

THEMATIC STRUCTURE AND COMPLEXITY

Introduction

PRIOR TO THIS CENTURY, THEMATIC criticism dealt primarily with the major thesis or theme *per se* of a literary work. More recently, thematic analysis has concentrated on the subthemes or theme components—image, motif, character, or plot episode—that support or comprise the major thesis.¹ This shift from thematic explication, often didactic, to structural analysis of thematic components is most apparent in two major areas of critical development: structuralism and computer criticism.

Structural considerations of thematic structure today appear in a variety of disciplines; however, almost all contemporary approaches are indebted to the folklorists working during the first half of this century. Emphasizing the motif aspect of thematic structure, modern folklore studies began with the Finnish School and their classification of tales, variants, and motifs epitomized by Aarne's index.² A widely known application of the Finnish Historical-Geographic method to a specific group of tales is Thompson's study of the Star-Husband tales.³ In a corpus of some 86 variations, Thompson noted fourteen basic actions or motifs. He then translated each tale into a symbolic representation using these fourteen basic motifs in their variant forms, and through laborious tabulations and groupings he was able to trace the probable geographical and literary development of the tale. The tabular approach was expanded by Vladimir Propp to include not only motifs and their variants but also the sequence or order of these motifs. Propp's "syntagmatic" approach is based on a concept of transformation that foreshadows more recent developments in linguistics: "The entire story of fairy tales ought to be examined

as a *chain* of variants. Were we able to unfold the pictures of transformations, it would be possible to satisfy ourselves that all of the tales given can be morphologically deduced . . . from that form which we are inclined to consider as basic."⁴ Calling for an "autonomous text theory," Lubomír Doležel has extended the study of thematic structure one step further by considering the symbolic representation of motif sequence as an independent semiotic structure to be analyzed and characterized in relation to the larger, "metastructure" that constitutes the interpretable dimension of literature.⁵ Doležel, after translating the action of Hemingway's "The Killers" into "dynamic motifs," denotes patterns in the symbol sequence, itself, defined by repetition, symmetry, contrariety, and equivalence. Literary interpretation, then, becomes a two step process beginning with delineation of structural patterns in the textual sequence which, in turn, produce and support abstract, metastructural interpretations.

The second major area where techniques for analyzing thematic structure have recently been developed is computer criticism. Folklorists have usually concentrated on motifs; computer criticism, in general, has concentrated more on imagistic aspects of thematic structure.⁶ While not involving a computer, Caroline Spurgeon's analysis of Shakespeare's imagery has greatly influenced subsequent computer studies.⁷ By systematically categorizing Shakespeare's imagery into thematic groups and tabulating the frequency of the groups in each play, she was able to show the fundamental relation between close, textual features and the larger, abstract dimensions of the plays. More recently, Howard Iker and Norman Harway have modified her approach by using the computer to assist with the categorization process, itself.⁸ They determine categories or themes contextually by having the computer statistically determine clusters of words that consistently occur close to one another in the text. To assist the critic in gaining perspective into patterns of interaction among groups or categories of words, Sally Y. Sedelow has developed a computer program, called MAPTEXT, that reproduces the text in condensed paragraph structure by substituting a single symbol (an "A," a "B," etc.) for each appearance of a word in a specified category and a period for all other words not included in one of the

categories under consideration.⁹ Examining the text in this compressed format, the critic is readily able to discern patterns of interaction among the categories or themes designated.

Both areas display a similar pattern of development: a major concern of earliest studies was the matter of categorization; this was followed by tabulation and numerical inference; the third stage of development was marked by contextual analysis and display of the linear sequence of theme components. While this brief review is obviously not complete,¹⁰ it does outline in skeletal form the growing tradition of thematic analysis on which the techniques I shall describe below are based. Collectively these new techniques represent an extension of techniques developed earlier to analyze and describe sequential patterns among themes or theme components. It will be apparent that they are all suitable for computational analysis, but the reader should realize that anything the computer can count, he or she can count; thus, they are general, and the critic does not have to confront the computer to use them.

Operative View of a Text

If we take a simple, mechanical view of a text, we may view it as a sequence of symbols: the characters of the alphabet and some set of special characters, such as punctuation marks. This sequence of symbols, which is more often and more profitably considered a sequence of word units, can be thought of as one long string, stretching from the first word in the text to the last, in ticker tape fashion; through an "accident" of history we generally encounter that long sequence divided into line-length units that are reproduced vertically on the page.¹¹ In considering thematic structure I shall rely rather heavily on the latter concept of a text; however, the reader should bear in mind that this perspective of the text must always be accompanied by other, overlapping perspectives. As we proceed through the text, from left to right, we also proceed through the fictive world that the words generate in our minds, often developed with a chronology that runs parallel to the textual sequence. A third level that grows out of the

textual and fictive levels is the abstract or interpretive level. It is the delineation of this third level that has become the major task of literary criticism. In examining thematic structure I shall try to show that patterns that exist in the operative, textual level support directly and sometimes constitute the patterns in the abstract, interpretive level.

Having distinguished between the mechanical notion of text as a sequence of words and the impressions those words produce in the reader, we may similarly distinguish between the textual manifestation of a concept, such as theme or image, and the concept itself. The critic may follow the lead of many folklorists and consider a theme to be a set of basic incidents or actions carried by the plot; following the lead of a number of recent literary scholars, the critic may choose to associate theme with a group of images suggesting a concept he feels is basic to an understanding of the text; or he may choose to concentrate on some other literary component. Whatever the critic's choice, in considering a specific literary work he must eventually come to the text and state that a theme is carried by a specific group of words, phrases, or sentences. Thus, by considering thematic structure and complexity operatively—as the relations among groups or collections of textual items—we may accommodate critics who use the term in widely different senses. In the discussion that follows I have used theme to mean a group of single words, usually images or words that might arouse any of the senses, that suggest the same basic concept. A theme may, thus, be thought of as a thesaural-like category of related words found in the text. For example, from an earlier study of James Joyce's *A Portrait of the Artist as a Young Man* I have included material related to four themes: fire/heat, water/cold, religion, and language. The theme, fire/heat, is carried by words such as blaze, burn, burned, burning, fire, hot, etc.¹² Another critic may wish to extend his definition to include phrases such as "the burning bush" or "the flaming radical": he may even wish to include whole scenes developed over several sentences or even paragraphs. From a structural point of view, all of these definitions can be used so long as the theme can be specified as a collection of specific textual items, regardless of whether those items are words, phrases, sentences, or combinations of all three.¹³

