The Analysis of Communication Content

Developments in Scientific Theories and Computer Techniques

Edited by

GEORGE GERBNER
OLE R. HOLSTI
KLAUS KRIPPENDORFF
WILLIAM J. PAISLEY
PHILIP J. STONE

JOHN WILEY & SONS, INC. New York · London · Sydney · Toronto

26 Categories and Procedures for Content Analysis in the Humanities[1]

Sally Yeates Sedelow

English and Information Science

Walter A. Sedelow, Jr.

Sociology and Information Science
University of North Carolina

As part of an exploration of style in language, we[2] have been concerned for several years with content analysis—not for the humanities alone but also for the social sciences. Humanistically, we have been interested, among other things, in the specification and location of literary themes; in other texts, we have been concerned with what are sometimes called structuring ideas. We undertook content analysis in this rather generalized way because we suspected that the distinctions between themes and ideas were more minimal than is sometimes argued or suggested. It is true that literature (and we are assuming throughout this chapter that the aspect of the humanities most relevant for content analysis is literature, although paintings, musical scores, and the like, also may be content analyzed) or, more properly, the analysis of literature, does differ in part from the study of texts in some other disciplines. The use made of ambiguity is probably a good example of a major distinguishing characteristic, where the variance has repercussions as to the appropriate analyses: for much literature revels in the kinds of ambiguity implied by puns and other word play, and such a use of language runs counter, for example, to that strain toward explicit or implicit operational definition which is criti-

[1] The research described in this chapter has been supported, in part, by the Office of Naval Research, Information Systems Branch.

[2] Other people who have been associated with this project include Terry Ruggles of the System Development Corporation, Joan Bardez, William Hickok, and William Buttelmann of the Information Science Department at the University of North Carolina (Chapel Hill), and Joan Peters of Washington University (St. Louis). For further elaboration, and relevant readings, see Sedelow (1964, 1965a, 1965b, 1966a, 1966b, 1967a, 1967b).

cally significant for much social analysis prose. But, then, humanistically one also enjoys and exploits the kind of ambiguous connotations which psychologists, too, may study when, for example, they are engaged in clinical analyses. And, insofar as an interest in using the language of motive, intention, and attitude as applied to verbal material is shared in other disciplines, such as political science and sociology, content analysis which is desirable for the humanities is also content analysis appropriate to social science. Thus, the categories and procedures that we shall discuss in this chapter are not, in fact, restricted to the humanities; and we have used them for such varied texts as *Hamlet*,[3] *Soviet Military Strategy*,[4] and very long strings from Hume's *History of England*.[5]

As is the case with much content analysis, the major question we faced was how to set up grouping procedures, the categories which would contain words forming the thematic or conceptual patterns in the texts; since, initially, we were not testing particular theories or looking for specific syndromes, but rather were searching—with an effort to avoid premature closure—for whatever patterns the text might reveal itself to exemplify, we did not want to begin with categories already set up for their relevance to a particular theory or concept scheme. Therefore, our content analysis program (really a series of computer programs grouped under the general title VIA, for Verbally-Indexed-Associations) begins simply by indexing the text, grouping the content words in the text together by root, and counting the occurrences of the words within the root group. Our procedure was next to assume that we might want to sort on any nonfunction word root and form subgroups around it, that is, that *any* nonfunction word or group of words with a common root might serve as a category key with which other words would eventually be grouped. The word or root group serving as a key we designated by the word, primary, and the words ultimately placed in that category we designated as associated words. This designation could be misunderstood, because the associated words may be no less valuable, no less critical, for the delineation of a theme than the primary word or root group. We used the designation simply to indicate which word or root group served as the key for the construction of a category because we needed to make such distinctions

[3] William Shakespeare, *Hamlet*, George Lyman Kittridge, ed., Ginn and Company, 1939.

