators	chief chiefs
	equip equipment equipped equipping	force, etc.	important importance	missile missiles	base bases based	peace peaceful peacetime peaceable	command commands commanding commanded commander commanders
	force forces forced	formation formations	military militarily	new newest	begin begins beginning began begun	prepare prepared preparing preparatory preparation preparations	committee committees
	general	form forms forming	missile missiles	nuclear	Britain Britain's Britainish	transport transported transporting transportation	communist communists communism
	German Germans Germany Germany's	front frontal	nuclear	operate operation operations operational	carry carrying carried carrier carriers		country countries country's
	increase	great greatly greatest greater	other others	organization organizational organize organized organizing			
	industry industries industrial industrially	important importance	people peoples peoples' people's	other others			
	large largely largest	increase increasing increased		social, etc.			

Figure 1. Root Groups with over 50 Occurrences per Chapter, Soviet Military Strategy, RAND translation

1	2	3	4	5	6	7	8
military militarily militarist militarists militaristic	military	line lines	politics political	state, etc.	combat combated combating	war wars warships wartime warfare	create creative creating creation created
morale	missile	main mainly	social socialist socialists socialism socialism's	troop troops	conduct conducted conducting		defend defending defensive defense
operation operations operational operating	more	military	Soviet	war, etc.	country, etc.		develop develops developing developed development
policy policies politics political politically	mostly	most	state, etc.	weapon weapons weapon's	defeat defeated defeating		direct directs directly directing direction directions directives directed director directors
social socialist socialism socialization	NATO	nineteen-eightee nineteen-forty-f nineteen-forty-o nineteen-forty-t nineteen-forty nineteen-ninetee nineteen-sixteen nineteen-thirty nineteen-twenty	unit, etc.		defend defended defenders defending defense defenseless defenses defensive defensives		force forces forces'
state states state's states' stated statement stating statesmen station stations	naval		war, etc.		deliver delivery delivering delivered deliveries		front fronts fronts' front's
strategic strategically strategy strategies strategists	nuclear		weapon weapons		deploy deployed deploying deployment		general, etc.
	number		world world's		destroy destroying destroyed destroyers destruction destructive		high highly higher highest
	operation,etc.	offensive offensives offense			develop, etc.		
	other	operate operated operating operation, etc.					
	plan planned plans	organize organized organizing organization organizations organizational organizationally					
	prepare prepared preparing preparedness preparation preparations	politics political politically					
	produce produces producing produced production product products productive	red					
	state, etc.						
	strategy, etc.						
	system systems						

Figure 1. (Continued)

1	2	3	6 contd.	6 contd.	6 contd.	8 contd.	8 contd.
unit units unity united unite unites unified union union's	time unit, etc.	reserve reserves Soviet	divide divided division divisions	main mainly	prepare prepared preparing preparedness preparation preparations	important importance	supreme
war wars warring warfare wartime	war, etc.	strategy strategic strategically	enemy enemy's enemies	method methods	social socialism socialist	lead leading leader leaders leadership	troop troops troops'
	weapon	time times timing	force forces forces' force's forced forcing	military militarily militaristic	Soviet	military militarily	unit units unity united unify unified unification union union's
	west western	troop troops troops'	form forms formation formations	missile missiles	state, etc.	most	
	world world's	unit, etc.	front fronts frontal	modern modernizing	strategy strategic strategically	navy navy's naval	war wars warfare wartime
		use used	German Germans Germans' Germany Germany's	most	strike strikes striking	operate, etc.	
		war wars warfare	great greatly greater	naval navy navies navy's	submarine submarines	organize, etc.	
			ground	nuclear	target targets targeting	party party's	
			imperial imperialist imperialists imperialism imperialism's	offense offensive offensives	theater theaters	people people's	
			important importance	operate operated operation operations operational	troop troops	plan plans planned planning	
				other others	unit, etc.	political	
				out outer outside	used useful useless	Soviet Soviets Soviet's	
				plan plans planned	use war, etc.	staff staffs staffs' staffed	
					weapon weapons	state, etc.	
					west western	strategy strategic	
					world world's worldwide		

Figure 1. (Continued)

The relevant totals for Figure 1 are as follows: in chapter one, there are 17 root groups having a total number of occurrences greater than 50 (the total number of content word occurrences within the chapter is 10,009); in chapter two, 37 root groups (13,380 content word occurrences); in chapter three, 33 root groups (12,649 content word occurrences); in chapter four, 22 root groups (9,376 content word occurrences); in chapter five, 19 root groups (7,793 content word occurrences); in chapter six, 61 root groups (17,755 content word occurrences); in chapter seven, 12 root groups (7,283 content word occurrences); and, in chapter eight, 37 root groups (10,458 content word occurrences).

Other than for the first two chapters, the best guess as to content might be expected to occur for chapter six, which had by far the largest number of root groups with 50 or more occurrences. This expectation may be a trifle misleading, though, because chapter six also had by far the largest number of content word occurrences. The guesses as to the major concept structuring a chapter compare with the title actually given that chapter in the RAND translation as follows:

Chapter One: Guess: Introduction;
Chapter Title: General Concepts.

Chapter Two: Guess: Background re Soviet strategy vis-à-vis Western world;
Chapter Title: Military Strategy of Imperialist States and their Preparation of New Wars.

Chapter Three: Guess: Resume of earlier wars, especially World War II, but also World War I;
Chapter Title: Development of Soviet Military Strategy, 1917-1945.

Chapter Four: Guess: Chapter has to do with capitalism, imperialism;
Chapter Title: The Nature of Modern War.

Chapter Five: Guess: Chapter concerned with modern warfare and equipment;
Chapter Title: The Problem of Organizing and Developing the Armed Forces.

Chapter Six: Guess: Again has to do with modern weapons and warfare vis-à-vis Western world;
Chapter Title: Methods of Conducting Warfare.

Chapter Seven: Guess: Has to do with developing military preparedness during peacetime;
Chapter Title: Preparation of a Country to Repel Aggression.

Chapter Eight: Guess: Summation of plans proposed for military organization security, including Party organization; calls for certain actions and organizational procedures;
Chapter Title: Command of the Armed Forces.

Of all the estimates as to chapter content, the guess for chapter four appears to be furthest off the mark. But an examination of chapter four shows that it's not so far off after all, because the emphasis of the chapter is upon the economic causes and bases for modern war. The estimates for chapters one, two, and six are quite good, as expected. But so is the estimate for chapter seven, which had only 12 root groups with more than 50 occurrences. The estimate for chapter three probably would not have been so successful had the words associated with dates in the nineteen forties, etc., not been included in the list. The appearance of this range of dates (the words are cut off because the maximum word length in the present program is 16 letters) gave rise to the guess that the chapter is concerned with some sort of survey of earlier military encounters.

What this little exercise shows is that for one reader with partial information about the contents of a book, lists of high-frequency root groups for major divisions in the book provided rather reliable guides to gross

content. For situations in which an intelligence analyst is looking at a very large body of information that has not been neatly done up in book form with chapter headings, close study of a fraction of the information combined with an examination of high-frequency root groups from other segments of the information may be a rather good way of isolating gross areas for closer examination with a full package of stylistic analysis programs. One would suppose, in fact, that many people concerned with the analysis of verbal data, whether in literature, intelligence reports, or some other form, have been, for years, intuitively using an analogous procedure for noncomputer-aided analyses. No matter how casually much of the material may be scanned, however, human effort as well as human inaccuracy is involved. Since the advent of computers, some verbal analysts have tended to derrogate the use of frequency counts as guides to content--perhaps in part because such counts do imply an accuracy that some find uncomfortable. Obviously, no one should be satisfied with lists of high-frequency words as the final commentary on the content of a given set of data, but on the other hand such lists can usefully serve as guides and bases for further detailed study with programs such as VIA and MAPTEXT. In fact, the computer-derived frequencies for chapter three, indicating an emphasis upon capitalism and imperialism, indicate more clearly the Marxist view of war than does the actual title for that chapter.

2. Comparison of Thesauri and a Synonym Dictionary. Most of the effort devoted to VIA this year was directed toward completely automating its operation (see pages 11-12 of this report). Since a major component of the human

intervention entails searches of thesauri and synonym dictionaries, it seemed sensible to begin the effort toward complete automation with some thesaurus and synonym dictionary comparisons. The thesauri and dictionary used most heavily for the manual compilation of lists were Roget's International Thesaurus (Third Edition, Thomas Y. Crowell Company, New York, 1962), Roget's University Thesaurus (Thomas Y. Crowell Company, New York, 1963), and Webster's Dictionary of Synonyms (G. & C. Merriam Company, Springfield, Massachusetts, 1951). Although Roget's University Thesaurus is described as "Originally published as Roget's International Thesaurus of English Words and Phrases" (copyright page, Roget's University Thesaurus), the indices for these two thesauri are quite different, thus producing rather disparate results when used with VIA. These three reference works were, then, the subject of the initial experiments.

In passing, it might be worth noting that the work with Roget's seems particularly valuable not only for this project but also for the many efforts focused upon machine translation, automatic abstracting, and associated computer-oriented natural-language research. Although many comments have been made about the need for new thesauri or the revision of existing thesauri for various computer-based research projects, there is a real paucity of specific detailed information about efforts to use thesauri that are currently available. As a result, directions for reform or revision are very vaguely defined or, indeed, lost in the various contemporary attempts to begin afresh the classification of words. Although the theories on which various classifications are based may seem far apart, it is possible that the practical

realizations of such theories are not so far apart. For example, an effort to classify words in terms of concepts basic to man's intellectual structuring of his environment or his responses to that environment may not produce results terribly far removed from the classification of words on the basis of their observed semantic usage. The real problem may be that a great many new words, representing new areas of knowledge, need to be absorbed into or added to existing classification schemes. Such absorption or addition may well change classifications as well as the emphasis of groups of classifications, but, nonetheless, much of the apparatus already existing presumably will be usable. In any event, whatever classification scheme ultimately proves most valuable for any given purpose, it does seem that a closer examination of the schemes currently in use is indicated and that the results of the examination should be made generally available to other research efforts faced with similar problems.

The procedure for initial experiments with thesauri entailed first arbitrarily selecting a set of words appearing in the RAND translation of Soviet Military Strategy.[6] The words chosen were those in the root groups associated with DEAD, DECLINE, DEFLECT, DELIBERATE, DELUSION, DETONATE, DEVELOP,

[6] Words appearing in just one of the translations were selected because of a related research interest in the nature of the disparities between the two translations. That is, does a difference in word choice merely reflect minor differences in the translators' taste or major differences of interpretation? A comparison of the linked lists associated with the words appearing in just one translation, with concepts revealed in comparable groups of linked words in the other translation, might help provide an answer.

DEVISE, DOUBLE, and DRILL. These root groups represent all those within the D-alphabetic range that occur in the RAND but not in the Praeger translation. Owing to limitations of computer space, some words within the index of a given reference thesaurus or within the synonym dictionary that might often be considered part of a designated root group, were not included in the list on which the search for associated words was based. For example, in the words taken from Roget's International Thesaurus for the search for words connected with DEAD, DEADEN and DEADENING were included but not DEATH, DEATHLY, etc. No hyphenated forms (such as DEAD-BEAT), or multiword idiomatic expressions (such as DEAD LETTER, DEAD TO THE WORLD, etc.) from the thesauri indices were included. However, hyphenated forms within the entries to which the index words pointed were included on the final list, described below. After the root groups had been selected and defined, the appropriate entries in the thesauri and synonym dictionary were consulted. In the case of the synonym dictionary, words were looked up directly in the dictionary; for the thesauri, the indices were used as guides to the entries. Having reached the proper section or sections in the thesauri and synonym dictionary, all words listed as in any way connected with the word being examined were recorded.[7] These words were then punched on cards and submitted

[7]Operationally, this means that all words (excluding multiple word idioms) in the appropriate thesauri entries were recorded. In Webster's Dictionary of Synonyms, the information under a given word often includes historical material that is, of course, full of words not synonymous with the word heading the entry; therefore, the synonym dictionary entries were perused "manually" for synonyms and antonyms.

to the computer, which searched the text--in this case chapter one of the RAND translation of Soviet Military Strategy--for word occurrences. A separate computer run was made for each thesaurus as well as for the dictionary. Figure 2 shows the words submitted for each reference word. An examination of these lists shows real disparities among both the number and apparent relevance of the words.

Webster's Dictionary of Synonyms

DEADLY	FATAL LETHAL MORTAL
DEAD	LATE LIFELESS QUICK ALIVE DECEASED DEFUNCT DEPARTED INANIMATE LIVING

Roget's International Thesaurus

DEADENING	NUMBING WEAKENING ANESTHETIC MITIGATING MODERATION RELIEF RELIEVING	DEADLINESS	DEATHLIKENESS FATALITY VIRULENCE
		DEADLY	EXTREMELY GREAT DEATHLIKE DEATHLY FATAL VIRULENT
DEADEN	WEAKEN CUSHION MODERATE MUFFLE NUMB RELIEVE	DEADNESS	DULLNESS INSIPIDNESS LACKLUSTER NUMBNESS

Figure 2. DEAD Root Group

DEAD	ANNIHILATED	BEAT-UP	DUST
	ARID	BEAT	ELEPHANTINE
	BARREN	BENUMBED	EXANIMATE
	BLUE	BLANK	EXPIRED
	CALLOUS	BLIND-ALLEY	EXTINCT
	CHARACTERLESS	BLIND	FALLEN
	CLEAR	BONES	FINISHED
	CLOSED	BREATHLESS	FISHY
	COLD	BUSHED	FLAT
	DEADER	CADAVER	FLAVORLESS
	DONE	CAECAL	FOUL
	DULLED	CARCASS	GLASSY
	EARTH	CARRION	GLAZED
	EMPTY	CLAY	GONE-BY
	EXHAUSTED	CLEAN	GONE
	EXPRESSIONLESS	COLORLESS	HALF-CONSCIOUS
	HEAVY	COMATOSE	HEBETUDINOUS
	LATE	CORPSE	HUSH
	LEADED	CROAKED	INACTIVE
	LEADEN	CROWBAIT	INAUDIBILITY
	LIFELESS	DAMPENED	INDIFFERENT
	NUMBED	DEADENED	INERT
	OBDURATE	DEAD-ALIVE	INEXPRESSIVE
	OBSOLETE	DEAD-BEAT	INSENSATE
	OUT	DEAD-END	INSENSIBLE
	PASSE	DEAD-TIRED	INSENTIENT
	PAST	DEADPAN	INSIPID
	PASSIVE	DEATH-STRICKEN	JEJUNE
	PEACE	DEATH-STRUCK	LACKLUSTER
	POINTLESS	DEATHLIKENESS	LACKLUSTROUS
	REGULAR	DECEASED	LANGUID
	REMAINS	DECEDENT	LANGUOROUS
	REST	DEFUNCT	LATE-LAMENTED
	SENSELESS	DEMISED	LATENT
	SLOW	DEPARTED	LENTITUDINOUS
	SPIRITLESS	DISUSED	LETHARGIC
	SUPPRESSED	DOG-TIRED	LISTLESS
	TONELESS	DOG-WEARY	LOGY
	TOTAL	DOPEY	LULL
	UTTER	DORMANT	LUMPISH
	WATERY	DRAB	LUSTERLESS
	WEAK	DREARISOME	LYMPHATIC
	ABEYANT	DREARY	MAT
	ABSOLUTE	DROOPY	MILD
	ASHES	DRY-AS-DUST	MILK-AND-WATER
	ASLEEP	DRY	MUFFLED
	BALD	DULL	MUMMIFICATION
	BEATEN	DUN	MUMMY

Figure 2. (Continued)

DEAD, contd.

MUM
MUTED
NAPOO
NOISELESSNESS
NONEXISTENT
NUMB
OBLIVIOUS
OBTUSE
OUT-WORN
PEDESTRIAN
PEPLESS
PERISHED
PHLEGMATIC
PLAIN
PLODDING
PLUMB
POKEY
POKY
PONDEROUS
POOPED
PROSTRATE
PURE
QUIETNESS
QUIET
QUIETUDE
RELEASED
RELICS
RELIQUIAE
REPOSING
REPRESSED
RUN-OUT
SAVORLESS
SEMICONSCIOUS
SHEER
SHUSH
SHUT
SILENCE
SILENTNESS
SLACK
SLEEPING
SLEEPY
SLUGGISH
SLUMBERING
SLUMBER
SMOLDERING
SMOTHERED
SOFTENED
SORDINE
SOUNDLESSNESS
SOURDINE
SPENT
STAGNANT
STALE
STANDING
STATIC
STIFF
STIFLED
STILLBORN
STILLNESS
STILL
STODGY
STUFFY
STUPID
SUBDUED
SUPINE
SUSPENDED
TAME
TASTELESS
TEDIOUS
THIN
TONELESSNESS
TORPID
UNAWAKENED
UNCONSCIOUS
UNEXPRESSIVE
UNFEELING
UNFELT
UNFLAVORED
UNINTELLIGENT
UNLIVELY
UNOPENED
UNOPEN
UNSAVORY
UNVENTED
UNVENTILATED
UNWAKENED
VACANT
VANISHED
VAPID
VERITABLE
WAKELESS
WASHED-UP
WASHY
WEARY
WHIST
WISHY-WASHY
WOODEN
WORN-OUT

Roget's University Thesaurus

DEAD*

ACCESSARY
ACCOUNT
ACTIONABLE
ADD
ADDITION
ADVANCE
ADVOCATE
AEGIS
AGAINST
AGE
AIDANCE
AIDING
AID
ALLY
ALLNESS
ALL
ALTOGETHER
ANALYTIC
ANALYSIS
ANARCHICAL
ANARCHIC
ANARCHY
ANCIENT
ANTAGONISM
ANTAGONIZE
ANTAGONISTIC
ANTIQUE
ANTIQUITY
APPRAISEMENT
APPRAISE

*Because of the length of the lists, only the entries for DEAD from the University Thesaurus are shown.