Returning to the notion of a text as a long sequence of words, we may consider the location of a word as its position in that sequence (word 9328, for example). The location of a phrase might be the location of its first word or, perhaps, its middle word; similarly we can designate the location of sentences and even longer textual units. In this manner, the location in a text where a particular theme appears can be specified. Instead of considering only the prevalence or frequency of a theme or group of themes for an entire text, we may now consider the importance of various themes in specific portions of the text by producing distributions of themes over the text. To do this, the text must be divided into portions of equal length and the frequency of occurrence for the theme under consideration tabulated for each such portion. A suitable unit might be a page; I prefer to use units measured in words (500 words is usually a little more than a page) so that thematic density in one section can be compared accurately with density in another.

Having tabulated the collective frequency for a theme in each textual unit, the critic may display the theme as a graph or chart. That is, if the horizontal line of the graph is used to represent the text (running from left to right) and the vertical dimension represents frequency, thematic density can be represented as a sequence of points, one for each text section, that rises and falls as thematic density rises and falls. Figure 1 illustrates this type of graph for the collection of fire and heat images occurring in James Joyce's *A Portrait of the Artist as a Young Man*. There one can see the fairly heavy concentration of this image group in Chapter 1; however, the graph is particularly useful for illustrating the degree of concentration present in Chapter 3 (during the hell fire and damnation sermons of the retreat) compared with all other sections.

While individual graphs are useful for comparing thematic density in various sections of text, graphs of several different themes can be compared to determine patterns of thematic association or interrelation. Figures 2, 3, and 4 represent thematic groups of images denoting religion, language, and water/cold, respectively. By comparing these four themes, we can see the rather parallel use of fire and water imagery in Chapter 1 when they are

closely associated as opposites in Stephen's mind: water standing indelibly for his traumatic fall into the ditch while fire represents its opposite, the security of home centered in the literal hearth fire. After Chapter 1, the two themes diverge with fire peaking dramatically in Chapter 3 and water peaking in Chapter 4. The implications of both are clarified by the religion and language representations. By Chapter 3, fire has shifted away from the positive connotations present in Chapter 1 in its close association with religion in the form of the fires of hell. Similarly, water imagery assumes entirely different associations by Chapter 4 when, in its close relation with language, it comes to suggest the mind of the artist, particularly his subconscious, in conjunction with epiphany or the esthetic experience. Used in this manner, distributions of a number of themes may verify and demonstrate comprehensive critical interpretations more clearly than is possible through recall and narrative, alone.

Thematic Transitions

Another important aspect of thematic structure is the pattern of transitions from theme to theme. In some works the progression from image to succeeding image or theme to theme carries considerable significance. For example, Joyce's *Portrait* is primarily concerned with the developing personality of Stephen Dedalus. All scenes in the novel are developed from Stephen's perceptual viewpoint, manifest in specific, sensual images. These components of experience, however, are not simply absorbed by Stephen but are fitted into a developing structure of associative relations with earlier images and experiences. Glimpses into Stephen's mind are usually in the form of references to what an image reminds him of or in the form of free associations. Paths from image to image that appear idiosyncratic late in the novel usually have their basis in individual links or associations that were formed in earlier experiences. Thus, the underlying structure of associations among images that defines, to a great extent, Stephen's fictive subconscious can be determined by noting patterns in the progression from image to image that occur in the text.

Other discipline areas have extensively employed diagrams that trace progressions from one state to another. One notable instance is automata theory in which a theoretical computing machine is defined mathematically by a set of states and the rules that govern transitions among them. Each state can be represented by a point and the possible transitions can be represented by lines between the appropriate points. State diagrams of this sort can be used to represent transitions among images or themes: in this case an image, represented by a point, would be regarded as a state and the relation of textual proximity would be a line or path between two points/images.

I have selected the set of words I feel are images in the first section of *Portrait* and traced this progression in Figure 5. The correspondence between the text and the figures can be noted by starting at the top, with the *moocow* coming down the *road* and meeting baby *tuckoo*; when an image is repeated, the diagram returns to the position representing that image and begins a new path. This can be seen with regard to *father* who tells a *story* and has a *hairy face* or *Betty Byrne* who sells *lemon platt* and sings a *song*. If we consider the entire figure we can note three major subsections—the early experience concerning Mr. Dedalus and Betty Byrne, the bed wetting episode, and the experience centering in Stephen's cowering under the table—represented by breaks in the sequence of connecting lines. The figure shows that these separate, very early childhood experiences are related through repeated references to songs and to Stephen's parents. It is primarily through associative patterns such as these that episodes that at times seem almost randomly selected are gathered into a coherent narrative structure.