[4] We used two translations of V. D. Sokolovsky's *Soviet Military Strategy*: I. Dinerstein D., Goure and Wolff, Prentice-Hall, Inc., 1963; II. Translation Services Branch, Foreign Technology Division, Wright-Patterson AFB, Frederick A. Praeger, Inc., 1963. Customarily we refer to the former as the RAND translation, and the latter as the Praeger translation.

[5] David Hume, *The History of England*. 6 Volumes, London, 1841.

in order to talk about our programming procedure. The selection of such a category key or root group was based upon frequency. In most cases we used high frequency, although in one instance when we were interested in comparing two translations of the same work we used any frequency, from one up, for words that occurred in one translation and not in the other. The researcher may, of course, choose any frequency level that he likes, high or low, for keying his categories. An important characteristic of this procedure thus far is that the words serving as keys to the categories are words which occur in the text with any frequency, *n*, of interest to the researcher.

The next and crucial step in our content analysis procedure has been to select the words which, if they occurred in the text, would be appropriate to one of the categories designated by the key word. Up to the present, the selection of these words has been made manually. We have consulted various thesauri, synonym dictionaries, and the context of the primary words in order to draw up lists of possible associated words. As a consequence of this use of various sources for the input lists of possible associated words, we call these lists the computer's thesaurus—and because of the way the program operates, we have been constructing clusters of words which might be used in a new thesaurus if one were building such a thesaurus from scratch. Such clusters differ from those in any available thesaurus not only because of the varied sources for the words in the input thesaurus but, much more essentially, because the program links words down through five levels although the researcher provides just the first level links. The links may be spoken of as implicit in the material the researcher supplies to the computer, but the computer program pulls them together into new clusters of semantically associated words. It systematically explores for the network of transitivity implications in sets of particular associations. For example, in Figure 1, the words HEAVEN, LAND, NATURE, and WORLD were all first level links to the word EARTH. On the other hand, the words LIFE and UNNATURAL did not appear on any list linked directly to EARTH; LIFE and UNNATURAL were on a list linked to NATURE and the operation of the computer program in turn links them to EARTH. As was the case with LIFE and UNNATURAL, the program links STARS to EARTH because STARS has appeared on a list linked with HEAVEN. Also in the list with STARS is a recurrence of EARTH, exemplifying the indication of cross-referencing often provided by the content analysis program. A new version of the content analysis program will permit the researcher to suppress the second occurrence of EARTH if he so desires.

It is very important to remember, of course, that these thesauri are text-specific; that is, we have an output consisting of groups of associated