Figure 2. (Continued)

DEAD, contd.

ARMET
ARMOR
ARMED
ARM
ASSIST
ASSISTANCE
BALANCE
BARREN
BASTION
BATTLEMENT
BENEFIT
BLACKNESS
BLACK
BOOKED
BREAK
CALCULATE
CALCULATOR
CALCULATION
CALLOUS
CLASSIC
COLD
COMMENSURATE
COMA
COMPETE
COMPETITIVE
COMPETITION
COMPLETENESS
COMPLETING
COMPLETION
COMPLETELY
COMPLETE
CONDUCE
CONFRONT
CONTRADICT
CONTRARINESS
CONTRARY
CONTRIBUTE
CONTROL
COUNTABLE
COUNT
CUSTOMARY
CUSTOM
DEAD
DEADS
DEADLY
DECLINE
DEFENDED
DEFEND
DEFENDER
DEFENDING
DEFENSIVELY
DEFENSE
DEFENSIVE
DEMONSTRATE
DEPOSE
DEPOSITION
DIFFERENTIATION
DIVIDERS
DIVIDE
DIVISION
EMPTY
ENCOURAGE
END
ENTERTAIN
ENTIRETY
ENTIRE
ENUMERATE
ENUMERATION
ESTIMATION
ESTIMATE
EVEN
EVOLUTION
EXHAUSTIVE
EXPEDITE
FASTNESS
FAVORABLE
FAVOR
FOOD
FORCE
FORTIFICATION
FRAUGHT
FRUITLESS
FULL
FULLY
FURTHERANCE
FURTHER
GAINLESS
GOOD
GO
GRADUATION
GRADUATE
HARD
HEAVINESS
HEAVY
HELPFUL
HELP
HINDER
HOLD
HOUSINGS
IDEAL
INADEQUACY
INADEQUATE
INDEX
INTEGRATION
INTEREST
KEEP
KILLINGTIME
LAGGARD
LAGGING
LAG
LATE
LAWLESS
LEADEN
LEAVINGS
LIFELESS
LIGHTLESS
LIMIT
LINE
LINES
LOAD
LOSS
LOWER
MAINTENANCE
MAN
MATHEMATICIAN
MEANS
MEASURABLE
MEASUREMENT
MEASURING
MEASURE
METE
METER
METRICAL
METRIC
METAGE
MORTON
NARCOSIS
NEGLECT
NOTATION
NUMBERABLE
NUMBERING

Figure 2. (Continued)

DEAD, contd.

NUMBER
OBLIGE
OBLIGING
OBSOLETE
OPPOSED
OPPOSE
OPPOSING
OPPOSITION
PAGE
PASSE
PAST
PASSIVITY
PASSIVE
PERFECTION
PERFECT
PERMANENCE
PRACTICE
PRESERVATION
PRIME
PROBE
PROFITLESS
PROHIBITED
PROOF
PROVE
PROPITIOUS
PROPORTION
PROTECTION
PROTECTIVE
PROTECTOR
PUTTER
QUICKEN
RACE
RADICAL
RATE
RECENSION
RECKONABLE
RECKONING
RECKON
REDUCTION
REFUSE
REGULAR
RESISTANT
RESISTANCE
REST
RIPENESS
RIVAL
RIVALRY
ROOTED
RULE
SCALE
SCREEN
SECOND
SELF-PRESERVATIO
SENSELESS
SERVE
SLOW
SOLIDARITY
SOUND
STEM
STEP
STRIKER
SUFFICIENCY
SUMMATION
SUMMARY
SUM
SUNLESS
SUPPLIES
SUPPORT
SUSTENANCE
SUSTAIN
TARGE
TARGET
TEND
THOUGH
THROUGHOUT
TOTALNESS
TOTALITY
TOTALLY
TOTAL
TRADITIONAL
TRADITIONARY
TRADITIVE
TRADITION
TRUANT
UNBRIDLED
UNFAVORABLY
UNFAVORABLE
UNITY
USELESS
UTTERLY
VALUATION
VALUELESS
VALUE
VIOLATE
VIOLATION
WEAK
WHITEN
WHITE
WHOLE
WHOLLY
WINTERLY
WINTRY
WINTER
WITHSTAND
ABUNDANT
INSENSIBLE
ECLIPSE
INSERVIENT
INSIGNIFICANT
HABERGEON
BREASTPLATE
INSUBORDINATION
BREAST
BREASTWORK
INTACTNESS
ABACIST
INTEGRALITY
INTEGRALNESS
HYPSOMETRICAL
INTEGRITY
HYPSOMETRIC
ECSTASIS
INTERPOLATION
INTERREGNUM
INTRENCHMENT
INTRENCH
ANESTHETIZE
EFFECTUALLY
INUTILE
INUTILITY
BRIGANDINE
BRIMFUL
BRIMMER
BRIMMING
INVETERATE
INVOLUTION
IRONCLAD
IRON-PLATED
AUTUMNAL
IRRESOLUTION
IRRESPONSIBILITY

Figure 2. (Continued)

DEAD, contd.	IRRESPONSIBLE	BUM	LANQUOR
	JABBER	EMULOUS	BRUMOUS
	EFFETE	LULLABY	AUTUMN
	JAMBE	ENCOUNTER	AUXILIARY
	LETHARGY	LULL	EMBANKMENT
	BULLET-PROOF	LUMBER	LATENCY
	LIBERTY	LUMPISH	LATENT
	LICENSED	LUNETTE	ILLEGALITY
	LICENSE	LURID	ANAEMIC
	LICENTIOUSNESS	LUSK	LAWLESSNESS
	ILLEGITIMATE	ATHWART	ILLEGALLY
	LIFT	MACHICOLATED	LAXITY
	LIGHT-COLORED	MACHICOLATION	LAXNESS
	LIGHTLESSNESS	MAIL	LAX
	ILLICITNESS	DEPARTURE	BANQUETTE
	ILLICIT	DOZY	LAZY
	DORMANT	MAKEWEIGHT	LAZZARONE
	DORMOUSE	JARGON	ILLEGAL
	ADVOCACY	ELDERN	DEATHBLOW
	LISTLESS	ELDER	ILLEGITIMACY
	LITTER	ELDERSHIP	GLOOM
	DOUT	ELDEST	GONE
	LOAFER	JORDAN	BUCKLER
	LOAF	ELD	BUFFER
	BULWARK	JUNK	LENTOR
	LOGISTICS	HYPSOMETRY	ADVERSE
	LOGLINE	KILA	EMULATE
	LOGOMETER	BARBICAN	LETHARGICAL
	LOG	HYPNOTIZE	LETHARGIC
	LOITER	ASLEEP	EMULATION
	LOLLOP	KILOWATT	ERGOPHOBIA
	LOLL	KNIGHT-ERRANT	MISRULE
	LOOPHOLED	LAAGER	ESCAPE
	LOOPHOLE	LACKADAISICAL	CADASTRATION
	LOOSENESS	LACKLUSTER	MITIGATE
	LOOSENING	LADEN	CADASTRE
	LOOSE	IDLENESS	CADAVEROUS
	LORICA	IDLER	MOAT
	ILL-SPENT	IDLE	MOBOCRACY
	LOUNGER	DECAY	CADGER
	LOUNGE	DECEASED	ANTEDILUVIAN
	DOZE	LAME	ESCARP
	BARAGOUIN	LANDMARK	DEATHWARRANT
	BUMMER	BRUMAL	CALCULABLE
	LUBBARD	LANGUID	MOLE
	LUBBER	LANGUISH	DESPOTIC
	BUMPER	LANGUOR	HOLLOW

Figure 2. (Continued)

DEAD, contd.	HAHA	DECEASE	MINE
	GUARDIANSHIP	HOBOISM	MINISTRANT
	BARRACOON	HOCK	MINISTRATION
	AZRAEL	HOBO	MINISTRY
	MONOCHROME	DRAWL	MIRKINESS
	MOONSHINE	ENTRENCHMENT	MIRK
	MOPUS	MANTELET	MIRKSOME
	ETHER	MANTLET	MIRKY
	ETIOLATED	MANURE	EQUATIONS
	ETIOLATE	BUTTRESS	BELIE
	MORIBUND	MARMOT	IMPERCIPIENT
	MORIENT	GLOOMY	BELLYFUL
	BEDTIME	ENTRENCH	CAPITOL
	MORPHEUS	MASK	GLACIS
	MORTALITY	HOCUS-POCUS	ACCESSORY
	ETIOLATION	HYPNOTISM	ABATTIS
	DIE	DESPOTISM	NUMB
	MORTIFICATION	MATURENESS	NUMERAL
	MORTUARY	MATURITY	NUMERATION
	MOTHEATEN	MAUDLIN	NUMERICAL
	MOTIONLESS	DEMISED	NURSE
	MOUND	MEANINGLESS	NURTURE
	MUDDY	DEFENCE	NUTRITION
	CALCULUS	GEODAESIA	OBFUSCATE
	CALIBRATE	ABDICATION	OBIT
	CALIGATION	BEDARKEN	OBITUARY
	CALIGINOUSNESS	AFTERNOON	DESPOTICAL
	MULTIPLICATION	IMBECILITY	BENIGHTED
	MULTIPLY	GEODESIA	EVENTIDE
	MUNIMENT	DREAM	OBSCURATION
	MUNITION	IMMEMORIAL	OBSCURE
	MURAL	MEDIEVAL	OBSCURITY
	BABBLE	MENSURATION	IMPLETION
	EUTHANASIA	DEMISE	OBSTINACY
	EUTHANASY	MERLON	CARAPACE
	MURKINESS	BEDIM	OBTENEBRATE
	MURK	IMMUNE	OBTENEBRATION
	MURKSOME	DREAMY	EVE
	MURKY	IMMUNITY	GLASSY
	CALIGINOUS	ABSURDITY	HIEMS
	MUSTER	METROLOGY	OBTUND
	CALIPER	MIDNIGHT	DIKE
	CALIPERS	GEODESIST	OBTUSENESS
	DETHRONEMENT	CABASSET	CARTOGRAPHER
	HOARY	CADASTER	OBTUSE
	DETHRONE	CADASTRAL	OBUMBRATE
	ENGARRISON	DIAL	OBUMBRATION

Figure 2. (Continued)

DEAD, contd.	BACKPLATE	NIGHTFULL	OSCITATION
	OCHLOCRACY	NIGHT	OTIOSE
	OFFSCOURINGS	NIGHTTIDE	OTIOSITY
	OFFUSCATE	NIHILISM	OUT-OF-DATE
	OFFUSCATION	NIHILISTIC	OUT-OF-FASHION
	CARTOGRAPHICAL	EVADE	OUTLAWRY
	OLDEN	NOCTIVAGANT	EXANIMATE
	OLD-FANGLED	NOCTIVAGOUS	OUTRIGHT
	OLD-FASHIONED	NOCTURNAL	OUTWORK
	OLD-WORLD	EVASION	OUTWORN
	OLDNESS	NODDING	DILATORY
	OLD	NOD	OVERCAST
	DISSOLUTION	EVENING	OVERHAUL
	CARTOGRAPHIC	GONIOMETER	OVER-THWART
	OPIATE	CAMAIL	OVERSHADOW
	OPITULATION	NONCOOPERATION	OVERSLEEP
	CARTOGRAPHY	NONCOOPERATOR	OVERTHWART
	OPIUM	NONIUS	PACE
	ABET	NONOBSERVANCE	AMICABLE
	DRONE	NONPERFORMANCE	PACHYDERMATOUS
	IMPOTENT	DONJON	ALGEBRAIST
	IMPUGNATION	CANDLELIGHT	ALGEBRA
	ANTEMUNDANE	NONSENSE	ALGEBRAIZE
	ANTINOMY	NONSENSICAL	IMPUGNMENT
	ACCOUTRE	IMPEACHMENT	PAGINATE
	NAPPY	NOURISHMENT	PAGINATION
	NAP	NOURISH	PAH
	ABACUS	CAPARISON	EXCULPATION
	GEODESY	NOVOCAIN	EXCUSABLE
	BEFRIEND	DISSERVICE	PALADIN
	ABATIS	EVENSONG	PALAEOCRYSTIC
	HELM	NUGACITY	PALAVER
	HALF-MOON	CAPITATION	PALE-FACED
	EUTROPHY	NUGATORY	PALENESS
	NARCOTIZATION	ALGEBRAIC	PALEOANTHROPIC
	ABUTMENT	HYPNOSIS	PALEOZOIC
	NECKGUARD	BATHOMETER	PALE
	NECROLOGY	ADMEASUREMENT	PALING
	NEEDLES	OPPUGNANCY	PALISADE
	GEODETICAL	OPPUGNANT	CASE-HARDENED
	DREGS	OPPUGNATION	CASEMATED
	NEIGHBORLY	OPPUGN	PALLIDITY
	ANAESTHETIZE	GIBBERISH	PALLID
	ANTIQUATED	ORTHOMETRY	PALLOR
	GEODETICS	ORTS	CASEMATE
	NIAISERIE	OSCITANCY	PALL
	NICELY	OSCITANT	CASQUE

Figure 2. (Continued)

DEAD, contd.	PALSIED	HYPSOGRAPHICAL	ALIMENTATION
	CASQUETEL	EXEMPT	ADUMBRATE
	CASTELLATED	GAUNTLET	PREADAMITE
	PALSY	ADESPOTIC	PREGLACIAL
	EXCUSE	DISUSED	CHIAROSCURO
	PANDICULATION	BANK	PREHISTORIC
	PANOPLIED	EXONERATE	PRE-RAPHAELITE
	PANOPLY	CHAMFRAIN	EYESERVER
	CASTLE	CHAMFRON	PRESCRIPTION
	PARALYSIS	PHOTOMETER	PRESCRIPTIVE
	ANAESTHESIA	EXONERATION	HYPSOGRAPHIC
	PARALYTIC	CHAMPION	DRONY
	HIBERNATE	CHAMPIONSHIP	PRIMEVAL
	CATALEPSY	PICKELHAUBE	PRIMIGENIAL
	ACCOMMODATE	PICKET	PRIMIGENIOUS
	PARALYZE	PIDDLE	PRIMIGENOUS
	PARAPET	HORNWORK	PRIMITIVE
	DECOLORIZATION	EXPIRE	PRIMOGENITURE
	PARRY	EXPLODED	PRIMORDIAL
	PA	EXTINCTION	PRIMORDIATE
	DECOLORATION	PITCHY	INANE
	CENSUS	DISMANTLE	PROCRASTINATION
	DRONISH	EXTINCT	INANIMATE
	PASSING-AWAY	EXTINGUISH	BERCEUSE
	CESSATION	PLANIMETRY	PROMOTE
	INACTIVELY	DISMAST	PROMOTION
	INACTION	PLATITUDE	APPRECIATE
	BENUMB	PLENARY	CHLORAL
	PATRIARCHAL	PLOT	BEREAVEMENT
	PATRONAGE	PLUMB	BLACK-LETTER
	PATRONIZE	POACHING	ANEMIC
	PATTER	POACH	PROP
	DECOLORIZE	CHECK	PROPUGNACULUM
	ALGORISM	CHEER	PROPUGNATION
	ALGORITHM	EXTRA-JUDICIAL	PROPUGNATOR
	EXEMAT	POISE	PROPUGNER
	EXEMPTION	EYESERVANT	PROPUGNOR
	CHAFF	ACQUIT	PROPUGN
	PEDDLE	POKE	APPROXIMATION
	PEDIEUX	POLL	APRON
	PEDOMETER	POPPYCOCK	INAPTITUDE
	PEELHOUSE	POPPY-COCK	DISPENSABLE
	PEEL	CHERISH	DECOLOR
	PENUMBRA	PORTCULLIS	PROTOHISTORIC
	INACTIVE	POTHOOKS	INANITY
	INACTIVITY	POTTERING	CHLOROFORM
	PERISH	POTTER	CHOCK-FULL

Figure 2. (Continued)