While transition diagrams are useful for noting associative relations among individual images, they are impractical to produce for long texts; Figure 5 represents a text section of less than a page! Part of this problem of proliferation can be solved by concentrating on thematic groups of images, such as those discussed above in relation to linear distributions, and denoting transitions among significant "clusters" of images rather than individual images. That is, one may consider a cluster to be a group of, say, three images from the same theme that occur within

100 words of one another. This cluster can be considered to be located at the first image of the group, the middle image, or some reasonable position in relation to the individual images comprising the group. Transitions can then be noted between significant clusters, emphasizing major thematic developments and reducing the number of states and paths. In another study (with Professors Bruce Rosenberg and Robert Brubaker) I have produced diagrams of this sort shown in Figures 6, 7, 8, and 9.¹⁴

The texts on which these diagrams are based are four sermons preached by the Rev. Rubin Lacy in Southern California during 1967 and 1968. Diagrams 6 and 7 represent the thematic structure of two different performances, delivered a year apart, of the same basic text, built around an extended metaphor relating a deck of cards to various Biblical numerological allusions. The second pair of sermons, also delivered approximately a year apart, is based on the Twenty-Third Psalm. Both pairs contain one performance that is markedly more successful than the other (DOC I-Fig. 6 and 23 II-Fig. 9 represent the two superior performances). However, the transcriber's notes and the tape recordings suggest that DOC I was slightly superior to 23 II which was delivered to a small, and hence voluble, audience. Of the two poorer performances, DOC II seems to have been a complete flop while 23 I was salvaged in part by Lacy's lapsing into a set piece, out of context, based on the Four Horsemen of the Apocalypse. Consequently, we feel that we can rank the sermons in order of esthetic quality or degree of success in the following order:

DOC I
23 II
23 I
DOC II

As indicated above, the difference between 23 II and 23 I is much more substantial than the other adjacent pairs in the sequence.

By visually inspecting these diagrams, we can note several features. The thematic structures for both DOC I and DOC II can be divided into two major clusters. The first cluster developed in both sermons is built around an axis between two themes, designated negative comments and auralty; the second cluster is centered

in allusions to a deck of cards.¹⁵ With DOC I (Fig. 6) the shift from the first cluster to the second comes almost exactly in the middle of the sermon: transition 17 out of a total of 32 transitions. In DOC II (Fig. 7) the shift comes two-thirds of the way through the sermon: transition 25 out of a total of 36 transitions. These thematic subunits are independent except for single shifts back from the cards cluster to the negative comment/aurality cluster in both sermons (transitions 26 and 35, respectively). The perfectly balanced structure of DOC I has an intuitive appeal because of its openness and symmetry that is missing from the more lop-sided structure of DOC II where the negative comment/aurality substructure involves more themes that are introduced and then dropped without development or integration into the main thematic structure of the text. It may be interesting to note that DOC I (Fig. 6) is the sermon we judged to be the best of the four while DOC II (Fig. 7) was judged the worst.

Turning to 23 I and 23 II (Figures 8 and 9, respectively), we can note similar features. That is, 23 II displays the same balance and symmetry present in DOC I. All major thematic developments are oriented around Biblical names: first with relation to negative comment, then aurality, and finally with animals. The sermon ends with these three subthemes drawn together (see transitions 23-30) and related to the church, *per se*. This pattern of introducing and developing thematic sub-units independent of one another before bringing them together in the conclusion is an example of pure, classical rhetorical form.

The only word I can use to describe the structure of 23 I is *cluttered*. It has half again as many paths as any other sermon. But the casual, unorganized way one theme is related to another on a singular basis rather than as part of a coherent subunit or cluster is a more telling factor. Rather than developing into clusters involving close, repeated interactions of a group of themes as in 23 I, themes of 23 II are related to a bewildering assortment of other themes. The performance is too short for a coherent pattern to evolve out of this high degree of thematic activity. One might infer that the audience was left bewildered and confused as to the content or inter-relation of ideas intended.

In summary, these diagrams reveal that two sermons (DOC I and 23 II) display thematic structures that are built from subunits

or clusters in a balanced, coherent way. DOC II and 23 I, on the other hand, display varying degrees of distortion and apparent randomness. It may be interesting to note that the first pair (DOC I and 23 II) was the pair judged by independent factors to be performances clearly superior to the second pair (DOC II and 23 I). Between the two members of both the superior pair and the inferior pair, it is difficult to make judgments. DOC I has an appealing balance in its two part structure; on the other hand, the three subdevelopments that are drawn together at the end of 23 II are more coherent and more tightly integrated. Between DOC II and 23 I, we must choose between the distortion of an obvious pattern (DOC II) and virtual randomness (23 I). However, there can be no confusion between the superior pair and the inferior pair.

While I believe the observations stated above are valid, they are precisely that—observations. It would be helpful if we could look a little closer at these structures to determine exactly what features suggest “balance,” “simplicity,” or “clutter.” If these features can be determined, inferences drawn from the diagrams could be placed on a firmer basis. In instances where the critic’s purpose is analytic as opposed to demonstrative, he may be able to bypass the rather laborious stage of drawing the state diagrams and go directly from data to feature and inference. To illustrate these concepts, I shall concentrate on diagrams 8 and 9, representing the two sermons based on the 23rd Psalm.

Features contributing to a simplicity/complexity impression that are immediately apparent are the number of themes represented in the diagram and the number of transitions or paths among them. 23 II (Fig. 9), which I described earlier as “classical” in its balance and form, contains eleven themes or points with thirty paths among them. 23 I (Fig. 8), on the other hand, contains sixteen themes with over fifty paths among them. This same relation holds for the Deck of Cards pair: DOC I (with the balanced structure) contains twelve themes and thirty-one paths while DOC II (the less balanced structure) contains fourteen themes and thirty-six paths. For this primitive, folk genre, greater thematic complexity may detract rather than contribute to a performance—a relation that might be reversed for a longer, more

sophisticated genre. Both the number of themes and the number of transitions among them can be determined directly from the data on which the state diagrams are based and used as general indications of complexity.

Complexity, however, is not limited to just the number of points and paths of a diagram; it is also influenced by the particular paths or relations that are represented. For example, in Figure 8, the diagram is distorted by the *particular* pattern of interrelation between Biblical names and aurality and between the negative comments and aurality, among others. Since the distances between each point or theme connected by a path should ideally be the same, the diagram cannot accurately reflect the relative "closeness" of these two themes while depicting the relations among physical geography, life/death, King James lexicon, etc. Specifically, to "pull" aurality closer to Biblical names would "bend" or distort the paths among the other themes just listed; if they were to remain rigid, these themes or points would be forced up or down, out of the plane of the diagram. If the degree of bend or stress necessary to force these points into the distance relations inherent in the data could be measured, this value would be quite useful as an indication of the relative complexity of the particular patterns or structure of relations among themes. Since it would be affected, but not completely determined, by the number of points and paths of the diagram, it should be a more comprehensive indicator of thematic complexity.