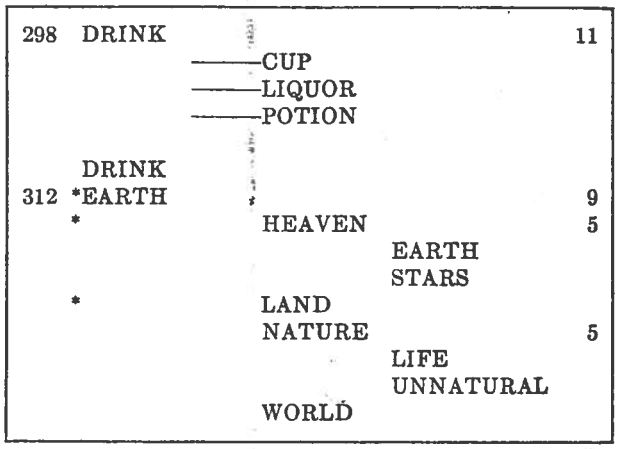

Figure 1 A sample of VIA's output for *Hamlet.*

words for *Hamlet* as well as separate outputs for both translations of *Soviet Military Strategy* and for Hume's *History of England.* A good deal of cross-referencing, as well as much additional use of the VIA program on extensive bodies of text, would be required if one were thinking of using the text-specific thesauri as bases for a new general thesaurus. As our research has progressed, the selection of words which, if they occur in the text, will form the categories for that text has become less eclectic and more "automatic." For example, initially, when looking at *Hamlet,* the tendency was not to include words appearing in the various thesauri or synonym dictionaries if it seemed reasonably certain that Shakespeare did not use such words. The reason for this selectivity was the desire to eliminate a good deal of manual work as well as computer processing time. Despite these considerations, it has seemed increasingly important to list *all* words in the appropriate categories found in thesauri and synonym dictionaries for two reasons: (1) the researcher probably does not know his author's vocabulary as thoroughly as he thinks he does, and (2) we began to think it would be desirable to automate this selection of possible associated words and wanted to see what the inclusion of all words in a category would do to our output. That is, we wanted to see if, by listing all words, our text-specific thesauri became overly swollen with extraneous words; and, if so, to begin thinking about controlling the excess by rule.

Since we do want to automate this manual section of our VIA procedure, and because we would like to utilize in some way an already existing thesaurus rather than construct a new subjectively structured group of

semantic categories, we conducted some experiments using *Webster's Dictionary of Synonyms, Roget's University Thesaurus,* and *Roget's International Thesaurus* to see whether, if the researcher were restricted to one of those thesaurus texts as the base for the computer thesaurus, one of them alone would prove satisfactory. We should quickly state that we do not envision our input thesaurus as being restricted to any one of these thesaurus texts. Instead, we were interested in seeing whether one of them looked promising enough so that it would be worthwhile to devise methods of amplification, or other modification, for its use. Our procedure was to choose a restricted number of words in the RAND translation of *Soviet Military Strategy,* and use *Webster's Dictionary of Synonyms, Roget's University Thesaurus,* and *Roget's International Thesaurus,* respectively, for the input thesaurus for each of three computer runs on chapter two. The number of words used for this experiment was restricted because the quantity of data involved turned out to be enormous, especially for the run using *Roget's University Thesaurus.* We were interested in trying the *University Thesaurus* because it does not subdivide categories as does the *International Thesaurus.* For example, the index for the word DEATH in the *University Thesaurus* refers to a category which includes both nouns and verbs in a great range of meanings. The index for DEATH in the *International Thesaurus* subdivides on the basis of meaning and syntax; the indexed section for DEATH as a noun in the *International Thesaurus* contains 26 entries as opposed to the single all-inclusive category of about 221 entries in the *University Thesaurus.* Our thought was that the *University Thesaurus,* because its entries were not subdivided, might provide a richer, more useful output than the *International Thesaurus.* In part this thought was confirmed, since the *University Thesaurus* produced a much more extensive text-specific thesaurus than did the *International Thesaurus.* Unfortunately, many of the words suggested by the *University Thesaurus* seemed extraneous to a given category. That was much less the case for the *International Thesaurus.* Extensive examination of the results[6] has convinced us that the greater richness of the results produced by the *University Thesaurus* is far outweighed by the extremely tangential nature of many of the words resulting in the output categories.

For example, when we were working with the word DEAD the number of words in our list based upon *Roget's International Thesaurus* was 268; the number of words from *Roget's University Thesaurus* was 2452. Given these numerically disparate lists as input, VIA's output, of course, reflected the disparity. The total number of words in the output list for the computer run using the *International Thesaurus* was 39; the computer run using

[6] Joan Peters has given us much valuable assistance in this phase of the project.

the *University Thesaurus* generated 229. Although the output produced by the use of the *University Thesaurus* was much more extensive, a great many words in that output seemed inappropriate in a list of words clustered around the word DEAD. Examples of such words are GO, METRIC, ANALYSIS, PRACTICE, PROPORTION, INDEX, LINE, LINES, and ESTIMATE. In our judgment—and the judgments used at this early stage of research are not based upon explicit, weighted criteria—four of the words in the output list based on an input from the *International Thesaurus* seemed irrelevant and two seemed only marginally relevant. By contrast, in our judgment, 110 words in the output produced by the *University Thesaurus* seemed irrelevant to some degree. So that, for the *University Thesaurus,* just under half of the total number of entries seemed questionable.