DEAD, contd.	CHOKE-FULL	SAFEGUARD	SENILE
	BLANCH	SALADE	FATHOM
	QUADRANT	SALLET	SENILITY
	EYESERVICE	SALLOW	SENIORITY
	QUIBBLE	SANGAR	GEOMETRICIAN
	QUIBBLING	COADJUVANCY	INDOLENT
	ANESTHETIC	DIM	REDOUBT
	QUIESCENCE	FAINT	INCOMMENSURATE
	CHOROGRAPHER	FAIR	REDUNDANCE
	CHOROGRAPHICAL	COADJUVANT	REENFORCEMENTS
	FACE	SATIATE	REENFORCE
	QUIETUS	SATURATED	CIRCUMVALLATION
	QUITCLAIM	SATURATION	REFRIGERATION
	QUITTE	SATURITY	INCOMMUNICABLE
	DROSS	INDIGITATION	CITADEL
	BLEACH	SCARP	FAIL
	RADIOGRAPH	FAKEER	DEATH
	RADIOMETER	FAKIR	REINLESS
	RAGS	FALL	CLASHING
	RAMPART	SCHLUMMERLIED	RELAXATION
	FADDLE	SCIAGRAM	RELAXED
	RANT	SCIAGRAPH	RELAX
	BLOCKHOUSE	SCIAGRAPHY	RELEASED
	RATH	COCAINE	RELEASE
	CHOROGRAPHIC	SCONE	BARRICADE
	RAVELIN	SCORE	RELIEF
	CHOROGRAPHY	SCOT-FREE	RELIEVE
	CHOROMETRY	SCOURINGS	REMISE
	RECAPITULATE	SCRABBLE	CLEAN
	RECAPITULATION	DITCH	REMISSION
	BLOND	SCRATCH	REMISSNESS
	RECITE	SCRAWL	REMISS
	CIPHER	INDILIGENCE	CLOUD
	DEPARTED	SCRIBBLE	REMIT
	INCH	SCUTUM	RENOUNCE
	INCOMMENSURABLE	COCKSHUT	RENUNCIATION
	RECRUIT	SECONDHAND	REPEL
	RECRUITS	ATTEND	REPLETE
	REDAN	ARBITRARY	DILLY-DALLY
	FADE	SEDATIVE	REPUDIATE
	CLOY	SEGNITUDE	REREBRACE
	RUST	SEGNITY	RESCUE
	RUSTY	SELF-DEFENSE	INCOMPETENT
	FAILURE	DINGY	DROWSE
	FAINEANCE	GUARDIAN	RESTRAINT
	FAINEANT	INDOLENCE	DROWSINESS
	DILUVIAN	ARCHAIC	RHABDOLOGICAL

Figure 2. (Continued)

DEAD, contd.	RHABDOLORY	SOMBROUS	HIBERNATION
	RIGMAROLE	SOMNIFACIENT	FENCE
	HYPSOGRAPHY	SOMNIFEROUS	SHODDY
	CLOUDY	SOMNIFIC	FENDER
	ABSOLVE	SOMNOLENCE	FEND
	DROWSY	SOMNOLENT	SHORE
	RODAMONTADE	SOMNUS	SHROUD
	RODOMONTADE	ASTROLABE	COLLISION
	ROD	ADJUVANT	SIESTA
	DUGOUT	SOOTINESS	COLORLESS
	DISOBEDIENCE	SOOTY	ACHROMATIC
	RUBBISH	HYETOMETER	SIMONY
	RUBBISHY	ACHROMATISM	ANCESTRAL
	RUBBLE	COMPUTABLE	ASSESSMENT
	ARBITRARINESS	COMPUTATION	SIN
	RUDERA	GEOMETRY	ABORIGINAL
	ANEMOMETER	COMPUTE	SKIAGRAM
	RUN	SOULLESS	SKIAGRAPH
	RUSTINESS	INEFFECTUAL	SKIAGRAPHY
	ADUMBRATION	SPAN	ARITHMETICAL
	SLUGGARDIZE	SPARE	SLACKER
	SLUGGARD	DECOLOURIZE	SLACKNESS
	SLUGGISHNESS	SPEED	SLACK
	COMPASS	SPINES	SLAG
	SLUGGISH	ALLSIDED	COMATOSE
	SLUG	ASSESS	FIDDLEDEEDEE
	SLUMBERER	DISMOUNT	FIDDLE-FADDLE
	SLUMBER	FLAGGING	FIELDWORK
	SMOLDERING	FLAG	FILIBUSTERISM
	SMOLDER	CONFLICT	FILIBUSTEROUS
	GLOAMING	FAULT	FILIBUSTER
	FIRSTBORN	FAUSSEBRAIE	SLEEPFUL
	HARDENED	INEFFABLE	SLEEPING
	SMOULDERING	FAUSSEBRAYE	COMMENSURABLE
	SMOULDER	SHADE	SLEEP
	SMUGGLE	SHADING	DISPENSATION
	SMUGGLING	SHADOW	SLEEPY
	SNOOZE	SHADY	SLIGHT
	SNORE	SHAKO	SLOTHFUL
	BARRIER	HYPNOTIC	SLOTH
	COMPLEMENT	HOSTILE	SLOUCH
	DECOLOUR	HAMSTRING	FILL
	BARTIZAN	ARCHAISM	SUCCORS
	ANAESTHETIC	SHEER	SUCKLE
	SOLIDITY	SHELL	INEFFICACY
	SOLID	SHIELDING	DISABLE
	SOLLERETS	SHIELD	FOLIATE
	SOMBER	SHILLY-SHALLY	

Figure 2. (Continued)

DEAD, contd.	SUFFOCATION	TALLY	STUFF
	FOLKLORE	COORDINATES	ARITHMETIC
	FOLLY	TAME	STUN
	INEFFICIENCY	TARES	FLY
	BOLSTER	BOOTLEGGING	CONTEST
	BOMBAST	INERITA	ACCOMMODATION
	SUNDOWNER	TARNISH	CONTINGENTS
	SUNDOWN	TAUTOLOGY	STUPEFY
	INEFFICIENT	SQUARE	CONTRABAND
	SUNSET	SQUIRE	STUPOR
	SUNYASI	FLAPDOODLE	HARNESS
	HARVEST	STAGNANT	CONTRARIETY
	CONTRARIOUSNESS	STAGNATION	SUBMINISTRATION
	SUPEREROGATION	STALE	SUBSERVE
	SUPERFLUOUS	STANDARDIZE	SUBSERVIENT
	CONTRARIWISE	STANDARD	SUBSIDIARY
	HASTEN	STARK	SUBSIDY
	SUPINE	STATISTICAL	SUBTRACTION
	SUPPLEMENTAL	STATISTICIAN	SUBTRACT
	SUPPLEMENTARY	STATISTICS	SUBVENTION
	SUPPLEMENT	STATUVOLENCE	SUCCOR
	BARRACK	STATUVOLENT	TELLING
	INEPTITUDE	STATUVOLIC	TELL
	SUPPUTATION	STATUVOLISM	BECLOUD
	SUPPUTE	GEOMETER	FOREFEND
	SURVEYING	STEELYARD	HEADPIECE
	SURVEYOR	BLUNT	COUNTENANCE
	SURVEY	INEFFICACIOUS	COUNTERACTION
	INEPT	STEREOMETRY	TENAILLE
	BOOM	STERILE	TENAIL
	SUSTENTATION	STETHOSCOPE	COUNTERACT
	CONTRAVALLATION	STIFF	INERTION
	CONTRAVENE	BALDERDASH	TENEBRIOUS
	CONTRAVENTION	FLAT	TENEBROSITY
	DISQUALIFY	STILLBORN	TENEBROUS
	FOMENT	HARNESSED	THEODOLITE
	HOUGH	STOCCADO	THERMOMETER
	HAUBERK	STOCKADE	COUNTERMINE
	SWEEPING	CONSUMMATE	THICK-SKINNED
	SWEEPINGS	ADJUNCT	COUNTERPLOT
	BEAM	GREAVE	THIMBLE
	DOOM	BALISTRARIA	THOROUGHGOING
	TACIT	FLUMMERY	THOROUGH
	BEAVER	FLUXIONS	BOOTLEG
	TALE	BODYGUARD	COUNTERSCARP
	TALLIES	STRONGHOLD	FORFEND
	TALLOW-FACED	STUBBLE	INERTNESS

Figure 2. (Continued)

DEAD, contd.	COUNTERVALLATION	ARMOUR	HEAUME
	BALL-PROOF	CRIPPLE	UNCONSTITUTIONAL
	FORSWEAR	CROSS	UNDEFINABLE
	FORTALICE	TRANSGRESSION	HYPNOLOGY
	HUELESS	TRANSGRESS	UNDERCURRENT
	COUNTERWORK	TRASH	ALTIMETRY
	FORTRESS	TRASHY	CUT-THROAT
	THWART	TRENCH	UNDESPOTIC
	TIME-HONORED	CRUMBLING	HUMOR
	TIME-WORN	TRIGONOMETRICIAN	UNDIVIDEDNESS
	TIMEKILLER	TRIVIALITY	UNDIVIDED
	BASINET	TRIVIAL	FREEDOM
	T-SQUARE	INFORMALITY	UNENCUMBERED
	TOLERATE	FOSSIL	FREE
	TOLERATION	TRUISM	DABBLE
	TOPFULL	TRUMPERY	DACTYLONOMY
	TOPOGRAPHER	TRUNCHEON	UNEXERTED
	TOPOGRAPHICAL	TWADDLE	UNEXPRESSIVE
	TOPOGRAPHIC	TWADDLING	INFRINGE
	TOPOGRAPHY	TWATTLE	INFRACTION
	DEATHWATCH	TWILIGHT	ACCOUTERED
	TORPESCENCE	TYRANNY	DALLY
	TORPESCENT	UMBRAGEOUSNESS	UNFEELING
	COWCATCHER	UMBRAGEOUS	DAMMERSCHLAF
	TORPIDITY	UMBRA	UNFITNESS
	CRADLE	ACCOUTER	ANESTHESIA
	TORPID	UNACCOUNTABLE	FREEZE
	CRAMMED	FOSTER	FRIBBLE
	GAUGE	ASYLUM	FRIENDLY
	ACCOUTRED	CUBITIERE	UNFRIENDLY
	TORPOR	CUIRASS	UNFRUITFULNESS
	INERT	UNANSWERABLE	BALMY
	INEXCITABILITY	CULTIVATE	HEMIPLEGIA
	BOOTLESS	BAROMETER	DEFUNCT
	INEXPRESSIBLE	UNAUTHORIZED	ACROSS
	TOW-HAIRED	UNAVAILING	BAVARDAGE
	TOW-HEADED	UNAWAKENED	UNILLUMINATED
	FORT	CURFEW	FUDGE
	FORWARD	UNBOUND	FULL-CHARGED
	FOSSE	INFORMAL	FULL-FRAUGHT
	INEXPRESSIVE	UNCHARTERED	ASCITITIOUS
	ABDICATE	UNCOLORED	UNINFLUENTIAL
	AUDIT	FRANCHISE	ASHEN
	DULLNESS	CURTAIN	UNINTELLIGIBLENE
	TRAMP	UNCONDITIONAL	DARKENED
	DORMANCY	UNCONFORMITY	INIMICAL
	TRANCE	GAUGING	UNIVERSALITY

Figure 2. (Continued)

DEAD, contd.	UNLAWFULNESS	UPHOLD	ACHROMATIZATION
	UNLAWFUL	USELESSLY	DUN
	FULL-LADEN	USELESSNESS	DUSK
	DARKEN	INJUDICIAL	VIOLENCE
	UNMEANINGNESS	USUPRATION	GALVANOMETER
	UNMEANING	DULL	GARRISON
	UNMEANT	FUSTIAN	VISOR
	UNMITIGATED	VACANT	VIZOR
	DARKLING	VAGUENESS	WALL
	DARKLINGS	VAGUE	ACHROMATIZE
	UNNECESSARY	VAIN	WAN
	UNNEEDED	VALLATION	WARD
	DARKNESS	VALLUM	DISCOLORED
	DARK	AUSPICES	ASPHYXIA
	HINDRANCE	DUMMY	WASH
	UNOCCUPIED	DUNNESS	WASHY
	UNOFFICIAL	VAMBRACE	WASTE
	UNPRODUCTIVE	VANFOSS	BEARSKIN
	HIPPOCRATIC	VANITY	GHASTLY
	UNPROFITABLE	FUSTY	BOSH
	UNPROPITIOUS	VAUNTMURE	AMBIGUITY
	DARKSOMENESS	VEDIC	WEEDS
	UNQUALIFIED	ASH-BLOND	WELL-DISPOSED
	UNRIG	VEGETATE	BASSINET
	DARKSOME	DEADEN	DEATHBED
	UNSALABLE	ASHY	INOPERATIVE
	HELMET	VELO	DUSKY
	UNSERVICEABLE	VENERABLE	DYING
	DATELESS	FUTILE	DYKE
	DAWDLER	FUTILITY	DISCORD
	UNSKILLFULNESS	VERBIAGE	BOUNTY
	DAWDLE	VERNIER	DYNAMOMETER
	DAWDLING	VERSUS	INSENSIBILITY
	ANCILLARY	VESPER	WISH-WASH
	ABORTIVE	VESPERS	WISHY-WASHY
	ABSOLUTE	VESPERTINE	EARTHWORK
	DO-LITTLE	VETUST	DEMILUNE
	ABSOLUTION	VIAGRAPH	WORTHLESSNESS
	UNUTTERABLE	GAGE	WORTHLESS
	UNWAKENED	ACTUARY	ASSIZE
	UNWARRANTABLE	DISCOLORATION	DEMIBASTION
	UNWARRANTED	DEBRIS	YAWN
	DISCHARGE	GUARD	DISFAVOR
	UNWRITTEN		

Figure 2. (Continued)

Notice that for the entries from Webster's Dictionary of Synonyms and from Roget's International Thesaurus there are two sequences of words in alphabetical order, e.g., the entries under DEAD for Webster's Dictionary of Synonyms begin with LATE, LIFELESS, QUICK, all in alphabetical order, then revert to ALIVE, ..., LIVING, also all in alphabetical order. These lists are taken directly from the computer printouts, which list first in alphabetical order all the thesaurus words that in fact do appear in the RAND text, followed by the words from the thesaurus, again in alphabetical order, which were also submitted to the computer for the search but which did not occur in the RAND text. The nonoccurring list from Roget's University Thesaurus is not in alphabetical order but it does follow the alphabetically ordered list of words that does appear in the text.

The difference in number between the words gleaned from the two entries in Webster's Dictionary of Synonyms (DEADLY and DEAD) and those from the entries from the thesauri is obviously considerable. In fact, the total number of words in the lists under the Dictionary of Synonyms is 12; the total for Roget's International Thesaurus is 268, and, for Roget's University Thesaurus, 2452. As one might expect from these inputs, the outputs also vary considerably. They are alike only in that there are not elaborate networks of sublists. Sublists are lacking for the Synonym Dictionary and Roget's International Thesaurus runs simply because so few primary root groups are being investigated.[8]

[8] VIA operates so that when a primary root group occurs on a list associated with or linked to another primary root group, a search is made for all words associated with the second primary root group; should one of the words on that sublist also be a member of a primary root group, a search is made for all the words associated with the latter group, etc. A glance at any of the VIA output shown in last year's report shows that VIA produces rather elaborate groupings of sublists. For this experiment, however, the number of root groups in any one run (sometimes just one root group) was not sufficient to produce sublists.

For Roget's University Thesaurus, primary root groups for any given computer run were even more stringently reduced, owing to the enormous size of the word list accompanying words in the primary root group; for one computer run, for example, only words associated with the DEAD root group could be submitted. Words associated with DECLINE, etc., had to be saved for another run. The output from the DEAD root group for each of the three reference works is shown in Figure 3.

Webster's Dictionary of Synonyms

```
DEADLY
        DEADLY
        DEAD
        --------LATE
        --------LIFELESS              (Different Form Appears in the Text)
        --------QUICK
```

Roget's International Thesaurus

```
DEADLY
        DEADLINESS
        DEADLY
        --------EXTREMELY
        --------GREAT
        DEADNESS
        DEAD
        --------ANNIHILATED           (Different Form Appears in the Text)
        --------ARID                  (Different Form Appears in the Text)
        --------BARREN
        --------BLUE
        --------BODY
        --------CALLOUS               (Different Form Appears in the Text)
        --------CHARACTERLESS         (Different Form Appears in the Text)
        --------CLEAR
        --------CLOSED
        --------COLD
        --------DEADER                (Different Form Appears in the Text)
        --------DONE
        --------DULLED                (Different Form Appears in the Text)
        --------EARTH
```

Figure 3. Results of Thesaurus Run for DEAD Root Group

--------EMPTY
--------EXHAUSTED
--------EXPRESSIONLESS (Different Form Appears in the Text)
--------HEAVY
--------LATE
--------LEADED (Different Form Appears in the Text)
--------LEADEN (Different Form Appears in the Text)
--------LIFELESS (Different Form Appears in the Text)
--------NUMBED (Different Form Appears in the Text)
--------OBDURATE
--------OUT
--------PASSE (Different Form Appears in the Text)
--------PAST
--------PASSIVE
--------PEACE
--------POINTLESS (Different Form Appears in the Text)
--------REGULAR
--------REMAINS
--------REST
--------SENSELESS (Different Form Appears in the Text)
--------SLOW
--------SPIRITLESS (Different Form Appears in the Text)
--------SUPPRESSED (Different Form Appears in the Text)
--------TONELESS (Different Form Appears in the Text)
--------TOTAL
--------UTTER
--------WATERY (Different Form Appears in the Text)
--------WEAK

Roget's University Thesaurus

DEADLY

DEADLY
--------BASE
--------BATTLE
--------BLOODY
--------BLOOD
--------CALAMITY
--------COURSING (Different Form Appears in the Text)
--------DEADLY
--------DESTROY
--------DESTRUCTIVE
--------DETRIMENTAL (Different Form Appears in the Text)
--------DISASTER
--------DISASTROUS
--------EXECUTIONER (Different Form Appears in the Text)
--------EXECUTE

Figure 3. (Continued)