The critic has at his disposal (with, perhaps, a little assistance) an analytic tool, called nonmetric multidimensional scaling, that produces such a measure. Available at many academic computing centers, this program begins with a table of distance relations between points—themes in this case—and attempts to find locations for the points so that the distance relations among these corresponding locations reflect the same distances in the table. (Distance, here, means the minimum number of paths or transitions between two points/themes.) If, instead of themes, we were dealing with coins (say, a nickel, dime, and a quarter) the program could be used to find spots on the floor where we could place the coins so that the distances among them are proportional to the distances represented in the table of data. In addition to finding the

positions, or coordinates, for the points, the scaling program also produces a value, called the stress, that indicates how much distortion or "bend" is present in the solution. This mathematically determined stress value corresponds to the intuitive concept of stress discussed above in relation to complexity.

Before looking at an actual application of this technique we need one other concept: dimensionality. The state diagrams, drawn on a flat piece of paper, are two-dimensional representations of the distance relations among themes. I suggested earlier that certain configurations, such as Figure 8, could be represented more accurately if the paths could "bend" up or down from the plane into the third dimension. Most nonmetric multi-dimensional scaling programs allow the critic to indicate whether he wants the points located in two dimensions, three dimensions, or, analogously, even higher dimensions. In general, the higher the dimensionality of the space in which the points are fitted the more accurately the distance relations can be represented and, hence, the lower the stress value. For most scaling programs, the stress value will range from 0.0 to 1.0 with a stress of 0.15 considered a "reasonable" fit. Thus, to use the concept of stress as an indicator of structural complexity we may compare the relative stresses for two dimensional (corresponding to the flat, state diagram) solutions for various thematic structures; but we may also use the program to find higher dimensional solutions and note at what dimension the stress falls below 0.15 or some other reasonable measure of distortion. In this way, dimensionality, itself, becomes an indication of complexity.

Figure 10 shows the relative stress values for all four sermons determined for solutions of dimensionality six down to dimensionality one. We can note immediately that at dimensionality two the stress values for DOC I and 23 II are quite low; these were the thematic structures described as "simple," "balanced," and "coherent." The stress values for DOC II and 23 I, on the other hand, are considerably higher; these sermons were described as "unbalanced," "distorted," and "cluttered." Little can be inferred by the slight difference between the 0.122 and 0.150 values for the first pair; but when the stress for one text is several times that of another, we may safely say that there is a significant

difference in complexity between the two. This impression is further confirmed when we note that the stress values for the pair of superior performances fall below the "reasonable," 0.15 value, at dimension 2; the two poor sermons don't reach this level until dimension 5.

Since the primary purpose of the scaling program is to locate points in a space that reflect the distance relations comprising the data with the stress value computed as an indication of the degree of success, it can also be used to isolate closely interrelated groups or clusters of themes. In Figure 6 this tendency of themes to join into independent sub-units is particularly apparent in the negative comment/aurality cluster early in the sermon followed by the later cluster involving card images and Biblical references. The two dimensional scaling of the data for Figure 6 is represented in Figure 11. Although the themes are not laid out in the same way that they appear in Figure 6, the two major clusters are apparent. Negative comment and aurality (points 9 and 1, respectively) appear close to one another and the line between them separates the other members of the cluster (chronological relations, life/death, and condemnation) from the rest of the themes. Similarly, cards and Biblical references mark the second cluster that includes chronological units, physical objects, and positive lexicon. Scriptural lexicon which serves as a transition between the two clusters is, appropriately, closer to the first cluster than the second but outside of it. King James lexicon, which continues the transition, is relatively unrelated to either major cluster. While representations such as Figure 11 may be sufficient to indicate major thematic units, the critic who wishes to produce an actual state diagram can use the configuration determined by the scaling program as an initial layout to be modified in successive drawings for increased legibility.

Summary

I have proposed several techniques for analyzing and representing thematic structures and complexity that build on a methodology that has been developing for a century. Since these

techniques are operatively defined, I began by considering a text to be a linear sequence of words and a theme to be a category of words or phrases (subsequences) that appear in that sequence. From this basis, the concept of distribution for a theme arises in which the frequency of occurrences for a category is determined for equal length subsections of the sequence (500 words, etc.). These diagrams allow comparison of various themes over the text, facilitating studies seeking thematic interaction. However, associative structure can be approached more directly by noting locations where concentrations of a theme occur in the text and tracing the sequence of transitions from theme to theme. The structure or pattern of these relations can be represented by state diagrams in which there is a point for each theme and the transition from one theme to another indicated by a line or path. A measure of the complexity of the thematic relations represented can, in turn, be derived from the concept of stress produced by nonmetric multidimensional scaling programs. In addition to providing a quantitative indication of complexity, the program can also be used to determine clusters of highly interrelated themes. While these techniques may be most applicable for computational analyses, they are general and most can be applied by hand. They are presented with the hope that they may assist the critic in extending the comprehensiveness of his analysis and enable him to portray his results more clearly and more dramatically.

NOTES

¹As with most terms fundamental to literary analysis, *theme* is used to denote a wide range of literary phenomena. Folklorists usually associate theme with a series of related actions or motifs; Thrall, Hibbard, and Holman, in *A Handbook of Literature* (New York, 1960), define theme as "The abstract concept which is made concrete through its representation in person, action, and image in the work" (486). Under this latter definition, a group of individual images that suggest the same basic concept could also be considered a theme. I shall attempt to bypass the problem of multiplicity of meanings for basic literary terms by distinguishing between conceptual definitions (such as that cited above) and practical or operative definitions (such as a list of words or phrases) that might be used for studying a specific text. By separating concept from the application of concept we may con-

sider analytic techniques that can be used by critics with widely different perspectives and purposes. See discussion entitled "Operative View of a Text," below.