We have made similar comparisons using thirteen other words: DECLINE, DECREE, DEFLECT, DELIBERATE, DELUSION, DEVELOP, DEVISE, DISARRAY, DISASTER, DISTRICT, DOUBLE, and DRILL. These other comparisons amplify the comments based upon the research using DEAD. Our conclusion was that the output based upon *Roget's International Thesaurus* is perhaps a little spare, but relatively free of irrelevant terms. The disadvantage of that slight spareness is obviated, in large measure, by VIA's operating procedure. Since VIA cross-links lists of words, the enrichment otherwise given by a bigger indexing net (such as that in the *University Thesaurus*) is available while, at the same time, the provenance of the enrichment is clearly shown. Our next step, if we were to pursue further the possible use of the *International Thesaurus* would be the very interesting project of exploring in some detail the biases of its indexing and internal structure. To accomplish this task without prohibitive cost, we will need a machine-readable version of the *International Thesaurus.*

Webster's Dictionary of Synonyms, per se, produced too little input and consequently too little output to be of interest for our purposes. However, if that dictionary were available in computer accessible form, it might be possible to move around within the dictionary and collect a good number of acccurately linked words for any given category. It may be that a large dictionary also would be usable as a base for thesaurus construction, and we have begun to think along those lines as well.

In addition to the comparison of thesauri, we have begun to think how we might enlarge any given thesaurus and, on something of another tack, how we might use the structure of a thesaurus to help resolve semantic ambiguities. Our work in enlarging a thesaurus is still in the thinking stage. But, in general, we are curious to see whether a word that does not appear in the thesaurus might, when appropriate, be added to the

thesaurus through an examination of the categories into which fall words near it in the text that do appear in the thesaurus. For example, if the word "ICBM" occurred a great many times followed by the word missile, and if "ICBM" did not appear in the thesaurus but missile did, one might feel safe in letting the computer program add "ICBM" to a category including missile; or, perhaps better, one might put "ICBM" in the output list containing missile with a notation that "ICBM's" assignment to this category is tentative, letting the researcher decide whether it should be entered permanently in that category in the thesaurus. If this sort of procedure is indeed helpful, there is still the problem of automating the initial selection of possible categories upon which the researcher would base his final allocation of words to specific categories.

Regarding the use of a thesaurus or a thesauruslike set of categories to resolve ambiguities, we are largely concerned with semantic rather than syntactic ambiguity. It is obvious that semantic and syntactic ambiguities do not always coincide. The word CONTROL, for example, does not differ significantly in meaning when its syntactic usage shifts from nominal to verbal, but the word STATE does. On the other hand, the word STATE has several different meanings even when used as a noun. One of the tacks that we have taken is to see whether thesaurus categories for words modifying, for example, STATE helped to clear STATE's ambiguity. This lead looks interesting enough to follow further and, if it continues to look fruitful, once again we shall begin to think about ways of automating such a procedure. In conjunction with this clarifying of ambiguity, we also think we may learn much from studies of intratextual semantic and syntactic stability—probabilities for consistencies in the opted uses of terms within any given strings.

Currently, we are considerably enlarging VIA's options as to the kinds of categories or word groupings it produces. We are conceptualizing this output in terms of ring structures which not only will show rings of words which are semantically associated but will show the ring of concepts associated with any given word. This part of the program system is being designed by William Buttelmann; in it, searches may be keyed on an individual word or root group, on an individual category, or on all of the categories in which a given word appears. Furthermore, by working back and forth between a thesaurus and the text, this program is going to have the power to show what words within a category an author avoids, or even what categories an author avoids as well as those he chooses. This power means that the researcher, if he chooses, will be able to key searches on words not appearing in the text as well an on those appearing in the text. Or, in fact, he could key a search entirely on words not appearing in the text.