--------EXECUTION
--------GUNMAN (Different Form Appears in the Text)
--------HARM (Different Form Appears in the Text)
--------HARMFUL
--------KILLING
--------KILL
--------MEAN
--------NARCOTIC
--------OPPRESSIVE (Different Form Appears in the Text)
--------OPPRESS (Different Form Appears in the Text)
--------OPPRESSION
--------POISONOUS (Different Form Appears in the Text)
--------ROTTEN
--------TAKING
--------UNHEALTHINESS (Different Form Appears in the Text)
--------UNHEALTHY
--------WARFARE
--------WAR
--------WRONG (Different Form Appears in the Text)
DEAD
--------ACCESSARY (Different Form Appears in the Text)
--------ACCOUNT
--------ACTIONABLE (Different Form Appears in the Text)
--------ADD (Different Form Appears in the Text)
--------ADDITION
--------ADVANCE
--------ADVOCATE (Different Form Appears in the Text)
--------AEGIS
--------AGAINST
--------AGE
--------AIDANCE (Different Form Appears in the Text)
--------AIDING (Different Form Appears in the Text)
--------AID
--------ALLY
--------ALLNESS (Different Form Appears in the Text)
--------ALL
--------ALTOGETHER
--------ANALYTIC (Different Form Appears in the Text)
--------ANALYSIS
--------ANARCHICAL (Different Form Appears in the Text)
--------ANARCHIC (Different Form Appears in the Text)
--------ANARCHY
--------ANCIENT
--------ANTAGONISM (Different Form Appears in the Text)
--------ANTAGONIZE (Different Form Appears in the Text)
--------ANTAGONISTIC
--------ANTIQUE (Different Form Appears in the Text)

Figure 3. (Continued)

--------ANTIQUITY
--------APPRAISEMENT (Different Form Appears in the Text)
--------APPRAISE
--------ARMET (Different Form Appears in the Text)
--------ARMOR (Different Form Appears in the Text)
--------ARMED
--------ARM
--------ASSIST (Different Form Appears in the Text)
--------ASSISTANCE
--------BALANCE
--------BARREN
--------BASTION (Different Form Appears in the Text)
--------BATTLEMENT (Different Form Appears in the Text)
--------BENEFIT
--------BLACKNESS (Different Form Appears in the Text)
--------BLACK
--------BOOKED (Different Form Appears in the Text)
--------BREAK
--------CALCULATE (Different Form Appears in the Text)
--------CALCULATOR (Different Form Appears in the Text)
--------CALCULATION
--------CALLOUS (Different Form Appears in the Text)
--------CLASSIC (Different Form Appears in the Text)
--------COLD
--------COMMENSURATE
--------COMA (Different Form Appears in the Text)
--------COMPETE (Different Form Appears in the Text)
--------COMPETITIVE (Different Form Appears in the Text)
--------COMPETITION
--------COMPLETENESS (Different Form Appears in the Text)
--------COMPLETING (Different Form Appears in the Text)
--------COMPLETION (Different Form Appears in the Text)
--------COMPLETELY
--------COMPLETE
--------CONDUCE (Different Form Appears in the Text)
--------CONFRONT (Different Form Appears in the Text)
--------CONTRADICT (Different Form Appears in the Text)
--------CONTRARINESS (Different Form Appears in the Text)
--------CONTRARY
--------CONTRIBUTE
--------CONTROL
--------COUNTABLE (Different Form Appears in the Text)
--------COUNT
--------CUSTOMARY (Different Form Appears in the Text)
--------CUSTOM
--------DEAD (Different Form Appears in the Text)

Figure 3. (Continued)

--------DEADS (Different Form Appears in the Text)
--------DEADLY
--------BASE
--------BATTLE
--------BLOODY (Different Form Appears in the Text)
--------BLOOD
--------CALAMITY
--------COURSING (Different Form Appears in the Text)
--------DEADLY
--------DESTROY
--------DESTRUCTIVE
--------DETRIMENTAL (Different Form Appears in the Text)
--------DISASTER
--------DISASTROUS
--------EXECUTIONER (Different Form Appears in the Text)
--------EXECUTE
--------EXECUTION
--------GUNMAN (Different Form Appears in the Text)
--------HARM (Different Form Appears in the Text)
--------HARMFUL
--------KILLING
--------KILL
--------MEAN
--------NARCOTIC
--------OPPRESSIVE (Different Form Appears in the Text)
--------OPPRESS (Different Form Appears in the Text)
--------OPPRESSION
--------POISONOUS (Different Form Appears in the Text)
--------ROTTEN
--------TAKING
--------UNHEALTHINESS (Different Form Appears in the Text)
--------UNHEALTHY
--------WARFARE
--------WAR
--------WRONG (Different Form Appears in the Text)
--------DECLINE
--------DEFENDED (Different Form Appears in the Text)
--------DEFEND (Different Form Appears in the Text)
--------DEFENDER (Different Form Appears in the Text)
--------DEFENDING
--------DEFENSIVELY (Different Form Appears in the Text)
--------DEFENSE
--------DEFENSIVE
--------DEMONSTRATE (Different Form Appears in the Text)
--------DEPOSE (Different Form Appears in the Text)
--------DEPOSITION (Different Form Appears in the Text)
--------DIFFERENTIATION

Figure 3. (Continued)

--------DIVIDERS (Different Form Appears in the Text)
--------DIVIDE (Different Form Appears in the Text)
--------DIVISION
--------EMPTY
--------ENCOURAGE (Different Form Appears in the Text)
--------END
--------ENTERTAIN (Different Form Appears in the Text)
--------ENTIRETY (Different Form Appears in the Text)
--------ENTIRE
--------ENUMERATE (Different Form Appears in the Text)
--------ENUMERATION (Different Form Appears in the Text)
--------ESTIMATION (Different Form Appears in the Text)
--------ESTIMATE
--------EVEN (Different Form Appears in the Text)
--------EVOLUTION (Different Form Appears in the Text)
--------EXHAUSTIVE (Different Form Appears in the Text)
--------EXPEDITE (Different Form Appears in the Text)
--------FASTNESS (Different Form Appears in the Text)
--------FAVORABLE
--------FAVOR
--------FOOD
--------FORCE
--------FORTIFICATION
--------FRAUGHT
--------FRUITLESS (Different Form Appears in the Text)
--------FULL (Different Form Appears in the Text)
--------FULLY
--------FURTHERANCE (Different Form Appears in the Text)
--------FURTHER
--------GAINLESS (Different Form Appears in the Text)
--------GOOD
--------GO
--------GRADUATION (Different Form Appears in the Text)
--------GRADUATE (Different Form Appears in the Text)
--------HARD (Different Form Appears in the Text)
--------HEAVINESS (Different Form Appears in the Text)
--------HEAVY
--------HELPFUL (Different Form Appears in the Text)
--------HELP
--------HINDER
--------HOLD
--------HOUSINGS (Different Form Appears in the Text)
--------IDEAL (Different Form Appears in the Text)
--------INADEQUACY (Different Form Appears in the Text)
--------INADEQUATE
--------INDEX
--------INTEGRATION

Figure 3. (Continued)

```
--------INTEREST
--------KEEP
--------KILLINGTIME          (Different Form Appears in the Text)
--------LAGGARD              (Different Form Appears in the Text)
--------LAGGING              (Different Form Appears in the Text)
--------LAG
--------LATE
--------LAWLESS              (Different Form Appears in the Text)
--------LEADEN               (Different Form Appears in the Text)
--------LEAVINGS
--------LIFELESS            (Different Form Appears in the Text)
--------LIGHTLESS           (Different Form Appears in the Text)
--------LIMIT
--------LINE
--------LINES
--------LOAD
--------LOSS
--------LOWER
--------MAINTENANCE
--------MAN
--------MATHEMATICIAN       (Different Form Appears in the Text)
--------MEANS
--------MEASURABLE          (Different Form Appears in the Text)
--------MEASUREMENT         (Different Form Appears in the Text)
--------MEASURING           (Different Form Appears in the Text)
--------MEASURE
--------METE                (Different Form Appears in the Text)
--------METER               (Different Form Appears in the Text)
--------METRICAL            (Different Form Appears in the Text)
--------METRIC
--------METAGE              (Different Form Appears in the Text)
--------MORION              (Different Form Appears in the Text)
--------NARCOSIS            (Different Form Appears in the Text)
--------NEGLECT
--------NOTATION            (Different Form Appears in the Text)
--------NUMBERABLE          (Different Form Appears in the Text)
--------NUMBERING           (Different Form Appears in the Text)
--------NUMBER
--------OBLIGE              (Different Form Appears in the Text)
--------OBLIGING            (Different Form Appears in the Text)
--------OBSOLETE
--------OPPOSED
--------OPPOSE
--------OPPOSING
--------OPPOSITION
--------PAGE
--------PASSE               (Different Form Appears in the Text)
```

Figure 3. (Continued)

--------PAST
--------PASSIVITY (Different Form Appears in the Text)
--------PASSIVE
--------PERFECTION
--------PERFECT
--------PERMANENCE (Different Form Appears in the Text)
--------PRACTICE
--------PRESERVATION (Different Form Appears in the Text)
--------PRIME (Different Form Appears in the Text)
--------PROBE
--------PROFITLESS (Different Form Appears in the Text)
--------PROHIBITED (Different Form Appears in the Text)
--------PROOF
--------PROVE
--------PROPITIOUS
--------PROPORTION
--------PROTECTION (Different Form Appears in the Text)
--------PROTECTIVE (Different Form Appears in the Text)
--------PROTECTOR (Different Form Appears in the Text)
--------PUTTER (Different Form Appears in the Text)
--------QUICKEN (Different Form Appears in the Text)
--------RACE
--------RADICAL
--------RATE
--------RECENSION (Different Form Appears in the Text)
--------RECKONABLE (Different Form Appears in the Text)
--------RECKONING (Different Form Appears in the Text)
--------RECKON
--------REDUCTION (Different Form Appears in the Text)
--------REFUSE (Different Form Appears in the Text)
--------REGULAR
--------RESISTANT (Different Form Appears in the Text)
--------RESISTANCE
--------REST
--------RIPENESS (Different Form Appears in the Text)
--------RIVAL (Different Form Appears in the Text)
--------RIVALRY
--------ROOTED
--------RULE
--------SCALE
--------SCREEN
--------SECOND
--------SELF-PRESERVATIO
--------SENSELESS (Different Form Appears in the Text)
--------SERVE
--------SLOW
--------SOLIDARITY

Figure 3. (Continued)

--------SOUND
--------STEM
--------STEP
--------STRIKER (Different Form Appears in the Text)
--------SUFFICIENCY (Different Form Appears in the Text)
--------SUMMATION (Different Form Appears in the Text)
--------SUMMARY
--------SUM
--------SUNLESS (Different Form Appears in the Text)
--------SUPPLIES (Different Form Appears in the Text)
--------SUPPORT
--------SUSTENANCE (Different Form Appears in the Text)
--------SUSTAIN
--------TARGE (Different Form Appears in the Text)
--------TARGET
--------TEND (Different Form Appears in the Text)
--------THOUGH (Different Form Appears in the Text)
--------THROUGHOUT
--------TOTALNESS (Different Form Appears in the Text)
--------TOTALITY
--------TOTALLY
--------TOTAL
--------TRADITIONAL (Different Form Appears in the Text)
--------TRADITIONARY (Different Form Appears in the Text)
--------TRADITIVE (Different Form Appears in the Text)
--------TRADITION
--------TRUANT (Different Form Appears in the Text)
--------UNBRIDLED
--------UNFAVORABLY (Different Form Appears in the Text)
--------UNFAVORABLE
--------UNITY
--------USELESS (Different Form Appears in the Text)
--------UTTERLY (Different Form Appears in the Text)
--------VALUATION (Different Form Appears in the Text)
--------VALUELESS (Different Form Appears in the Text)
--------VALUE
--------VIOLATE
--------VIOLATION
--------WEAK
--------WHITEN (Different Form Appears in the Text)
--------WHITE
--------WHOLE
--------WHOLLY
--------WINTERLY (Different Form Appears in the Text)
--------WINTRY (Different Form Appears in the Text)
--------WINTER
--------WITHSTAND

Figure 3. (Continued)

DEALING
DEAL
DEALS
DEALT
DEAN
DECEMBER
DECISIONS
DECISIVELY
DECISIVENESS
DECISIVE
DECLARATION
DECLARATIONS
DECLARED
DECLARES
DECLARING
DECLINE
DECREE
DEEMS
DEEP
DEFEAT
DEFEATS
DEFENDING
DEFENSE-INDUSTRI
DEFENSE
DEFENSIVE
DEFINED
DEFINES
DEFINING
DEFINITE
DEFINITION
DEFINITIVE
DEFLECT
DEGREE
DELAYED
DELAYING

Figure 3. (Continued)

Notice that the output is formatted so that, in each case, the word DEADLY appears initially at the far left. This location means that DEADLY is the one word in the root group associated with DEAD that actually occurs in the text. The other words, including DEAD, located in the column just to the right of the initial DEADLY, are simply words within the root group upon which searches

for associated words will be based. DEADLY appears in this column as well because it "initiates" a search. For a discussion of the reason behind this aspect of the program, see pages 69-70 in this report.

The number of words in the output clearly reflects the disparate numbers in the input. Other than the words in the root group, the total number of words in the output from the computer run using Webster's Dictionary of Synonyms, is 3; for Roget's International Thesaurus, the total is 44; for Roget's University Theaurus, the total is 310 (the repeated 32 words associated with DEADLY are not counted in this total).

Before examining the outputs in more detail, the reason for the enormous discrepancy between the numbers of words produced by the two Roget's thesauri--one of which is said to be a version of the other--is worth pointing out. The reason lies in the methods used for indexing the two thesauri. In some cases the indices are concerned with comparable concepts. For example, in Roget's University Thesaurus, some of the entries under DEAD are as follows:

lifeless 360
inert 172
obsolete 124

In Roget's International Thesaurus, the analogous entries under DEAD are:

lifeless 407.24
inert 267.12
passé 123.15

The reference to "lifeless 360" in the University Thesaurus sends the investigator to a general heading, 360, entitled DEATH, which contains approximately two columns of synonyms and antonyms. The reference for "lifeless" in Roget's International Thesaurus is to a subsection of the general entry DEATH. This

subsection contains just 24 words suitable for inclusion in the suggested-word list (remember that multiple word idioms, such as "gone to a better land," are excluded) as opposed to some 65 words in the University Thesaurus. An examination of the other entries would show a similar pattern. Thus, although the entry under DEATH in the International Thesaurus is actually fuller than that in the University Thesaurus, the practice of indexing by subsection in the International Thesaurus considerably reduces the number of words referenced by any one index entry.

In addition to the differences between the suggested word lists caused by reference to subsection in one index and not in the other, the major headings referenced in the two thesauri under any given word are not identical. For example, the headings under DEAD from the University Thesaurus that were used for this comparison are shown in Figure 4.

dead	complete 52	(Completeness)
	obsolete 124	(Oldness)
	inert 172	(Inertness)
	colorless 429	(Achromatism)
	lifeless 360	(Death)
	insensible 376	(Physical Insensibility)
	oppose 708	(Opposition)
	sponger 886	(Servility)
	unmeaning 517	(Unmeaningness)
	useless 645	(Inutility)
	laxity 738	(Laxity)
	exempt 927	(Exemption)
	illegal 964	(Illegality)
	difficulty 707	(Aid--a thesaurus error--should be 706 Hindrance)
	midnight 126	(Evening)
	dark 421	(Darkness)
	numeration 85	(Numeration)
	measurement 466	(Measurement)
	defense 717	(Defense)
	against 14	(Contrariety)

Figure 4. Headings under DEAD, University Thesaurus

The index headings from the International Thesaurus were as follows (Figure 5):

dead		
nouns	450.1	(Silence)
adjs.	nonexistent 2.10	(Nonexistence)
	utter 56.11	(Completeness)
	passé 123.15	(Oldness)
	dead-end 265.9	(Closure)
	inert 267.12	(Quiesence)
	lackluster 336.19	(Darkness)
	lifeless 407.24	(Death)
	numb 422.5	(Insensibility)
	unconscious 422.7	(Insensibility)
	insipid 429.2	(Insipidness)
	muffled 451.17	(Faintness)
	expressionless 545.11	(Meaninglessness)
	listless 706.18	(Inactivity)
	asleep 710.20	(Sleep)
	tired out 715.8	(Fatigue)
	dull 881.6	(Dullness)
advs.	absolutely 34.21	(Greatness)
	directly 289.25	(Direction)
	exactly 515.19	(Truth)

Figure 5. Headings under DEAD, International Thesaurus

The words in parentheses in Figures 4 and 5 are the names of the main thesaurus headings to which the index entries refer. Of these main headings, only four are common to both lists; the four common headings are Completeness, Oldness, Death, and Darkness. A number of other headings are similar. Inertness in the University Thesaurus is not unlike Quiescence in the International Thesaurus; other somewhat parallel headings include Physical Insensibility in the University Thesaurus (U.T.), and Insensibility in the International Thesaurus (I.T.), as well as Unmeaningness and Meaninglessness, and perhaps Laxity and Inactivity. Other headings are not so easily matched up. Achromatism (U.T.) may have something to do with Darkness (both thesauri) or Insipidness (I.T.), and Evening (U.T.) is probably related to Darkness (both thesauri); nonexistence

(I.T.) is also connected with Death (both thesauri). There do not seem to be any very good matches, though, for Opposition, Servility, Inutility, Exemption, Illegality, Hindrance, Numeration, Measurement, and Defense, all in the University Thesaurus, or, on the other hand, for Silence, Closure, Faintness, Sleep, Dullness, Fatigue, Greatness, Direction, and Truth in the International Thesaurus. A relaxation of our selection rule to permit the use of words preceded by dashes in the University Thesaurus index (dashes indicate compounds or multiple word entries) would have produced an entry for -Dead Silence (Silence occurs in the International Thesaurus) and for -Dead Asleep (Asleep occurs in the International Thesaurus). Still, there are a number of emphases that distinguish one index from the other. For instance, the stress on conflict verging over into militancy in the University Thesaurus (Opposition, Hindrance, and Defense) is not found in the International Thesaurus, nor is the emphasis upon precise enumeration (Numeration, Measurement). In fact, these headings, in addition to others such as Illegality and Servility, all suggest connections with human manipulativeness--an emphasis lacking in the more abstract entries (Greatness, Direction, Truth, Dullness) used in the International Thesaurus.