²A. Aarne, *Verzeichnis der Marchentypen*, Folklore Fellows Communications, No. 3 (Helsinki, 1911).

³Stith Thompson, "The Star-Husband Tale," Alan Dundes, ed., *The Study of Folklore* (Englewood Cliffs, 1965), pp. 416-59.

⁴V. Propp, *Morphology of the Folktale*, trans. Lawrence Scott, Second English Edition (Austin, 1968), p. 114.

⁵Lubomír Doležal, "Toward a Structural Theory of Content in Prose Fiction," Seymour Chatman, ed., *Literary Style: A Symposium* (Oxford, 1971), pp. 95-110.

⁶There have been, of course, folklore studies that have used computers. Alan Dundes ("On Computers and Folktales," *Western Folklore* 24:185-89) has used the computer to generate tales; Bruce Rosenberg and I ("The Computer and the Finnish Historical-Geographical Methods," *Journal of American Folklore*, 87:149-54) have recently developed a series of analytic programs specifically designed to aid studies modelled after the Finnish Historical-Geographic approach.

⁷Caroline Spurgeon, *Shakespeare's Imagery and What it Tells Us* (Boston, 1958).

⁸Howard P. Iker and Norman I. Harway, "A Computer System Approach Towards the Recognition and Analysis of Content," *Computer Studies in the Humanities and Verbal Behavior*, vol. 1, no. 3 (Oct., 1965), 134-54.

⁹Sally Y. Sedelow, et al., *Automated Language Analysis: 1968-69* (Chapel Hill, 1969), pp. 28-31, 231-44.

¹⁰One major approach to thematic analysis, not discussed above but which should be mentioned, is that of Claude Lévi-Strauss. His paradigmatic approach differs substantially in that it emphasizes logical or associative relations across texts rather than relations defined by sequence within a text. Because the techniques I shall describe are based on the latter concept, I shall not discuss this important approach but refer the reader to Lévi-Strauss' excellent introductory essay, "The Structural Study of Myth," which first appeared in *Anthropologie Structurale* (Paris, 1958), pp. 227-55; a slightly revised version, translated by Clare Jacobson and Brooks Schoepf, appears in their edition of *Structural Anthropology* (New York, 1963), pp. 206-31.

¹¹This is more properly a "typesetting" than a "linguistic" view of language since it assumes that the text exists in a physical, printed medium, not in oral speech. Problems of segmentation, so important for considering spoken language, can here be reduced to operative terms. The basic unit is the character or grapheme, the set of which can be enumerated by listing the keys struck on the linotype, typewriter, or keypunch machine used in preparing the text. Larger units, words or morphemes, can also be defined operatively by partitioning the set of characters into the set of syntactic punctuation characters (commas, semicolons, etc.), the blank or space, and the characters of the alphabet; a word, then, becomes the sequence of alphabetic characters bounded by two spaces or by a space and a syntactic mark of punctuation. Viewed in this manner, segmentation is equivalent to Kenneth Pike's notion of particle (item-and-arrangement) as opposed to wave (item-and-process) or field (See Kenneth L. Pike, *Language in Relation to a Unified Theory of the Structure of Human Behavior* (The Hague, 1967), pp. 545-55.

¹²The themes of water/cold and religion were derived analogously in a rather straightforward manner; the theme I designated language, however, may need clarification. Appearing primarily in the latter part of Chapter IV and through most of Chapter V, the theme of Language is associated most strongly with Stephen's growing awareness of the written and spoken word and his ensuing decision to become a writer. The theme, manifest in words such as period, meter, word, language, etc., is most often present in specific discussions of language, *per se*, as with the aesthetic theory Stephen outlines in his discussion with Lynch.

¹³The concept of category has been referred to by a variety of terms in other thematic studies. Aarne used the term *types*; Propp grouped manifestations of motifs under the term *function*; Barthes uses *paradigm* to refer to a group of "reservoir." These and other approaches may employ the structural techniques described below by recognizing the common underlying notion of collection.

¹⁴That study involves an attempt to develop techniques for determining and representing comprehensive narrative structures that indicate degree of esthetic success. We are, thus, extending Parry's and Lord's notion of formula to include any principle of structural organization. While the study focuses at three different levels (global patterns involving the dynamics of the entire text, intermediate patterns that are primarily thematic, and micro patterns involving stress relations in individual lines) references here will be only to intermediate, thematic aspects. For more details of this study see "Rhythm in Speech: The Formulaic Structure of Four Fundamentalist Sermons" forthcoming in *Computer Studies* and "Thematic Structure in Four Fundamentalist Sermons" forthcoming in *Western Folklore*.

For the reader unfamiliar with state diagrams, some instructions on "reading" them may be useful. To trace the progression from theme to theme

linearly through the text, start with the theme marked "START" and follow the path designated "1" to the second theme; then find the path, beginning with that theme, marked "2" that leads to the third theme; etc. Themes that cluster in more than one section of text are represented by multiple paths to the same point; hence some paths will have several sequence numbers designated.

¹⁵In general, I have found it necessary to establish a different set of themes (and the list of words that constitute those themes) for each specific text. The unusual nature of the folk sermon led us to specify as themes concepts that may seem unfamiliar. Because of the extreme importance of the manner of delivery as well as the substance delivered, the preacher's references to his own speech and aural communication in general function as a most important theme. It is carried by words such as *talk* in the phrase "I'm gonna *talk* short tonight," or *hear* in his request for response in the refrain, "Do ya *hear* me?" Negative comment is our rather innocuous label for the preacher's very vigorous threats of damnation. The other themes cited, I think, are fairly apparent.