The current version of VIA is being retained with some minor modifications. To indicate its adequacy for use in content analysis, we shall describe and briefly compare the text specific thesauri for *Hamlet* and the RAND translation of *Soviet Military Strategy.*

For *Hamlet,* our "cut-off" point for designation of primary root groups was 10 or more occurrences; this figure resulted in 37 primary root groups in ACT I, 35 in Act II, 48 in Act III, and 24 in each of Acts IV and V. To give these figures relative meaning, it should be noted that there are approximately 1450 content word types in Act I; remember that 37 content word root groups occur 10 or more times. There was a great deal of overlap from one act to the next; with 10 as the threshold, the number of primary root groups for all of *Hamlet* totaled 65. Even some of these words, such as the quantifiers MUCH, MANY, and MOST, can be considered content words only marginally.

A detailed study of output would be beneficial both to the individual who has never read *Hamlet* and to the scholar who knew *Hamlet* very well. For the person *unacquainted* with *Hamlet,* a glance at the primary root groups for Act V would suggest, for example, that a drink or the act of drinking was of some importance in this act. The presence of the word POTION on the list linked to DRINK might imply something unusual about the drink. The researcher would guess that a KING had an important role in this act and would note that the words associated with KING have connotations of a king's court and of his kingdom. Among proper names used in this act, he would find that the names of Hamlet, Horatio, and Laertes qualify as primary words, with Hamlet having the highest frequency. The researcher would notice that the more generic word, MAN, occurs with some frequency and that some of the words associated with it imply familial relationships. If he were to glance at those words in the output which had qualified as primary words in earlier acts, the researcher would discover some of the familial terms FATHER, DAUGHTER, and MOTHER. He would take note of the fact that these familial terms no longer had primacy in the final act (in fact, DAUGHTER does not occur at all) but that the more general, more impersonal terms KING and MAN did. If the word POTION associated with DRINK had overtones of the unusual, the presence of MADNESS as a primary word amplifies them. The word ILL linked to MADNESS may suggest disease, as does the linking of ILL to WELL.[7]

[7] WELL is an example of a word that qualifies as a primary word because it is multifaceted as to function and meaning. In *Hamlet,* WELL is used, among other ways, as an interjection (as in "Well, again"), as an idiom implying some sort of assent or agreement ("Very well"), and as a descriptor of health or condition. Thus, if this primary word group appeared in isolation its implications as to content

Another primary word of possible interest in this act is the word TELL. The researcher would notice that a number of words associated with TELL suggest an emphasis upon speaking or talking. He would note, in fact, that the word SAY also qualified as a primary word in this act and that the word SPEAK has had primacy in an earlier act or acts, as had the word WORDS. There would seem to be an interest in language, especially the spoken language (the word WRITE appears for the first time, and only once, in this final act), as well as perhaps an interest in reporting or relaying something. The latter interest seems especially important in this last act, with the emphasis upon TELL (a glance at the output from earlier acts would show the researcher that TELL had not qualified as a primary word in those acts). The researcher would also notice a premium upon knowledge; the words KNOW, KNOWS, KNOWN, and KNOWING all occur as well as the words LEARNING and UNDERSTAND which are associated with them. The contrasting word IGNORANCE is also present. Although the word DEATH does not quite qualify as a primary word in this act (it occurs eight times), 14 of the words on the computer's thesaurus list which might be associated with death occur in this act. An examination of this list reveals words which definitely imply violence: BLEED, MURD'ROUS, MURTHER, and SLAIN. The word PLOT, which also appears on this list, may provide some insight into the importance of knowledge, of KNOWING. In keeping with the theme of violence, it may be of interest that three of the words associated with LOVE, which does not quite qualify as a primary word, imply violence rather than tenderness. The fourth word, DOTE, also suggests a lack of restraint, although not quite of the same sort as that connoted by DESIRE, HOT, and PASSION. Other words among those which do not qualify as primary in this act but did qualify earlier—and which catch the eye because of the relatively large numbers of words linked to them—include SLEEP, SOUL, HEAVEN, EARTH, HEARERS, and ACT. In this act, SLEEP occurs just once but, even so, there is nothing of restfulness or soothingness in the words associated with it. Instead, there is the violence associated with DEATH which is linked to SLEEP. The presence of words clustered around SOUL, HEAVEN, and EARTH, indicates that *Hamlet* is not simply a description of derring-do for the sake of derring-do. The words FEELING, CONSCIENCE, and HEART are associated with