Having in mind the general distinctions between the indices used for DEAD in the two thesauri, the next step is to turn to the computer outputs showing the results of the textual searches. Which reference works produced the most meaningful and relevant lists of associated words? A glance at the outputs in Figure 3 shows that two classes of words appear in the lists--those that occur in the text in exactly the form appearing in the list, and those that

occur in a different form.[9] For assessment of the relevance of the outputs, the words on the lists were separated into these two groups and each group was judged separately.

The output from the run using input from Webster's Dictionary of Synonyms is so modest as to require little attention. Of the three words linked to DEAD, LATE and QUICK occur in the text as shown. LATE is clearly a synonym and QUICK an antonym for DEAD. Although LIFELESS appears in the list, the warning "Different Form Appears in the Text" sends us to an index for chapter one of Soviet Military Strategy where we discover that the form actually occurring in the text is an antonym for DEAD--LIFE. Since one of the principles of the VIA research has been to include antonyms as well as synonyms on suggested word lists, the three words linked to DEAD may all be considered relevant. The reasons for including antonyms on such lists have been formally justified by linguists interested in "contrast" methods. The reasons for including them in our lists are: (1) When a given concept has structural importance in a portion of text, the importance is often more clearly understood if the relative strength of its opposite concept is also revealed--for example, the emphasis upon nuclear war vis-à-vis conventional war. (2) The presence of one concept in the mind of a writer or speaker often seems to imply the presence of the concept's opposite as well--for example, when Hamlet speculates as to whether it would be better to live or die, "to be or not to be." For more subtle analyses of motive, word choices revealing opposite moods can lead to further

[9] For a complete discussion of this search procedure, see pages 69-70 of this report.

exploration of the verbal interaction between these moods and thus to more informed estimates of a preference for one mode of action as opposed to another. The above list from Webster's Dictionary of Synonyms is too modest to afford any basis for more than a casual impression that it could be interesting to submit a much more extensive list of words that might be associated with DEAD as well as a list that might be associated with LIFE. Varying emphases upon these concepts from one section of Soviet Military Strategy to another might well be revelatory of shifting attitudes toward particular military strategies.

The outputs from the computer runs using the thesauri are considerably more complex. To reduce this discussion to manageable proportions, only those words linked to DEAD (as opposed to DEADLY, DEADEN, etc.) will be discussed. For Roget's International Thesaurus, 25 words from the suggested word list occur in a form identical to that in the text. These words are (Figure 6):

Barren	(Dullness)
Blue	(Completeness)
Body	(Death)
Clear	(Completeness)
Closed	(Closure)
Cold	(Darkness--Insensibility--Dullness)
Done	(Fatigue)
Earth	(Death)
Empty	(Meaninglessness)
Exhausted	(Fatigue)
Heavy	(Inactivity--Quiescence--Dullness)
Late	(Death)
Obdurate	(Insensibility)
Obsolete	(Oldness)
Out	(Insensibility)
Past	(Oldness)
Passive	(Quiescence)
Peace	(Silence)
Regular	(Completeness)

Figure 6. Words Identical in Form in Text and Suggested Word List, International Thesaurus

Remains	(Death)
Rest	(Silence)
Slow	(Dullness)
Total	(Completeness)
Utter	(Completeness)
Weak	(Insipidness)

Figure 6. (Continued)

The words within parentheses are the headings in the thesaurus under which the words occur. In the judgment of this investigator, the words on the list seeming to have the most questionable relevance are Clear and Regular. Obdurate also seems somewhat tangential. Therefore, on the above list, 22 or 23 words seem relevant, meaningful output. Clear and Regular are both introduced through the heading, Completeness. But if that entry were removed from the index, other words that seem relevant (Blue, Total, and Utter) would be lost. Obdurate is selected through Insensibility, but so, too, is Out.

Eighteen words bearing the "Different Form Occurs in Text" message are linked to DEAD in the output from the International Thesaurus. Of these 18 words, only 14 should be considered here. The other four--a very atypically high number--were mislinked by SUFFIX and should not have appeared in the list. One of the four, DULLED, appears because the SUFFIX program took it to be of the same root as DULLES, which does appear; the other three (ARID with Arise, CALLOUS with Call, and NUMBED with Number) are genuine mislinks and have since been entered in the appropriate exception list for SUFFIX's operation. The 14 words properly appearing on the list were (Figure 7):

Word	Form Actually Occurring in Text	Thesaurus Heading
ANNIHILATED	Annihilate, Annihilation	(Nonexistence)
CHARACTERLESS	Character, Characteristic, Characteristics, Characterized	(Dullness)
DEADER	Deadly	(Death)
EXPRESSIONLESS	Expression, Expressing, Expressed, Expresses	(Meaninglessness)
LEADED	Lead, Leads, Leading, Leader, Leaders, Leadership	(Quiescence)
LEADEN	(same as above)	(Quiescence)
LIFELESS	Life	(Death, Quiescence, Darkness, Inactivity)
PASSE	Passed, Past	(Oldness)
POINTLESS	Point, Pointed, Pointing	(Dullness)
SENSELESS	Sense	(Insensibility)
SPIRITLESS	Spirit, Spiritually, Spiritual	(Dullness)
SUPPRESSED	Suppress	(Faintness)
TONELESS	Tone	(Faintness)
WATERY	Water, Waters	(Insipidness)

Figure 7. Words of the Same Root but not Identical Form, International Thesaurus

Of the two words on this list, those that are most suspect when compared with the forms actually occurring in the text are LEADED and LEADEN.[10] LEADER, LEADERS, and LEADERSHIP are neither synonyms nor antonyms for the LEADED and LEADEN associated with DEAD and, given this usage context, one might suppose that LEAD, LEADS, and LEADING are not associated with DEAD either. A glance at the use of the latter words in the text shows that they are used in the sense of guide or direct; they are not describing a condition associated with death. The other words on the list seem acceptable either as antonyms or

[10] The current VIA program does not print out the form or forms actually occurring in the text, but the rewritten program will include those forms. For a more complete discussion of the philosophy underlying this procedure, see pages 90-91 of this report.

synonyms for words actually occurring in the text. WATERY may seem more dubious than the other words, but water is often metaphorically associated with death. WATERY is introduced through Insipidness, to which the acceptable WEAK on the first list is also owing.

For both lists from the computer run using the International Thesaurus, the score for relevance seems rather high. Of the 39 words considered, only 4 seem definitely out of place or misleading and 2 others seem marginally relevant.

The output from the University Thesaurus run is considerably more extensive and complicated than that from the International Thesaurus. The words among those on the suggested list that occurred in identical form in the text were as follows (Figure 8):

ACCOUNT	(Numeration)	BREAK	(Exempt)
ADDITION	(Numeration)	CALCULATION	(Numeration)
ADVANCE	(Aid)	COLD	(Achromatism)
AEGIS	(Defense)	COMMENSURATE	(Numeration)
AGAINST	(Opposition)	COMPETITION	(Opposition)
AGE	(Oldness)	COMPLETELY	(Completeness)
AID	(Aid)	COMPLETE	(Completeness)
ALLY	(Aid)	CONTRARY	(Contrariety)
ALL	(Completeness)	CONTRIBUTE	(Aid)
ALTOGETHER	(Completeness)	CONTROL	(Opposition)
ANALYSIS	(Numeration)	COUNT	(Numeration)
ANARCHY	(Laxity)	CUSTOM	(Oldness)
ANCIENT	(Oldness)	DECLINE	(Oldness)
ANTAGONISTIC	(Contrariety)	DEFENDING	(Defense)
ANTIQUITY	(Oldness)	DEFENSE	(Defense)
APPRAISE	(Numeration)	DEFENSIVE	(Defense)
ARMED	(Defense)	DIFFERENTIATION	(Numeration)
ARM	(Defense)	DIVISION	(Numeration)
ASSISTANCE	(Aid)	EMPTY	(Inutility)
BALANCE	(Numeration)	END	(Death)
BARREN	(Inutility)	ENTIRE	(Completeness)
BENEFIT	(Aid)	ESTIMATE	(Measurement)
BLACK	(Darkness)	FAVORABLE	(Aid)

Figure 8. Words Identical in Form in Text and Suggested Word List, University Thesaurus

FAVOR	(Aid)
FOOD	(Aid)
FORCE	(Aid)
FORTIFICATION	(Defense)
FRAUGHT	(Completeness)
FULLY	(Completeness)
FURTHER	(Aid)
GOOD	(Completeness)
GO	(Achromatism)
HEAVY	(Inertness)
HELP	(Aid)
HINDER	(Defense)
HOLD	(Defense)
INADEQUATE	(Inutility)
INDEX	(Measurement)
INTEGRATION	(Completeness)
INTEREST	(Aid)
KEEP	(Defense)
LAG	(Inactivity)
LATE	(Death)
LEAVING	(Inutility)
LIMIT	(Completeness)
LINE	(Measurement)
LINES	(Measurement--Defense)
LOAD	(Completeness)
LOSS	(Death)
LOWER	(Darkness)
MAINTENANCE	(Aid)
MAN	(Defense)
MEANS	(Aid)
MEASURE	(Numeration--Measurement)
METRIC	(Measurement)
NEGLECT	(Exempt)
NUMBER	(Numeration)
OBSOLETE	(Inutility)
OPPOSED	(Contrariety--Opposition)
OPPOSE	(Contrariety--Opposition)
OPPOSING	(Contrariety--Opposition)
OPPOSITION	(Contrariety--Opposition)
PAGE	(Numeration)
PAST	(Oldness)
PASSIVE	(Inertness)
PERFECTION	(Completeness)
PERFECT	(Completeness)
PRACTICE	(Numeration)
PROBE	(Measurement)
PROOF	(Physical Insensibility)
PROVE	(Numeration)
PROPITIOUS	(Aid)
PROPORTION	(Numeration)
RACE	(Opposition)
RADICAL	(Completeness)
RATE	(Measurement)
RECKON	(Numeration)
REGULAR	(Completeness)
RESISTANCE	(Defense)
REST	(Death)
RIVALRY	(Opposition)
ROOTED	(Oldness)
RULE	(Measurement)
SCALE	(Measurement)
SCREEN	(Defense)
SECOND	(Aid)
SELF-PRESERVATION	(Defense)
SERVE	(Servility)
SLOW	(Inertness)
SOLIDARITY	(Completeness)
SOUND	(Measurement)
STEM	(Opposition)
STEP	(Measurement)
SUMMARY	(Numeration)
SUM	(Numeration)
SUPPORT	(Aid)
SUSTAIN	(Aid)
TARGET	(Defense)
THROUGHOUT	(Completeness)
TOTALITY	(Completeness)
TOTALLY	(Completeness)
TOTAL	(Completeness--Numeration)
TRADITION	(Oldness)
UNBRIDLED	(Laxity)
UNFAVORABLE	(Opposition)
UNITY	(Completeness)
VALUE	(Measurement)
VIOLATE	(Exempt)
VIOLATION	(Exempt)

Figure 8. (Continued)

WEAK	(Laxity)	WHOLLY	(Completeness)
WHITE	(Achromatism)	WINTER	(Evening)
WHOLE	(Completeness)	WITHSTAND	(Opposition)

Figure 8. (Continued)

Of the 138 words on this list, 19 were introduced through the thesaurus heading Aid, and therefore should be ignored.[11] Of the remaining 19 words, a great many seem either partially or totally irrelevant to the concept DEAD. Among those seeming totally irrelevant are such words as GO, METRIC, PAGE, ANALYSIS, and so on up to the approximate total of 38. The word "approximate" is used because judgments of relevance are, of course, subjective as to both person and time. My own selection of words on this list that seemed totally irrelevant varied by a total of two or three words on each occasion that I searched the list; 38 represents the average for the searches.[12] Thesaurus headings most productive of words judged irrelevant were Numeration and Measurement. Through Numeration, the words ADDITION, ANALYSIS, APPRAISE, CALCULATION, COMMENSURATE, COUNT, DIFFERENTIATION, PAGE, PRACTICE, PROVE, and PROPORTION were selected. Measurement accounted for ESTIMATE, INDEX, LINE, LINES, METRIC, PROBE, SOUND, STEP, and VALUE. Thus, these two headings accounted for more than half the words that seem not to contribute meaningfully

[11] See comment in Figure 4 re this error in the thesaurus.

[12] Although it would be enormously useful to have precise measures of relevance, such measures are not yet forthcoming. All that any investigator does, finally, is sort words on the basis of a collection of a number of different judgments as to how the words are used. In recent years, some judgments have been collected in a controlled experimental situation involving a number of people working with the same words at the same moments in time. Whether, for many purposes, this mode of collecting judgments is more useful than are observations of usage collected over a period of time from more widely varied sources is open to question. A combination of the two approaches may prove most fruitful.

to the list. The heading Completeness accounted for the next largest number of words--ALTOGETHER, PERFECTION, PERFECT, REGULAR, SOLIDARITY, and UNITY.

In addition to the words seeming totally irrelevant, approximately 20 others seem somewhat questionable. Among these words, the headings Numeration and Measurement account for the following: DIVISION, MEASURE, NUMBER, RATE, RECKON, RULE, and SCALE. The heading Defense is responsible for HINDER, HOLD, SCREEN, and TARGET; Completeness accounts for INTEGRATION, LIMIT, RADICAL, and THROUGHOUT; Opposition introduces COMPETITION, CONTROL, and RACE. The two other words in this group were BREAK, from Exempt, and TRADITION, from Oldness.

Among the words remaining on the list, many of those introduced through the heading Completeness--such as WHOLE, WHOLLY, TOTALITY, TOTAL, etc.--seem appropriate as synonyms for the aspect of finality associated with deadness. Another group of words, those selected through DEFENSE and, in some cases, OPPOSITION, are acceptable only because they imply reaction to death or to forces associated with death. In a complete VIA run for chapter one of Soviet Military Strategy, the words ARMED and ARM would link words associated with DEAD to concepts associated with WAR and MILITARY--a desirable linkage for that text.

Among the words left on the list that seem acceptable, only four were selected through the heading Numeration; no words judged thoroughly acceptable were selected through the heading Measurement. Of the four--ACCOUNT, BALANCE, SUMMARY, and SUM--from Numeration, the sense of SUMMARY and SUM is implied by other words, such as TOTAL and TOTALITY, which are on the list; since ACCOUNT,

as associated with DEAD, implies some sort of final accounting, its sense, too, may be represented by a word such as TOTAL. BALANCE, in the sense of a balancing out of accounts, can obviously also be gleaned from other words, but the sense of a balance between death and life is not so easily inferred from other words. Nonetheless, by any accounting, the headings Numeration and Measurement would seem expendable. Their absence from the index for Roget's International Thesaurus helps account for the much higher ratio of relevance in its output.