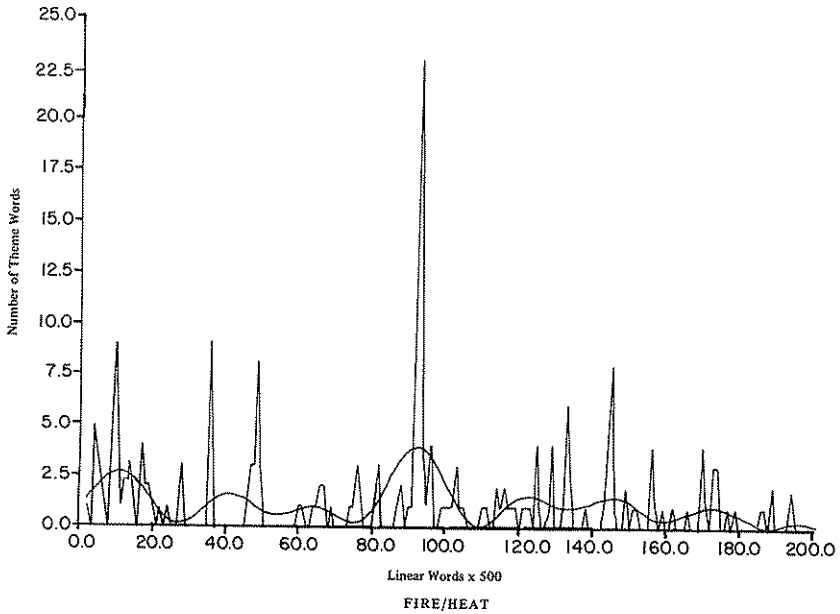


Figure 1

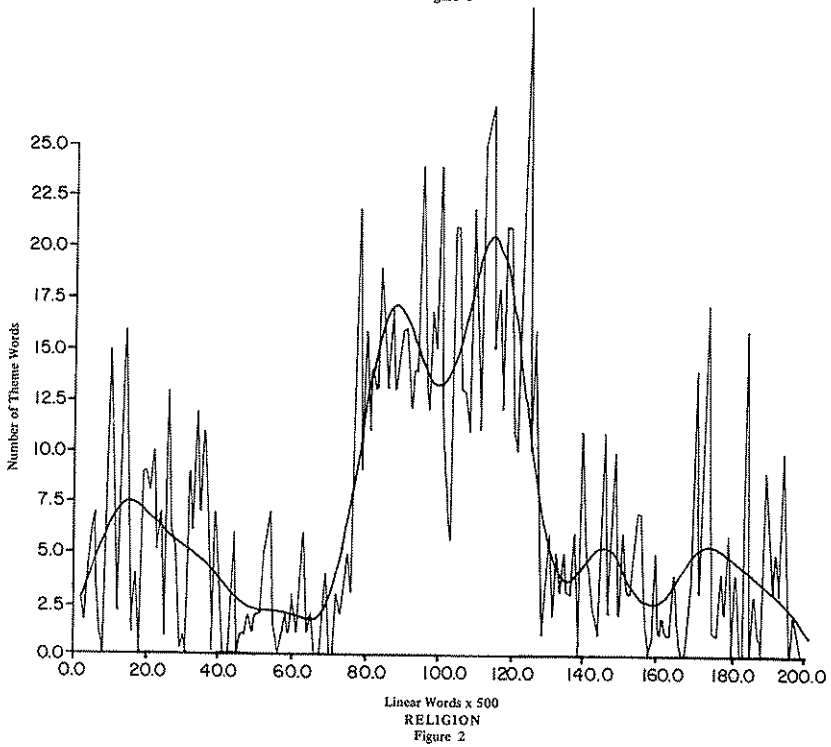
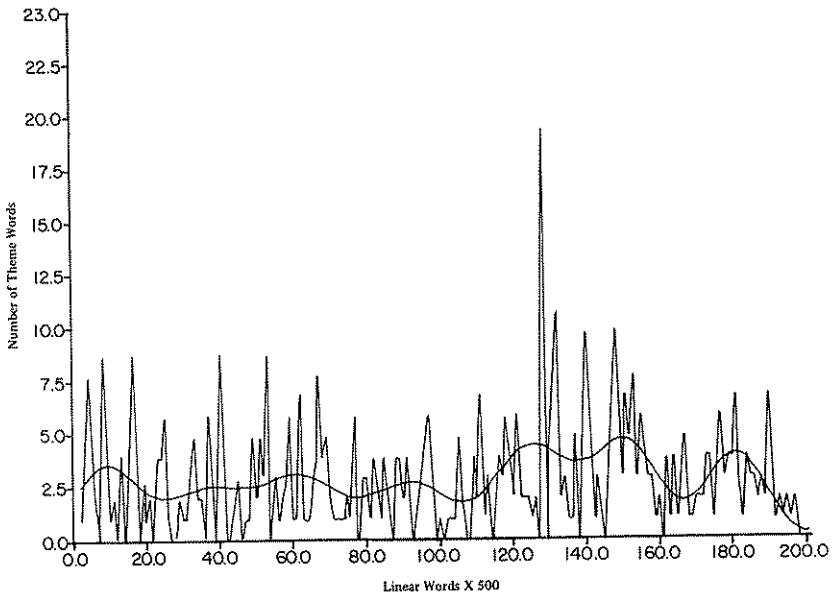
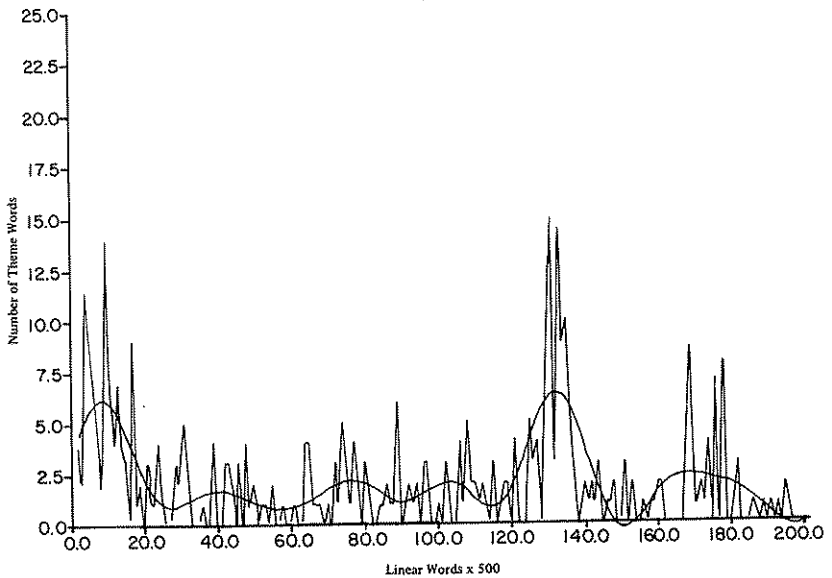