---

would have to be investigated thoroughly by examination of context. However, because of the presence of MADNESS as a primary word, and the high incidence of words associated with DEATH in this chapter, the presence of ILL as an antonym to WELL would not seem too misleading insofar as a general assessment of content (or "theme" or "tone") was concerned.

SOUL. Words associated with both HEAVEN and EARTH include NATURE, WORLD, STARS, LIFE, and UNNATURAL. When, bearing these groups in mind, the researcher sees that GOD, THINK, and THOUGHT have also qualified as primary words in earlier acts, and that in this act the words BELIEVE and REASON are linked to THINK, he might suspect that there is some emphasis upon abstract speculation and upon speculation about the abstract in this play. Considering the already noted emphasis upon saying and speaking, the presence of HEARERS as a sometime primary word will not surprise the researcher. He might also suspect that ACT, another earlier primary word, has some connection with SPEAK; the word STAGE, which is linked to ACT, would strengthen this suspicion.

High-frequency words which seem unequivocally to be verbs—words such as COME, DO, and GO—would present more difficulty. The researcher might conclude that there is a good deal of coming and going but he would not be able to attach these actions to individuals. If he knew anything about Shakespeare, he would realize that many of the verbal references to coming and going are simply a part of Shakespeare's staging technique. These verbs provide stage directions in the text itself, and also they help to fill intervals during which characters are making entrances and exits. Lacking this information, the researcher would need to turn to the context of the occurrences of the verbs in order to determine who was coming and going and the reasons for the activity.

Even after this quick initial glance at VIA's output from Act V, the researcher would have a good notion of important content, as to both plot and theme, of the fifth act of *Hamlet*. It is in this act that the King (Claudius) successfully plots to kill Hamlet. The murder will occur during a duel, which Hamlet supposes to be a friendly sporting match between Hamlet and Laertes. The tip of Laertes' foil will have been coated with a poison. Although Laertes is an excellent fencer, the King takes the additional precaution of preparing a poison potion (if the researcher's suspicions had been aroused by POTION he could have consulted the complete output for Act V and could have discovered that both POISON'D and POISON occur) in the event that Hamlet should have occasion to drink a toast in victory. In the general carnage at the end of the play, Hamlet, Laertes, and the King are all dispatched by the poisoned foil, and the Queen, in an effort to drink to Hamlet's early success in the duel, dies of the poisoned potion. The final act's emphasis on TELL is clearly pointed out when Hamlet, dying, urges Horatio to "Absent thee from felicity awhile,/ And in this harsh world draw thy breath in pain,/ To tell my story."

A separate chapter would be required to delineate the aspects of VIA's

output on *Hamlet* that would be of interest to the Shakespearean scholar. Suffice it to say that there are shifts in thematic emphasis—such as the decline in familial reference and the increase in more general impersonal reference at the end of the play—which the traditional scholar may not have noticed and may find illuminating or strengthening for this interpretation of *Hamlet.*