Before considering further the desirability of deletions from or additions to index headings for either of the thesauri, it might be well to look at the second portion of the output from the run using Roget's University Thesaurus--the portion showing words for which a different form appears in the text:

Word	Form Actually Occurring in Text	Thesaurus Headings
ACCESSARY	Access	(Aid)
ACTIONABLE	Act, Action, Actions, Acted	(Illegality)
ADD	Addition	(Numeration)
ADVOCATE	Advocating	(Aid)
AIDANCE	Aid	(Aid)
AIDING	Aid	(Aid)
ALLNESS	All	(Completeness)
ANALYTIC	Analysis	(Numeration)
ANARCHICAL	Anarchy	(Laxity)
ANARCHIC	Anarchy	(Laxity)
ANTAGONISM	Antagonistic	(Contrariety)
ANTAGONIZE	Antagonistic	(Contrariety)
ANTIQUE	Antiquity	(Oldness)
APPRAISEMENT	Appraise	(Numeration)
ARMET	Armed, Arm	(Defense)
ARMOR	Armed, Arm	(Defense)
ASSIST	Assistance	(Aid)
BASTION	Base	(Defense)
BATTLEMENT	Battle, Battles, Battlefield(s)	(Defense)
BLACKNESS	Black	(Darkness)
BOOKED	Book	(Death)
CALCULATE	Calculation	(Numeration)
CALCULATOR	Calculation	(Numeration)

Figure 9. Words of the Same Root but not Identical Form, University Thesaurus

Word	Form Actually Occurring in Text	Thesaurus Headings
CLASSIC	Classification	(Oldness)
COMPETE	Competition	(Opposition)
COMPETITIVE	Competition	(Opposition)
COMPLETENESS	Completely	(Completeness)
COMPLETING	Completely	(Completeness)
COMPLETION	Completely	(Completeness)
CONDUCE	Conduct	(Aid)
CONFRONT	Confronted	(Opposition)
CONTRADICT	Contradictions	(Opposition)
CONTRARINESS	Contrary	(Contrariety)
COUNTABLE	Count	(Numeration)
CUSTOMARY	Custom	(Oldness)
DEAD	Deadly	(Death)
DEFENDED	Defending	(Defense)
DEFENDER	Defending	(Defense)
DEFENSIVELY	Defense, Defensive	(Defense)
DEMONSTRATE	Demonstrated, Demonstration	(Numeration)
DEPOSE	Deposits	(Laxity)
DEPOSITION	Deposits	(Laxity)
DIVIDERS	Division	(Numeration)
DIVIDE	Division	(Numeration)
ENCOURAGE	Encouraged, Encouraging	(Aid)
ENTERTAIN	Entertained	(Aid)
ENTIRETY	Entire	(Completeness)
ENUMERATE	Enumeration	(Numeration)
ENUMERATION	Enumerated	(Numeration)
EVOLUTION	Evolutionary	(Numeration)
EXHAUSTIVE	Exhausted, Exhaustion	(Completeness)
EXPEDITE	Expeditions	(Aid)
FASTNESS	Fast	(Defense)
FRUITLESS	Fruit(s)	(Inutility)
FULL	Fully	(Completeness)
FURTHERANCE	Further	(Aid)
GAINLESS	Gains, Gaining	(Inutility)
GRADUATION	Graduated	(Measurement)
GRADUATE	Graduated	(Measurement)
HARD	Hardly	(Physical Insensibility)
HEAVINESS	Heavy	(Inertness)
HELPFUL	Help	(Aid)
HOUSINGS	House	(Defense)
IDEAL	Idealistic	(Completeness)
INADEQUACY	Inadequate	(Inutility)
KILLINGTIME	Killing	(Evening)
LAGGARD	Lag	(Inactivity)
LAGGING	Lag	(Inactivity)
LAWLESS	Law(s)	(Illegality)

Figure 9. (Continued)

Word	Form Actually Occurring in Text	Thesaurus Headings
LEADEN	Lead(s), Leader(s), Leading, Leadership	(Achromatism)
LIFELESS	Life	(Inertness)
LIGHTLESS	Light	(Darkness)
MATHEMATICIAN	Mathematics	(Numeration)
MEASURABLE	Measure	(Numeration--Measurement)
MEASUREMENT	Measure	(Numeration--Measurement)
MEASURING	Measure	(Numeration--Measurement)
METER	Metric	(Measurement)
METRICAL	Metric	(Measurement)
METAGE	Metric	(Measurement)
NARCOSIS	Narcotic	(Physical Insensibility)
NOTATION	Noted	(Numeration)
NUMBERABLE	Number	(Numeration)
NUMBERING	Number	(Numeration)
OBLIGE	Obligation	(Aid)
OBLIGING	Obligation	(Aid)
PASSE	Past	(Oldness)
PASSIVITY	Passive	(Inertness)
PERMANENCE	Permanently	(Inertness)
PRIME	Primary, Primarily	(Oldness)
PROFITLESS	Profits, Profitable	(Inutility)
PROHIBITED	Prohibit	(Illegality)
PROTECTION	Protect	(Defense)
PROTECTIVE	Protect	(Defense)
PROTECTOR	Protect	(Defense)
QUICKEN	Quick	(Aid)
RECENSION	Recent, Recently	(Numeration)
RECKONABLE	Reckon	(Numeration)
RECKONING	Reckon	(Numeration)
REFUSE	Refused, Refusing	(Inutility)
RESISTANT	Resistance	(Defense)
RIPENESS	Ripe	(Completeness)
RIVAL	Rivalry	(Opposition)
SENSELESS	Sense	(Physical Insensibility)
STRIKER	Strike(s), Striking	(Exemption)
SUFFICIENCY	Sufficient	(Completeness)
SUMMATION	Summary	(Numeration)
SUNLESS	Sun	(Darkness)
SUPPLIES	Suppliers, Supply	(Aid)
SUSTENANCE	Sustain	(Aid)
TARGE	Target	(Defense)
TEND	Tendency, Tendencies	(Aid)
TOTALNESS	Total	(Completeness)
TRADITIONAL	Tradition	(Oldness)

Figure 9. (Continued)

Word	Form Actually Occurring in Text	Thesaurus Headings
TRADITIONARY	Tradition	(Oldness)
TRADITIVE	Tradition	(Oldness)
UNFAVORABLY	Unfavorable	(Opposition)
USELESS	Use(s), Used, Using	(Inutility)
UTTERLY	Utter	(Completeness)
VALUATION	Value	(Measurement)
VALUELESS	Value	(Measurement)
WHITEN	White	(Achromatism)
WINTERLY	Winter	(Evening)
WINTRY	Winter	(Evening)

Figure 9. (Continued)

Of the 123 words on this list, 17 can again be disregarded because they occurred under the thesaurus heading, Aid. Among the 110 remaining words, it is quite clear that the thesaurus headings, Measurement and Numeration, are again responsible for a number of irrelevant entries. In fact, in some cases, the irrelevance from the earlier list has been considerably amplified because of the repetition of a number of words in the same root group.[13] Thus METER, METRICAL, and METAGE all appear on this list, whereas METRIC was the only word in this root group actually appearing in the text. If Measurement and Numeration were removed from the index, a total of 30 words would disappear from this list. Of the other 21 words on the list that seem irrelevant in some degree, the 3 owing to the thesaurus heading Illegality--ACTIONABLE, LAWLESS, and PROHIBITED--are all dubious entries. The thesaurus entry, Defense, again accounts for some unacceptable words: ARMET, BASTION, HOUSINGS, and TARGE; Inutility contributes two: INADEQUACY and PROFITLESS; Oldness is

[13] See discussion on pages 64-65.

responsible for four: CLASSIC, TRADITIONAL, TRADITIONARY, and TRADITIVE--the latter three are again an amplification of the one word, TRADITION, which does occur in the text.

The relevance score for the words retrieved on the computer run using Roget's International Thesaurus is clearly much higher than that for the University Thesaurus. To recapitulate, the total for the International Thesaurus (counting both lists) was 39; of this number, 4 words definitely seemed out of place and 2 seemed marginally relevant. The total for the University Thesaurus was 229 (excluding those words introduced through AID. Had the index correctly referred to Hindrance, the total would have been larger). Of the total number, 110 words seemed irrelevant to some degree. For the University Thesaurus, just under half of the total number of entries are questionable. For the International Thesaurus, 6 of 39, or two-thirteenths, are questionable. If the words accounted for by the headings Measurement and Numeration in the University Thesaurus were removed, the total number of words would be reduced to 168, while the total of questionable words would be cut to 51. The resulting unacceptability ratio of less than one-third would be a considerable improvement upon the earlier ratio of approximately one-half. The ratio could be improved a slight degree by removing the index entry Illegality, which introduced three words, all of which seemed inappropriate. The resulting ratio would then be 48 to 165. Further improvement of the ratio for the University Thesaurus would be difficult, but a reduction in false amplification in the output can be achieved by a change in the program, which will be implemented during its rewriting. The change will be made in the operation of

SUFFIX for words appearing in thesaurus or synonym dictionary entries. If one word, such as TRADITION, occurs in both a thesaurus and the text, that word will be printed in the output; then, if other words in that root group, such as TRADITIONAL, TRADITIONARY, and TRADITIVE, occur in the thesaurus but not in the text, they will not be included in the output. SUFFIX will continue to operate so as to reveal continuity of root group when no exact match is found; for example, KILLINGTIME is included in the output although the thesaurus entry is Killing. Implementation of this change would remove 36 more words from the University Thesaurus computer run output, resulting in a total of 129. Of the 36 words removed, 11 are on the unacceptable list. Thus the new ratio of unacceptable to acceptable words would be 37 to 129--a ratio about midway between the one-quarter and one-third mark.

Oddly, considering the shortness of the output from the run using the International Thesaurus relative to that from the University Thesaurus, 15 of the 39 words or root groups in the output using the former do not appear in the output for which the University Thesaurus was employed. Those words are: BLUE, BODY, CLEAR, CLOSED, DONE, EARTH, OBDURATE, OUT, PEACE, REMAINS, ANNIHILATED, CHARACTERLESS, EXPRESSIONLESS, SUPPRESSED, TONELESS. Bearing in mind these words, it is worth noticing once more that the thesaurus headings Defense and Opposition in the University Thesaurus were responsible for a number of irrelevant words as well as for words that seemed appropriate. The words that seemed appropriate from these headings, in addition to those introduced by a third heading, Contrariety, were those implying opposing forces, often involving violent action. Although the list under DEAD from the

International Thesaurus does not contain a heavy representation of such words, PEACE, ANNIHILATED, and possibly OBDURATE, which seemed irrelevant in the earlier assessment in this report, do carry such connotations. Furthermore, in a complete VIA run, PEACE would very probably be linked, antonymously, to WAR, which in turn would supply, in sublists under DEAD, the militant concepts supplied by DEFENSE and in some measure by OPPOSITION in the output for the University Thesaurus run. Notions associated with Oldness--OBSOLETE, PAST, and PASSE--do appear in the International Thesaurus, as do those associated with Completeness (e.g., TOTAL, UTTER).

On the whole, then, this survey of words associated with DEAD in the outputs from the three reference works would suggest that Roget's International Thesaurus would be preferable to the other two, if each were being consulted in a literal, exhaustive search by a computer. The output using Webster's Dictionary of Synonyms is clearly too slight to be very informative. The output using Roget's University Thesaurus contains too much semantic "noise"; alterations in the index and SUFFIX program operation would reduce that noise somewhat but not to an acceptable level for many purposes. The output using Roget's International Thesaurus is perhaps a little spare, but relatively free of irrelevant terms. Because of the way VIA operates, spareness in any given entry is not serious if that entry is part of a larger conceptual structure in the text being examined; VIA's system of cross-linking lists of words will provide the enrichment that a wider indexing net such as that in the University Thesaurus furnishes while, at the same time, showing clearly the provenance of the enrichment.

A similar study of words associated with DECLINE, in outputs using the three references, supports the conclusions above. The run using Webster's Dictionary of Synonyms produced a list of eight words; six of the eight were antonyms. The computer run using the International Thesaurus resulted in a total of 28 words, one of which seemed questionable. Again, the University Thesaurus produced a much more extensive list of words, 96, as well as a higher number that seemed irrelevant, 21. The outputs for DECLINE are shown in the Appendix.

3. Relevance of Context. The preceding comparison of reference works has ignored the question of the context in which the words appearing in the output occurred in the text. Discussions of context were avoided, first, because they would have obscured the basic purpose of the foregoing discussion--to compare the three reference works; and, second, because the analysis of style is concerned with subtle cues as to emphasis and tone in any set of words. Modern psychology and personality studies as well as contemporary techniques of literary criticism have amply demonstrated that word choice in any given context may be governed by more overriding considerations than those supplied by the immediate context. Words such as UTTER and EMPTY, which are clearly ambiguous, may indeed carry meaningful connotations of a number of different denotations even though context might seem to confine them to one denotation. Such ambiguities are of particular significance and, very probably, of higher frequency in words associated with common human experiences, such as death and birth, which have been generalized to describe many perceptions and

responses in addition to the literal events known as death and birth. As it happens, the context for the word UTTER in chapter one of Soviet Military Strategy is concerned with literal death; the phrase is "utter mutual annihilation." But whenever UTTER is used, even as a verb, there may well be connotations of the completeness and finality associated with death; as a verb, UTTER refers to spending, using up, completing utterances, and, as an adjective, the sense of completeness, of being finished, is always there. Words such as MILITARY or AMERICA or RUSSIA tend to have more specific connotations, as do many of the words clustered around them. One of the interesting aspects of VIA's output is that the investigator can often see the relationship between specifically focused content and more emotionally focused content; for example, the word ARMED associated with MILITARY, might, through a series of cross-links, bring lower levels of sublists associated with DEAD into a cluster of words heavily concerned with the details of military organization. Given this lead, the investigator might then use a MAPTEXT representation to explore spatial relationships between the two concepts.

It is in the areas having to do with specific denotation of aspects of contemporary events or areas of knowledge that thesauri and synonym dictionaries are known to be weak, and for which, therefore, context may be more valuable. It is possible that, while reference works are being expanded to provide a more adequate verbal key to new medicines, new weapons, etc., computer search programs can be designed to turn to context when reference works fail. For example, if the investigator wished to run a search keyed on ICBM, and no such

entry appeared in the reference work, the search would immediately explore any contexts in which ICBM occurred and then return to the reference work with more information. In the meantime, the human being could intervene in such situations by asking for a concordance output for ICBM and submitting words from that context for further computer search, as well as using such words, when possible, as alternate entries to the reference work.

4. Suffix Root Retrieval Feature. More than a year ago, the THESAUR portion of the VIA program was altered to incorporate the operation of SUFFIX during searches of verbal text. This change was introduced because theme continuities were being lost. If the word on the search list were WARFARE and the word in the text were WAR, the earlier search program would not have matched WAR and WARFARE; thus, neither WAR nor the many words associated with it would have appeared in the output. The solutions to this difficulty were either to try to put all possible forms of a given root group in the search list or to let the computer check for common roots. The former solution seemed both tedious and full of the risk of omission; therefore, SUFFIX, the root-matching program used in an earlier stage of VIA, was reintroduced during the program stage actually concerned with constructing the linked output lists.

Although the use of SUFFIX still results in some mislinks, the incidence of those mislinks is not high and will decrease as the mislinks are detected and the appropriate words added to the SUFFIX exception list. Of the four mislinks occurring during the computer run using Roget's International Thesaurus, only one, the linking of DULLED with the proper name DULLES, might possibly

occur again. The others will be entered in the exception list. Of the 7 mislinks (in a total of 123 links) occurring during the computer run using Roget's University Thesaurus, all can be prevented from happening in the future. In any event, mislinks that do occur will be immediately apparent in the rewritten version of VIA because the form(s) actually occuring in the text will be output along with the form(s) occurring in the search list. We initially considered the option of simply printing out the words appearing in the text, but decided that the form on the search list often shows the conceptual link with the main word, e.g., DEAD, more clearly. On the other hand, the addition in the output of forms actually occurring will guard against misleading indications of putative or supposed relevance.

The other problem posed by the use of SUFFIX at this stage of VIA was made apparent by the run using Roget's University Thesaurus, which often provides multiple forms for any given root group. This problem, already discussed on pages 64-65 of this report, was the amplification of either relevant or irrelevant root groups because of the multiplicity of forms within root groups. Although the University Thesaurus indexing procedure does not seem the most promising for VIA runs, and the problem of amplification is not serious with the other reference works used, the program change described on pages 64-65 will be implemented so as to reduce possibilities of misleading amplification.

B. MAPTEXT as a Delineator of Style

MAPTEXT is a program that provides abstract, i.e., nonverbal, representations of a text. For analysis of a given set of data, the researcher may specify

particular words, types of words, punctuation marks, etc., that he would like to have displayed, in terms of relative position within the data. The representation is nonverbal in order to prevent semantic distractions.

MAPTEXT can serve a number of different functions. For example, using VIA output as its input, MAPTEXT can reveal gross distribution patterns that are much less obvious in index entries for individual words than when they are graphically displayed on the page. Thus, MAPTEXT can quickly direct the information analyst to the section or sections of the data that might warrant more extensive, detailed analysis. The condensed outputs produced by MAPTEXT may be used for quick comparisons of one distribution pattern with others in order to reveal variations or continuities of style.

As noted in last year's annual report, VIA revealed ambivalent attitudes toward nuclear and conventional warfare in the first chapter of Soviet Military Strategy. Although both methods of warfare were, of course, discussed, an emphasis of favor seemed to fall upon nonnuclear warfare. Subsequent to last year's report, in an effort to see how words describing nuclear and conventional warfare were distributed throughout the text as a whole, a MAPTEXT representation of terms associated with the two forms of warfare was obtained for all chapters of both translations of Soviet Military Strategy. Words appearing in the MAPTEXT output associated with nuclear warfare were represented by an A. The words represented by A were:

AIR-TO-AIR
AIR-TO-GROUND
AIR-TO-SURFACE
ANTIATOMIC
ANTIMISSILE
ANTINUCLEAR
ANTIROCKET
ATLAS
ATOMIC
ATOM

BRINKSMANSHIP
DECONTAMINATION
EARLY-WARNING
FALLOUT
FISSIONABLE
FISSION
FUSION
HYDROGEN
ICBM
ICBMS
IRBM
IRBMS
LAUNCH
LAUNCHED
LAUNCHING
LAUNCHINGS
MEGATON
MISSILE-CARRYING
MISSILE-LAUNCHING
MISSILE
MISSILES'
MISSILES
NUCLEAR-MISSILE
NUCLEAR-ROCKET
PLUTONIUM
POLARIS
POLARIS-CARRYING
RADIATION
ROCKET
ROCKETS
THERMONUCLEAR
WARHEAD

Words appearing in the MAPTEXT output associated with conventional, or non-nuclear, warfare were represented by a C. These words included:

ADMIRAL
ADMIRALS
ADMIRALTY
ANNUMITION
AMPHIBIOUS
ANTIAIRCRAFT
ANTI-AIRCRAFT
ANTISUBMARINE
ANTITANK
ARM
ARMAMENT
ARMAMENTS
ARMED
ARMIES'
ARMIES
ARMING
ARMORED
ARMOR
ARMS
ARMY'S
ARMY
ARTILLERY
ASSAULT-LANDING
BATTLESHIPS
BATTALION
BATTLEFIELD
BATTLEFIELDS
BATTLESHIP
B-FIFTY-EIGHT
B-FIFTY-EIGHTS
B-FIFTY-TWO
B-FIFTY-TWOS
B-FIFTY-SEVEN
B-FIFTY-SEVENS
B-FIFTY-NINE
B-FIFTY-NINES
BRIDGEHEAD
BRIDGEHEADS
BRIGADE
BRIGADES
BROWNING
CANNON
CARBINES
CARRIER-BASED
CARRIER
COLUMN
COMPANY
CONSCRIPTION
CONVOY
CONVOYS
CORPS
COUNTEROFFENSIVE
CRUISER
DEMOBILIZATION
DEMOBILIZED
DEPOT
DESTROYER
DISARMED
DIVISION
DYNA-SOAR
ECHELONED
ECHELON
ECHELONS
ENCIRCLED
ENCIRCLEMENT
ENCIRCLE
ENCIRCLING
FIGHTER
FIREARM
FLAGSHIP
FLEET
FLOTILLAS
FORTIFICATION
FORTIFICATIONS
GARRISON
GENERAL
GROUND-SUPPORT
GUN

GUNS
HOWITZER
HOWITZERS
LANDED
LANDINGS
LIMITED-WAR
MAGAZINE-LOADING
MARINE
MINE
MORTAR
MUNITIONS
NAVAL
NAVIES
NAVY
NAVY'S
PRISONER
PROPULSION
RANK
REGIMENT
REMILITARIZATION
REPEATER
RIFLE
SHIP
SHIPS
SOLDIER
SOLDIERS
SQUADRON
SUBDIVISION
SUBMARINE
SUBMARINES
SWORD
TANKER
TNT
TORPEDOES
TORPEDO
TROOP
TROOPS
TROOPS'
UNIT
WARSHIP
WEAPONRY
WEAPON
WEAPONS

A number of words that might be associated with either nuclear or conventional warfare were included in the output. This group included words such as BOMB, JET, PILOT, MILITARY, ATTACK, ARSENAL, etc. The 67 words in this group were represented in the output by the characters H, G, F, E, D, and B.[14] A few words associated with peace were also included in the MAPTEXT representation. These words--ACCORD, DISARMAMENT, PEACE, PEACEFUL, PEACEFULNESS, PEACETIME, PEACE-LOVING, PEACELOVING--were represented by the letter K.