Figure 2



LANGUAGE

Figure 3



WATER/COLD

Figure 4

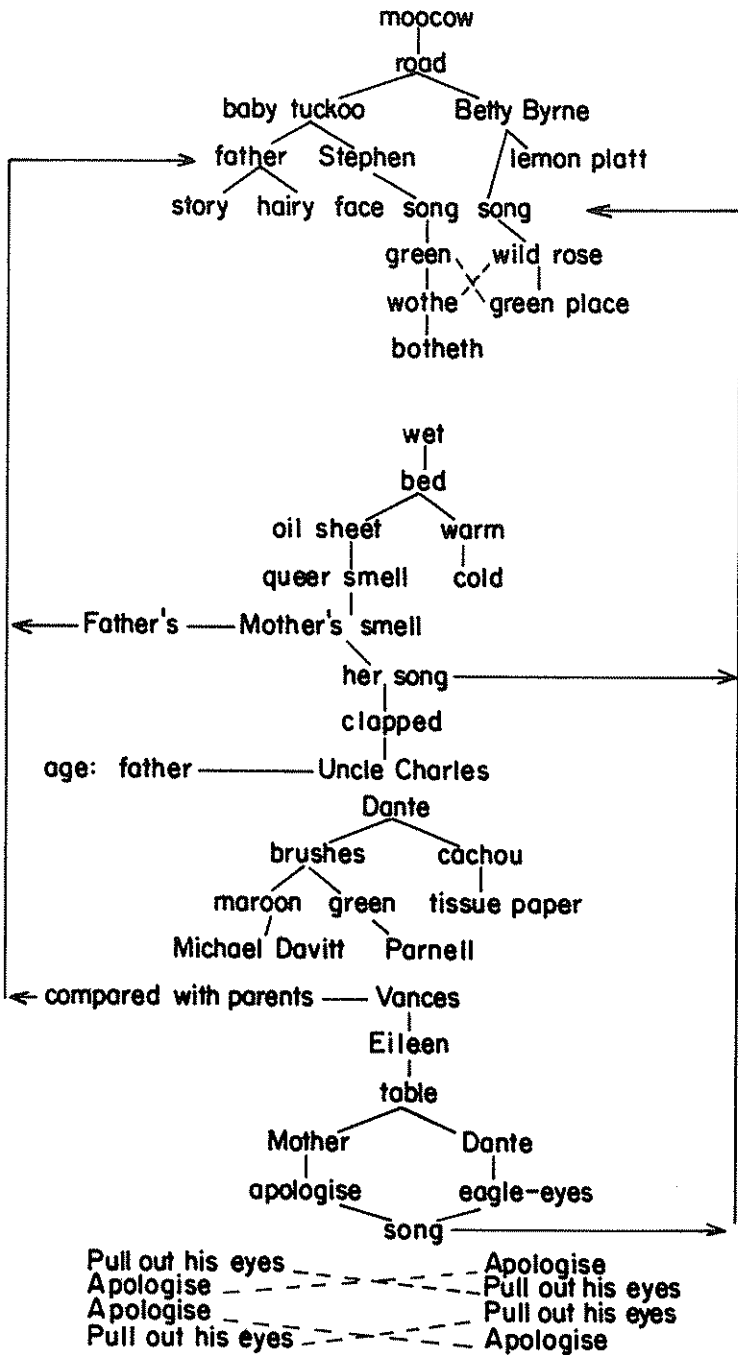
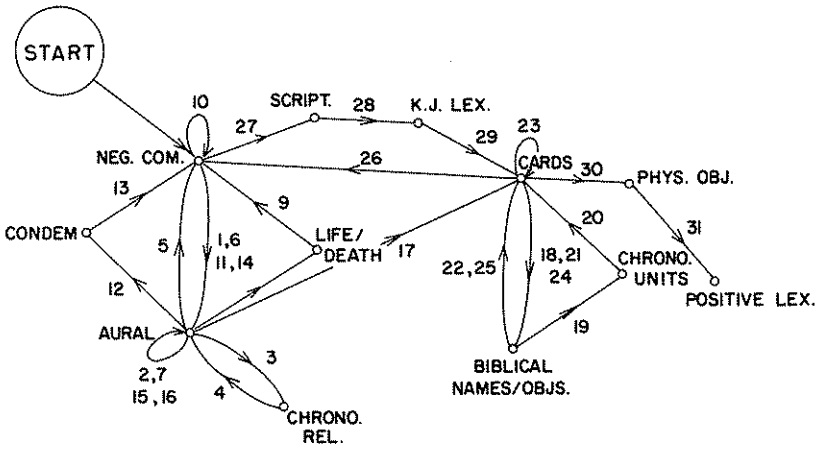
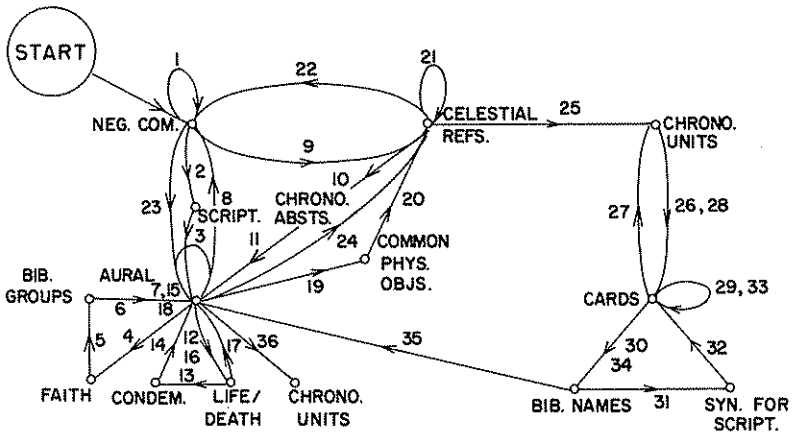