As might be supposed, VIA's output for the translation of *Soviet Military Strategy* differs both in kind and detail from that for *Hamlet.* For example, it was possible to use a cut-off point of 50 or more for primary word designations in the first chapter of the Praeger translation of *Soviet Military Strategy.* Unlike *Hamlet,* there were root groups with very high frequency. For example, the root group containing WAR has 293 occurrences; the root group containing MILITARY has 285 occurrences; and the root group containing STRATEGY has 283 occurrences. In Act V of *Hamlet* the two root groups with the highest frequency (containing GO and KNOW) have only 32 and 30 occurrences, respectively. Because of the very high occurrence of certain root groups in *Soviet Military Strategy,* a reading of the chapter from which they are taken would make the information they convey seem, for the most part, obvious. Thus, VIA's output, using a high cut-off point, for *Soviet Military Strategy* would be of value to the information analyst or content analyst who had no knowledge of the book. Nevertheless, even though the high cut-off point was used, the information conveyed by VIA is more subtle than one might at first suppose. For example, the sublists under WAR include the word AVIATION as well as the word BALLISTICS. The latter might bring missiles to mind, and a glance at the sublist under ARMS would reveal the presence of MISSILE, ROCKET, and NUCLEAR. The information analyst should also notice that FIREARM and GUN appear on that list and that GUNPOWDER is on the list linked to WAR. The terms cited in the above two sentences might imply two kinds of strategy: that involving nuclear missiles and that involving conventional weapons. Pursuing this lead further, the investigator could check the complete output and discover that words associated with NUCLEAR occur 30 times and the terms MISSILES and ROCKETS a total of 7 times. On the other hand, words associated with FIREARMS and GUNS occur a total of 14 times, and the word ARMED occurs in the phrase ARMED FORCES 47 times (the latter information can be obtained from a perusal of the indexes for ARMED and FORCES, or from a MAPTEXT[8] printout, or by going directly from the index entries for ARMED to the indicated context). A further investigation of the contexts for the occurrences of

[8] A separate program that provides abstract representation, or graphs, of the text.

ARMED FORCES would show that the phrase seems most often to refer to conventional warfare. Thus insofar as frequency is concerned, conventional warfare receives a greater emphasis than nuclear warfare (approximately 61 to 47), although the latter is by no means ignored. A complete reading of Chapter One shows this picture to be accurate; there is, in fact, an ambivalence toward nuclear warfare, revealed in such passages as the following:

> The country that finds itself in a catastrophic situation as the result of mass nuclear-missile strikes may be forced to surrender even before its armed forces have suffered any decisive defeat. But we must remember that such results can be accomplished only by means of force, by means of armed conflict.[9]

Or, again, having earlier noted that nuclear warfare had completely changed military strategy, this statement is made: "Modern war involves mass armies . . ."[10]

Hamlet and Sokolovsky are certainly very different types of text, and the researcher's interest is not necessarily the same in both cases. For example, no one is likely to use VIA on *Hamlet* to determine whether he wants to read *Hamlet*. Instead, VIA might be used to provide a more detailed basis for interpretation or to make a more general study of the genre of tragedy. With respect to documents like Sokolovsky's, VIA could be used both for gross information retrieval purposes (that is: Should the document be designated for close examination?) and for close analysis. The man-machine interaction which VIA provides can be used both to guide close reading to important sections of a document and to enrich comprehension of that reading so as, for example, to speed effective response to emergency diplomatic communications from other states. For this range of purposes, VIA seems to show considerable promise.

It is through the revelation of content detail and its interconnections within a work that VIA achieves something of a fusion of thematic analysis with idea analysis. The elaboration of an associational structure for terms provides us with some of the conceptual understanding of a work which otherwise would depend on a detailed, human reading. That is, we seem to be finding some of the meaning that eludes nonassociational content analysis by developing through VIA, a structure of intratextual content term relationships somewhat to replace syntactical structures, which content analysis does not reflect. An enriching of theme detail analysis begins to give us some of the understanding of a text which we generally have

[9] Praeger, translation, p. 107 (see footnote 4).
[10] Praeger, translation, p. 36 (see footnote 4).

expected to find only when we had, in a traditional way, dealt with concepts by examining themes within their syntactic structures.

In conclusion, then, we are sufficiently encouraged by our content analysis procedures to develop VIA further, along the lines described, including the ring structure analysis option, as well as to continue work on the sizable ancillary research project—thesaurus analysis and construction—necessary to make our content analysis procedure more completely automatic.