Some of the words on the lists associated with nuclear and conventional warfare may themselves seem ambiguous to the present-day reader. AIR-TO-AIR, for example, does not in many contexts imply nuclear warfare. In Soviet Military Strategy, however, words describing missiles, rockets, etc., are, in the main, associated with nuclear war. Thus, for the purposes of this

[14] Different characters were used for these ambiguous words because subsets of these words seemed to fall into groups that might be analyzed separately; e.g., BOMBARDMENT, BOMBED, BOMBER, BOMBERS, BOMBING, BOMBINGS, BOMB, BOMBS, and BOMBARDING were all represented by the letter B.

representation, such words were grouped under the nuclear heading and represented by an A.

An example of the 46-page MAPTEXT representation of the complete RAND translation is shown in Figure 10.

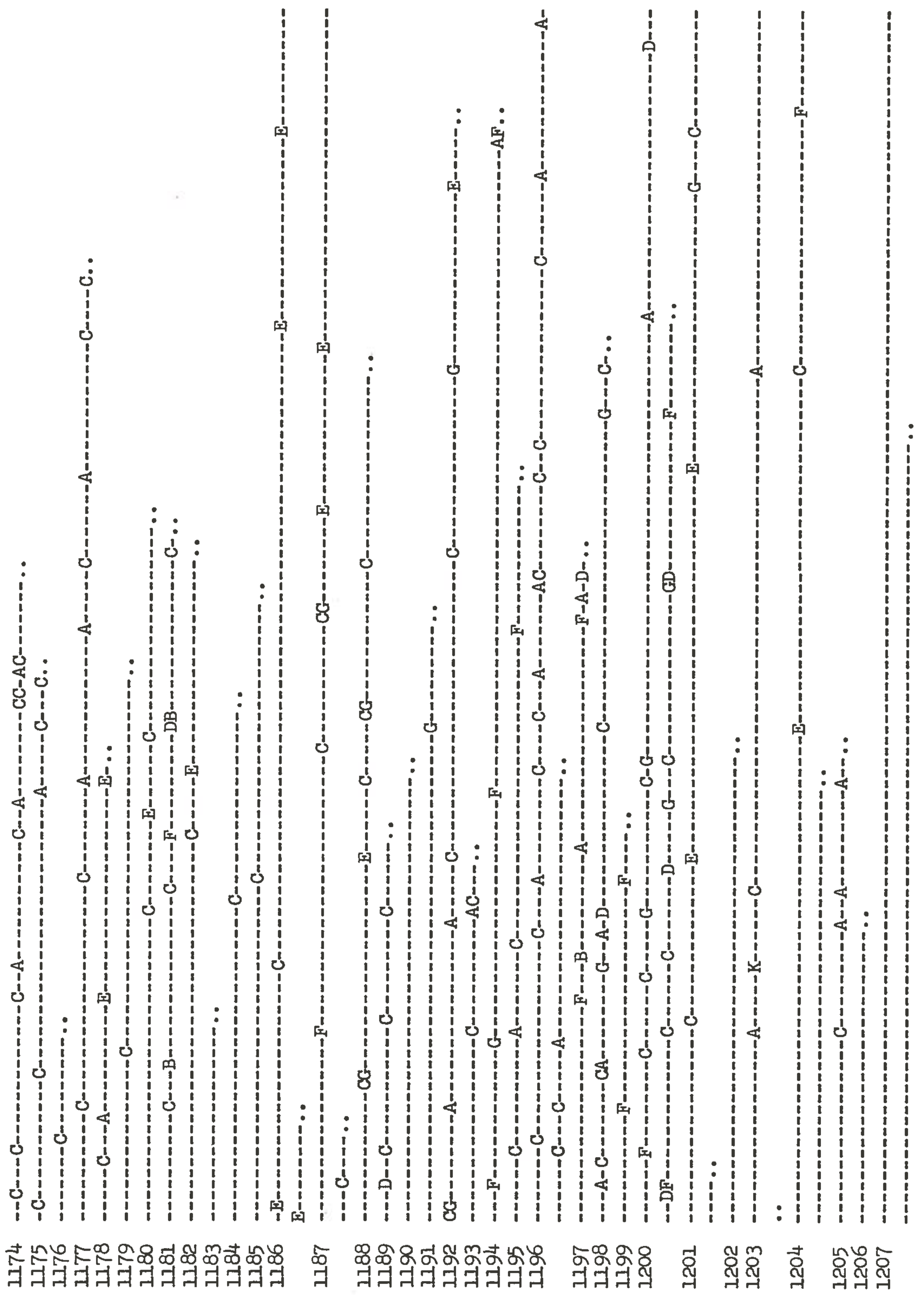

Figure 10. Page from Chapter Four, RAND Translation, Beginning with Paragraph 1174

```
1208 -----------------------------------------------------..
1209 ---------E------------------------------------------..
1210 --------------------------E------------..
1211 ---------C---------------C---------------------------------------------------------C---------------------------
     --------------------------A-----------------C------------E----C-------..
1212 ---------------------AC-E----C--------C-----------G------------C-----C-..
1213 ---------C--------E-----E-------C---------------C-------------------C----------------------------------------F--
     -..
1214 -------------------C------C------C-------C---..
1215 --------------------..
1216 --------------------------------------C--E-----..
1217 -------------------C--------C--------------C--G------------------E-------------C--------F---------C-
1218 -C-----------------------------------------------------------------..
1219 --------------------C-------------C---------------------C------E---------------FC--------------C---K-----------
     ----------..
1220 -----------C--K-----------C------..
1221 -----C-------------------..
1222 --------------------C-------------A--G-..
```

Each dash stands for a word that does not fall within the categories represented by A, C, etc. The numbers (e.g., 1174, 1175) at the left sequence paragraphs. The double period indicates the end of a paragraph.

Figure 10. (Continued)

An inspection of the entire output reveals some rather interesting patterns. The first chapter, "General Concepts," looks as the VIA output had lead us to expect: there is a heavy, quantitative emphasis upon words associated with conventional warfare. Of the 291 paragraphs in this chapter, only a few groups of paragraphs have noticeable collections of words associated with nuclear warfare. In paragraphs 56-78, 14 of the 82 words represented by special characters are in the A (nuclear) category; in paragraphs 266-268, 8 of the 20 words specially designated are represented with an A; and in paragraphs 287-289, the ratio is 8 to 20. These are the instances of high concentration of words associated with nuclear warfare in chapter one. There are long stretches in this chapter that have no A words at all: paragraphs 1-55, 79-129, 131-259. Of course a number of the non-A words are also non-C (conventional warfare) words; but, because of the much higher incidence of C words, one might guess that the ambiguous words in this chapter would most often resolve on the side of conventional warfare. The work with VIA, described in last year's annual report, indicated that this is, in fact, the case.

The first chapter, "General Concepts," is a general description of the techniques of war, with some slight emphasis upon Soviet methods of warfare. The second chapter is concerned with the "Military Strategy of Imperialist States and Their Preparation of New Wars." In this chapter, the incidence of words associated with nuclear warfare greatly increases. For example, in paragraphs 413-460, which look to have about a normal density (for this chapter) of words represented with special characters, 75 of the 221 words thus represented are nuclear warfare words. Notice that the number of paragraphs

used for this sample is considerably greater than the longest sample that could be obtained for chapter one and that, even within this large sample, the ratio of nuclear to other words is more than one-third. In fact, other than the first 55 paragraphs in this chapter, which concentrate upon the military policies of "imperialist states" during World War I, it is difficult to find any section of this chapter that has no words associated with nuclear warfare. The next longest such section, about 20 paragraphs, is a passage concerned with economic preparation for war.

In contrast to chapter two, chapter three, "Development of Soviet Military Strategy (1917-1945)," has an extremely low incidence of words associated with nuclear warfare. In fact, in a chapter consisting of 331 paragraphs, there are only five A words in the entire chapter. Naturally, the time span covered by this chapter is largely responsible for the tremendous decline in words associated with nuclear warfare. The interesting thing about the book thus far, though, is that it has been structured so that nuclear warfare has definitely been associated with countries antagonistic to Soviet Russia. Bearing this fact in mind, it is well to notice that in chapter four, "The Nature of Modern War," the first heavy clustering of words associated with peace (represented by K) occurs. There are two such clusterings early in the chapter. The first, covering 16 paragraphs, has 9 such words in a total of 44 words; the second, covering 12 paragraphs, has a ratio of 18 to 37, or nearly half. A glance at the text will show that these emphases on peace, occurring early in a chapter on modern war, appear in passages describing the Marxist-Leninist approach to the problem of war and peace, with an emphasis upon peace.

The pattern that has been established--imperialist nations and nuclear warfare vs. the Soviet Union, conventional warfare and peace--is one that recurs in the rest of the book. For example, in chapter six, "Methods of Conducting Warfare," two sections with a high incidence of nuclear warfare words turn out to be descriptions of American atomic strikes in Japan, at the close of World War II, and of "imperialists'" methods of waging war. Chapter seven, "Preparation of a Country to Repel Aggression," has a very high incidence of words associated with peace, which are in turn associated with preparations the Soviet Union might make during peacetime to maintain peace and repel aggression, including nuclear aggression, from nations in the West. There are, of course, passages in these later chapters in which nuclear weapons are associated with the Soviet Union, because the authors feel the need to declare the supremacy of the Soviet Union so far as nuclear potential is concerned. But the main rhetorical structure of the book mirrors many of the individual statements within the book which describe the Soviet Union as devoted to peace or, at the least, a more humane war than the imperialist, war-loving, nuclearly-destructive aggressors. The first nuclear "attack" in the book is associated with the Western nations; the first emphasis upon peace in the book is associated with the Soviet Union, and the initial and continued emphasis upon conventional war is associated with the Soviet Union.

The importance of the perception of such a structure is that it may suggest something about the authorship of the book. Although it is true that many potential or actual authors in the Soviet Union doubtless have the biases expressed in this book and would thus be perfectly capable of making many of the statements

in the book, it is another matter to say that a work thrown together by many different authors will have the rather coherent rhetorical structure which MAPTEXT has revealed in Soviet Military Strategy. Such a structure would seem to suggest, at the least, a rather close supervision of the content and arrangement of the book by one, or, at the most, a very few individuals. Taken in conjunction with the evaluation of the first chapter produced by VIA, which indicated an author preference for conventional, especially ground, warfare, this rhetorical structure would also seem to suggest that if the book is in part a propaganda effort, the propaganda effort in this case is intended not only to persuade the reader of the moral rectitude of the Soviet Union but to couple moral rectitude with the conventional warfare in which the author of the first chapter, at any rate, has more confidence. Thus, it would seem in part not only an attempt to evoke revulsion for the West and its nuclear warfare in order to ensure moral condemnation of the West, but also in order to ensure condemnation of a military technique that has not been accepted as thoroughly practical by the person who controlled the rhetorical structure of Soviet Military Strategy.

Because the Philco 2000, the computer for which our programs were written, was dismantled, we have had to defer some other MAPTEXT runs--which might prove additionally illuminating--until the program is rewritten. For example, we would like to produce a MAPTEXT representation that combines words implying emotional attitude with the words already used for nuclear-conventional warfare. Any additional evidence of bias, or lack of bias, supplied by such a combined run, would help establish the validity or lack of validity of the hypothesis

outlined in the above paragraph. Separate, condensed MAPTEXT representations of the nuclear and conventional warfare distribution patterns would also help display the rhetorical structure more clearly.

To exemplify the use of the condensed output, a question posed in last year's report might be considered. The question concerned the use of the word GENERAL, and words that are in some senses synonymous with it--"natural," "total," "typical," "universal," "whole," and "common." An investigation of the use of GENERAL in chapter one of the Praeger translation of Soviet Military Strategy showed it to be used in an abstract, "in general" fashion. The suggestion made in the report was that, if words synonymous with GENERAL were also used in such fashion, a comparison of the distribution patterns of GENERAL and its synonyms might tell us something about the style of the translator, if not about the style of the original work underlying the translation. Figure 11 shows the condensed MAPTEXT outputs for, first, GENERAL and, second, GENERAL plus its synonyms. Each asterisk stands for one occurrence. The characters above some of the asterisks have been added after "manual" consultation of the noncondensed MAPTEXT output, to indicate which of the synonyms the asterisk represents. T represents TYPICAL; W, WHOLE; U, UNIVERSAL; Y, TOTAL; C, COMMON; and N, NATURAL. Paragraphs are condensed into groups of four, so that the first row in Figure 11 represents occurrences in paragraphs 1-4, the second row in paragraphs 5-8, etc.

Paragraph	General	General and Synonyms
4	*	*
8	***	T ****
12		
16		
20		
24	*****	W N *******
28	**	W ***
32	*	U **
36	**	**
40	*	WW ***
44		W *
48		
52		W *
56	**	**
60		WW **
64		
68	*	*
72		
76	*	U W ***
80	***	***
84	***	***
88	**	T ***
92		

Figure 11. Condensed MAPTEXT Outputs

Paragraph	General	General and Synonyms
96		
100		
104		
		W
108		*
112		
116		
120		
124		
128	**	**
		T
132		*
		W W
136	***	*****
140	*	*
144		
		C
148		*
		T
152		*
156		
160	***	***
		W
164		*
168		
		W
172		*
176		
180		
		W
184	*	**

Figure 11. (Continued)

Paragraph	General	General and Synonyms
188		
192	*	*
196	**	**
200		
		T
204		*
		W
208		*
212		
216		
		W
220		*
224	*	*
228	*	*
232		

Figure 11. (Continued)

The thought was that if GENERAL and its synonyms tended to cluster in the same locations, one might conclude that the translator was striving for some variety even though he, or the original text, had fallen into a habit, of a few paragraphs' duration, of using vague descriptors such as "in general." As Figure 11 shows, the results of the combined comparison are not clear-cut, but there is a tendency toward clustering in the same location. The only synonyms at a distance of more than four paragraphs from a paragraph group containing the word GENERAL are WHOLE, in paragraph row 108; COMMON, in paragraph row 148; TOTAL, in paragraph row 152; WHOLE, in 172; TOTAL, in 204; and WHOLE, in 208. Of the synonyms used, WHOLE occurs almost twice as often (17 occurrences) as

all the others combined (9 occurrences). Eleven occurrences of WHOLE appear in the phrase "as a whole," and two in the phrase "on the whole"; both phrases are roughly synonymous to the use of GENERAL in the phrase "in general." WHOLE is used adjectivally four times ("whole depths," "whole territory," "whole continents," "whole world"). One of these uses is in paragraph row 208, and the other three are in paragraph rows 52 and 60; thus, in no case are the adjectival uses of WHOLE part of a paragraph row "string" that also includes uses of GENERAL. When WHOLE does appear in such a string it does, then, tend to provide a varied locution for the rather vague collocation implied by GENERAL. Thus, the translator, at least, can be given some credit for an effort to vary his style somewhat. When MAPTEXT is running again, we'll compare the RAND translation with the Praeger translation; if the same sort of pattern appears in RAND, we might reasonably conclude that the modest push for stylistic variety may exist in the original.

MAPTEXT, in its current version, obviously can provide a quick guide to some of the structural and distributional aspects of style in verbal material. In its future version, options for the condensed output feature will be enlarged so that the researcher can obtain, on one run, composite and individual graphs for the words represented by special characters within that run. Then we could request, for example, individual graphs for each of the synonyms for GENERAL, in addition to that for GENERAL, as well as the composite graph shown in the right-hand column of Figure 11.

C. Some Stylistic Implications of a Comparison of Content Words in the Two Translations of SOVIET MILITARY STRATEGY.

This comparison was initially suggested by a survey of the VIA outputs for chapter one in both the RAND and Praeger translations of Soviet Military Strategy. It was clear from that output that there were some important word choice distinctions between the two translations. The most striking difference was the use of the word CONFLICT, 24 times in the Praeger translation and not at all in the RAND translation; the word COMBAT appeared a compensatory number of times in the latter. We had planned to do extensive numerical comparisons of the remaining chapters in both translations but were unable to complete the task before the Philco 2000 was dismantled. We did, however, succeed in getting an output that simply lists words occurring in the Praeger translation and not in RAND, and vice versa. The output is a surprisingly long 52 pages. Within those pages a few stylistic patterns are clearly apparent, even without the statistical summaries we had planned to have.