Figure 5



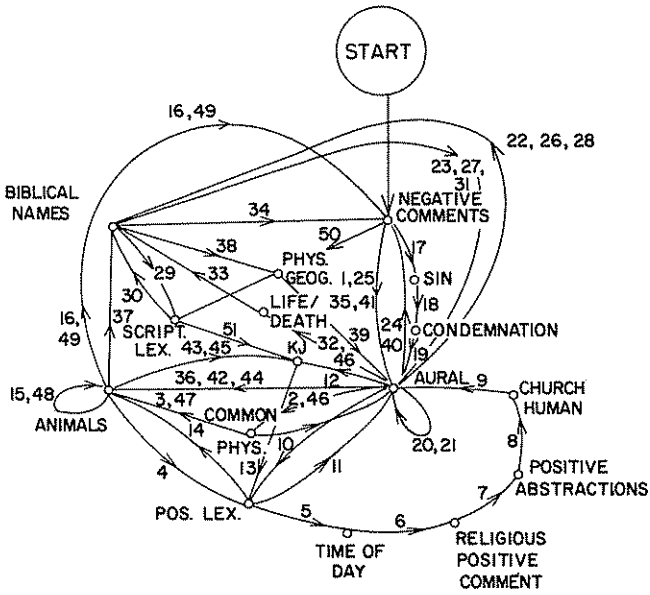
DOC I

Figure 6

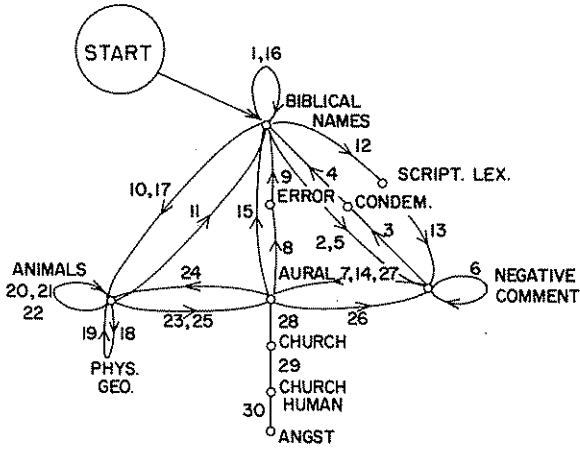


DOC II

Figure 7



23 I
Figure 8



23 II
Figure 9

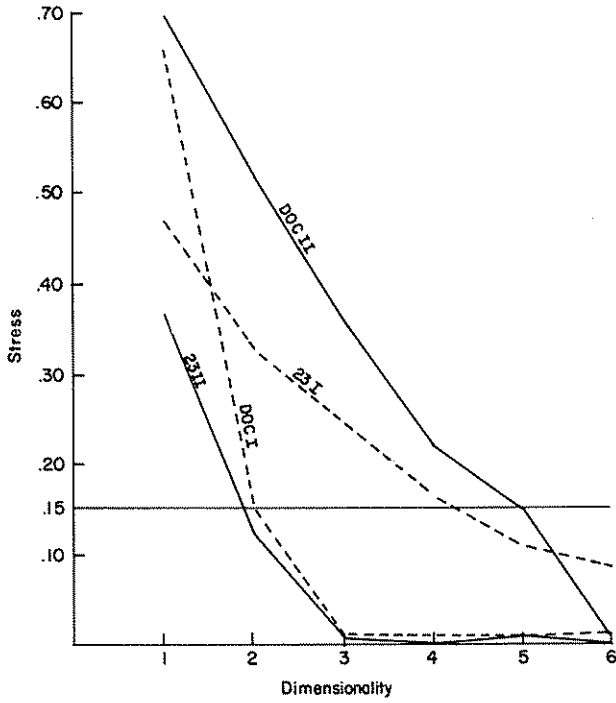


Figure 10

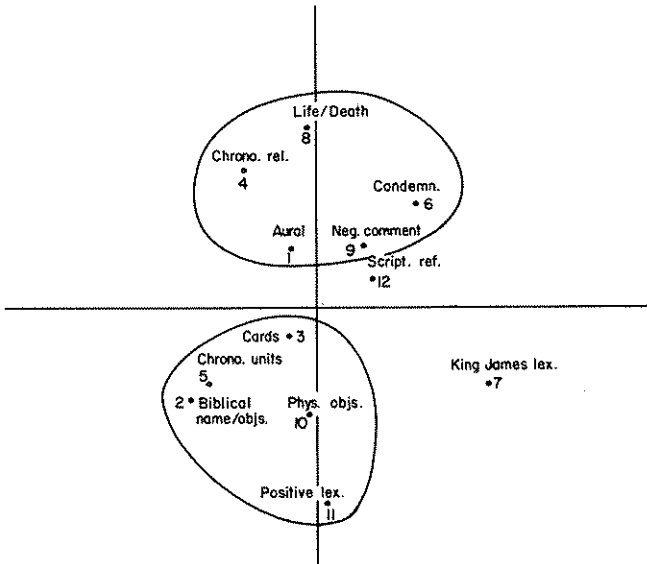


Figure 11