In an effort to determine whether either of the translations showed syntactic preferences, the lists were perused for instances in which one or more forms of a word appeared in one list and different forms of the same word in the other list. For example, the Praeger translation used the words ADAPTABILITY and ADAPTION, whereas RAND ignored those forms but did use the word ADAPTABLE. The latter form did not, of course, occur in the Praeger translation. In all, there were 127 of these instances of disparate forms of the same word. Within these 127 instances, many of the forms chosen do not consistently appear in either one list or the other. There are two exceptions,

though, to the generally random character of the use of disparate forms. In 11 cases, the Praeger translation uses an -ing form whereas the RAND translation uses simply a present, active verb form. For example, Praeger uses ENLISTING, which does not occur in RAND, and RAND uses ENLIST, which does not occur in Praeger. The reverse of this pattern--that is, the present active form in Praeger and the -ing form in RAND--occurs just twice. The other discernible syntactic preference pattern is for the -ion nominative form in the Praeger translation. For example, Praeger has COMPILATION, RAND has COMPILED and COMPILE; Praeger has COMPLICATIONS, RAND has COMPLICATE and COMPLICATES; Praeger has CONCEPTIONS, RAND has CONCEPTUAL, etc. In total, there are 13 such instances; the reverse situation (the -ion ending in RAND and not in Praeger) occurs 4 times. Although these usage patterns may not seem, on the surface, striking, it was just such seemingly trivial "subconscious" patterns of stylistic choice that enabled Mosteller and Wallace to produce such satisfactory statistical estimates as to the authorship of the disputed Federalist Papers.[15]

It would be interesting to pursue further the syntactic distinctions cited above by getting condensed MAPTEXT distribution patterns for the -ion, -ing forms and their counterparts. If they cluster, they might well represent the stylistic preference of one of the several translators associated with the project.

---

[15] F. Mosteller and D. L. Wallace, Inference and Disputed Authorship: The Federalist, Addison-Wesley Publishing Co., Inc., Reading, Massachusetts, 1964.

Other than the patterns of syntactic preference, one other much more general but nonetheless interesting pattern emerges from a study of the Praeger and RAND lists. That is, that once an author or, in our case, translator, begins to use a particular word he is likely to use several forms of that word. For example, Praeger has CULMINATED, CULMINATES, CULMINATING, CULMINATION, whereas no form of the word appears in RAND. On the other hand, RAND has DECIMATE and DECIMATED, but no form of this word occurs in Praeger. There are 120 such instances in the two lists: 90, or three-fourths of the instances, occur in the RAND translation. This fact might seem to indicate that the RAND translators were less inventive than the Praeger translators, but, on the other hand, of the approximately 2930 words occurring in the total output, 1599 appear in the RAND translation and not in the Praeger, whereas only 1331 words appear in Praeger and not in RAND. The multiform use of a particular word or words might indeed indicate the hand of a particular translator, however. Again, some condensed MAPTEXT outputs might prove informative.

When our programs are again operative, we will pursue this kind of comparison in more detail. The significance of the results obtained thus far lies not only in the techniques indicated for similar and further research, but also in the fact that the choice of content words is so demonstrably an aspect of stylistic choice. The almost 3000 content words appearing in the lists of words occurring in the RAND translation but not in Praeger, and vice versa, clearly show that one writer will describe the same situation (in this case, the "situation" is a verbal text) with one set of words, while a

second writer will use a variant set of words. The use of two translations of one word for analysis probably provides the strongest evidence obtainable that content word choice is an aspect of style. This point certainly should no longer be a matter for the dispute implicitly contained in all discussions of style that entail distinctions between style and content, or form and content. So long as such distinctions are made, the problem of the analysis of style will be obscured.

III. Plans for the Future

The future directions of research on the stylistic analysis project will be based on the accomplishments already achieved. Since VIA works well as a content analysis program, its operation and output will be refined to make it increasingly useful. The MAPTEXT representations have already been shown to provide very quick leads for the perception of spatial clustering of concepts as well as for the spatial relationships among concepts which reveal rhetorical, or structural, aspects of style; to improve the ease with which MAPTEXT can be used, programming modifications to increase its output flexibility and clarity will be introduced. In addition, more basic research on thesauri and on statistical measures to be used in the stylistic analysis systems will be carried out.

The introduction of computer program modifications will take place concomitantly with the rewriting of programs for the IBM S/360 computers. The higher order language being used for much of the reprogramming is PL/I, a language much better adapted to verbal data processing than is FORTRAN, but which should have the wide availability for which FORTRAN was chosen for the original programs.

The new VIA program will incorporate a number of changes. For example, the INDEX phase of VIA will be made more general so that it can index oral as well as written verbal material. The THESAUR portion of VIA will be given the power to retrieve word pairs and triplets, as well as single words. THESAUR will also be altered so that the printout will show the text form as well as the thesaurus form for associated words of the same root; the printout will not,

however, "balloon" output by including many different thesaurus forms with a given root when only one or two of those forms appear in the text--in these circumstances, only the form or forms appearing in the text will appear in the printout. Some additions to THESAUR, incorporated since VIA was last described, will be retained. For example, the output now subdivides associated word lists so that the appropriate words are associated with DEADLY, another group of appropriate words is associated with DEAD, etc. This change was necessitated by the thesaurus experiments, which subdivided root groups in this manner and listed thesaurus entries for each of the subdivisions, or root forms. A concordance feature has been added to the VIA package so that the researcher can automatically have a concordance for primary words, if he chooses. This feature will be retained with an improved output format.

Work toward the complete automation of VIA will continue. The emphasis here will be upon further work on thesauri. Additional comparisons of reference thesauri and synonym dictionaries will be made so that, for example, more general deductions as to the indexing bias in a given thesaurus can be drawn. Efforts directed toward experimenting with various combinations of thesauri and context will begin in an attempt to see whether context can indeed supplement weak areas in a thesaurus.

The output potential of MAPTEXT will be increased so that in addition to the MAPTEXT of the complete text (see the sample in Figure 10), condensed output graphs for each of the words represented by a special character in the complete text representation will be produced (see sample in Figure 11). The readability of the complete text representation will be improved by providing larger spatial demarcations between major data units such as chapters.

Efforts directed toward the introduction of statistical tests to specific sections of the programming system will also be initiated. For example, when our programs are again operative, we would like to obtain more precise quantitative data regarding the differences in word choice between the two translations of Soviet Military Strategy than was possible during the current year. We would also like to begin exploration of word sequencing choices, both as to type and individual word, by using Markov procedures where applicable. When we begin to have a number of different operating procedures for the determination of style, some of the implications of the output from those procedures (such as those implications already seen in the MAPTEXT representations of Soviet Military Strategy) should begin to coalesce toward an extensive, coherent picture of the style of any given body of verbal data.

IV. Professional Activities

1. Sedelow, S. Y. The Computer and Verbal Data Processing in the Humanities. Presented at the University of North Carolina at Chapel Hill, February 25, 1966.

2. Sedelow, S. Y. Stylistic Analysis, Report on the Second Year of Research. SDC document TM-1908/200/00, Santa Monica, California, System Development Corporation, March 1, 1966.

3. Sedelow, S. Y. The Computer and the Study of Literature. Presented at the California Institute of Technology, March 4, 1966.

4. Sedelow, S. Y. The Computer and the Study of Language. Presented at the University of Connecticut, April 22, 1966.

5. Sedelow, S. Y. Quarterly Report, 1 March 1966 to 1 June 1966. SDC document TM-(L)-1908/011/00, Santa Monica, California, System Development Corporation, May 25, 1966.

6. Sedelow, S. Y. and Sedelow, W. A., Jr. "A Preface to Computational Stylistics," in J. Leed (Ed.), The Computer and Literary Style, Kent State University Press, Ohio, 1966.

7. Sedelow, S. Y. "Form Recognition in Literature," Proceedings of IFIP Congress, 1965, Vol. II, Spartan Books, pp. 626-627.

8. Sedelow, S. Y. Quarterly Report, 1 June 1966 to 1 September 1966. SDC document TM-(L)-1908/012/00, Santa Monica, California, System Development Corporation, August 31, 1966.

9. Sedelow, S. Y. The Computer and Linguistic Analysis. Presented to the Central Carolina ACM Chapter, Durham, North Carolina, November 11, 1966.

10. Sedelow, S. Y. Teaching Style with the Aid of the Computer. Presented at the National Council of Teachers of English Annual Convention, Houston, Texas, November 25, 1966.

11. Sedelow, S. Y. Quarterly Report, 1 September 1966 to 1 December 1966, SDC document TM-(L)-1908/013/00, Santa Monica, California, System Development Corporation, November 30, 1966.

Appendix:

VIA Output for Word DECLINE, RAND Chapter One, Showing Comparative Retrieval Capacities for Webster's Dictionary of Synonyms, Roget's International Thesaurus, and Roget's University Thesaurus.

Webster's Dictionary of Synonyms

DECLINE
DECLINE
--------ACCEPT
--------ADVANCE
--------PROGRESS
--------RECEIVE
--------REFUSE
--------RISE
--------SPURN
--------STICK
--------TAKE

Roget's International Thesaurus

DECLINE
DECLINED
--------DECLINING (Different Form Appears in the Text)
DECLINE
--------AGE
--------BACKWARDATION (Different Form Appears in the Text)
--------BACKWARDIZATION (Different Form Appears in the Text)
--------BREAK
--------CLOSE
--------CLOSING (Different Form Appears in the Text)
--------DECLINATION (Different Form Appears in the Text)
--------DENY
--------EVEN (Different Form Appears in the Text)
--------FOUNDER
--------GO
--------IMPERFECTION (Different Form Appears in the Text)
--------INCLINE (Different Form Appears in the Text)
--------INSUFFICIENCY (Different Form Appears in the Text)
--------LACK
--------LESSEN (Different Form Appears in the Text)
--------LOWERING (Different Form Appears in the Text)
--------LOWER
--------PARE (Different Form Appears in the Text)
--------REDUCE (Different Form Appears in the Text)
--------REDUCTION (Different Form Appears in the Text)
--------REFUSE (Different Form Appears in the Text)
--------RISE
--------SET
--------SHORTAGE
--------SHORTEN (Different Form Appears in the Text)
--------SHORTFALL (Different Form Appears in the Text)
--------WEAKEN (Different Form Appears in the Text)
--------WINTER
--------WITHDRAW

Roget's University Thesaurus

DECLINE

DECLINE

--------ACCIDENT (Different Form Appears in the Text)
--------AGE
--------AIDLESS (Different Form Appears in the Text)
--------ANALYSIS
--------ANCIENT
--------ANTIQUE (Different Form Appears in the Text)
--------ANTIQUITY
--------ATTACK
--------BLOW
--------CALAMITOUS (Different Form Appears in the Text)
--------CALAMITY
--------CARE (Different Form Appears in the Text)
--------CASE
--------CATASTROPHE
--------CLASSIC (Different Form Appears in the Text)
--------COLD
--------CONFINED
--------CONSUMPTION (Different Form Appears in the Text)
--------CUSTOMARY (Different Form Appears in the Text)
--------CUSTOM
--------DEAD (Different Form Appears in the Text)
--------DECLINATION (Different Form Appears in the Text)
--------DECLINING (Different Form Appears in the Text)
--------DECLINE
--------DECREMENT (Different Form Appears in the Text)
--------DEFENSELESS (Different Form Appears in the Text)
--------DENIAL (Different Form Appears in the Text)
--------DENY
--------DEVOTED
--------DISASTER
--------DISASTROUS
--------DOWNWARD (Different Form Appears in the Text)
--------DOWNWARDS (Different Form Appears in the Text)
--------DOWNFALL
--------DOWN
--------ERODE (Different Form Appears in the Text)
--------EROSION (Different Form Appears in the Text)
--------EXCLUSION (Different Form Appears in the Text)
--------EXCLUDE
--------EXTREMITY (Different Form Appears in the Text)
--------FIT
--------GATHERING (Different Form Appears in the Text)
--------HALTING (Different Form Appears in the Text)

--------HALT
--------HOSPITAL
--------IMPOSSIBLE
--------INCLINATION (Different Form Appears in the Text)
--------INFLAMMATION (Different Form Appears in the Text)
--------ISSUE
--------LESSEN (Different Form Appears in the Text)
--------LIGHT
--------LOAD
--------LOWER
--------MINIMIZE
--------NEITHER
--------NO
--------OBSOLETE
--------PAPERY (Different Form Appears in the Text)
--------PASSE (Different Form Appears in the Text)
--------PASS (Different Form Appears in the Text)
--------PAST
--------POISONED (Different Form Appears in the Text)
--------POORLY (Different Form Appears in the Text)
--------POOR
--------POWERLESS (Different Form Appears in the Text)
--------PRESSURE
--------PRIME (Different Form Appears in the Text)
--------REDUCE (Different Form Appears in the Text)
--------REDUCTION (Different Form Appears in the Text)
--------REFUSAL (Different Form Appears in the Text)
--------REFUSE (Different Form Appears in the Text)
--------REFUSED
--------REFUSING
--------REPULSE (Different Form Appears in the Text)
--------RESIST (Different Form Appears in the Text)
--------RESTIVE (Different Form Appears in the Text)
--------RIPENESS (Different Form Appears in the Text)
--------ROT (Different Form Appears in the Text)
--------ROTTEN
--------SEIZURE
--------SET
--------SHORTEN (Different Form Appears in the Text)
--------SPURN (Different Form Appears in the Text)
--------STRENGTHLESS (Different Form Appears in the Text)
--------SUFFER (Different Form Appears in the Text)
--------TRADITIONAL (Different Form Appears in the Text)
--------TRADITIONARY (Different Form Appears in the Text)
--------TRADITIVE (Different Form Appears in the Text)
--------TRADITION
--------TRIAL

--------UNFAVORABLE
--------UNHEALTHY
--------WEAKEN (Different Form Appears in the Text)
--------WEAKLY (Different Form Appears in the Text)
--------WEAKNESS (Different Form Appears in the Text)
--------WEAKENED
--------WEAK
--------WITHDRAWAL (Different Form Appears in the Text)

Unclassified
Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)* System Development Corporation, Santa Monica, California	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

3. REPORT TITLE
Stylistic Analysis Report on the Third Year of Research
(Final Report under this Contract)

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*

5. AUTHOR(S) *(Last name, first name, initial)*
Sedelow, Sally Yeates

6. REPORT DATE 1 March 1967	7a. TOTAL NO. OF PAGES 99p.	7b. NO. OF REFS
8a. CONTRACT OR GRANT NO. Nonr-4427(00) b. PROJECT NO. Office of Naval Research, U.S. Navy Task No. NR 348-005 c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) TM-1908/300/00 9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*	

10. AVAILABILITY/LIMITATION NOTICES
Distribution of this document is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY

13. ABSTRACT

This report describes current research associated with the computer-based stylistic analysis project. Work directed toward the complete automation of thesaurus construction for the VIA (content analysis) program is reported in some detail; the focus of this work has been upon comparative VIA runs using Webster's Dictionary of Synonyms, Roget's International Thesaurus, and Roget's University Thesaurus. As a result of these runs, conclusions are drawn regarding preferable sources for the computer thesaurus as well as regarding some technical aspects of VIA's operation. The power of MAPTEXT for quickly revealing stylistic patterns to the information analyst or intelligence officer is illustrated.

DD FORM 1 JAN 64 1473

Unclassified
Security Classification

Unclassified

Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Stylistic Analysis Computer Based Stylistic Analysis Thesaurus Construction						

INSTRUCTIONS

1. ORIGINATING ACTIVITY: Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization *(corporate author)* issuing the report.

2*a*. REPORT SECURITY CLASSIFICATION: Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.

2*b*. GROUP: Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. REPORT TITLE: Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. DESCRIPTIVE NOTES: If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. AUTHOR(S): Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

6. REPORT DATE: Enter the date of the report as day, month, year; or month, year. If more than one date appears on the report, use date of publication.

7*a*. TOTAL NUMBER OF PAGES: The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7*b*. NUMBER OF REFERENCES: Enter the total number of references cited in the report.

8*a*. CONTRACT OR GRANT NUMBER: If appropriate, enter the applicable number of the contract or grant under which the report was written.

8*b*, 8*c*, & 8*d*. PROJECT NUMBER: Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9*a*. ORIGINATOR'S REPORT NUMBER(S): Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

9*b*. OTHER REPORT NUMBER(S): If the report has been assigned any other report numbers *(either by the originator or by the sponsor)*, also enter this number(s).

10. AVAILABILITY/LIMITATION NOTICES: Enter any limitations on further dissemination of the report, other than those imposed by security classification, using standard statements such as:

(1) "Qualified requesters may obtain copies of this report from DDC."

(2) "Foreign announcement and dissemination of this report by DDC is not authorized."

(3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through

_____."

(4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through

_____."

(5) "All distribution of this report is controlled. Qualified DDC users shall request through

_____."

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. SUPPLEMENTARY NOTES: Use for additional explanatory notes.

12. SPONSORING MILITARY ACTIVITY: Enter the name of the departmental project office or laboratory sponsoring *(paying for)* the research and development. Include address.

13. ABSTRACT: Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as *(TS)*, *(S)*, *(C)*, or *(U)*

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. KEY WORDS: Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, rules, and weights is optional.

Unclassified

Security Classification

System Development Corporation/2500 Colorado Avenue/Santa Monica, California 90406