INFORMATION TO USERS

This was produced from a copy of a document sent to us for microfilming. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help you understand markings or notations which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure you of complete continuity.

2. When an image on the film is obliterated with a round black mark it is an indication that the film inspector noticed either blurred copy because of movement during exposure, or duplicate copy. Unless we meant to delete copyrighted materials that should not have been filmed, you will find a good image of the page in the adjacent frame. If copyrighted materials were deleted you will find a target note listing the pages in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed the photographer has followed a definite method in "sectioning" the material. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For any illustrations that cannot be reproduced satisfactorily by xerography, photographic prints can be purchased at additional cost and tipped into your xerographic copy. Requests can be made to our Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases we have filmed the best available copy.
MARKELS, ROBIN BELL

COHESION PATTERNS IN ENGLISH EXPOSITORY PARAGRAPHS

The Ohio State University

University Microfilms International 300 N. Zeeb Road, Ann Arbor, MI 48106

Copyright 1981 by Markels, Robin Bell

All Rights Reserved
PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy. Problems encountered with this document have been identified here with a check mark √.

1. Glossy photographs or pages ☐
2. Colored illustrations, paper or print ☐
3. Photographs with dark background ☐
4. Illustrations are poor copy ☐
5. Pages with black marks, not original copy ☐
6. Print shows through as there is text on both sides of page ☐
7. Indistinct, broken or small print on several pages ☐
8. Print exceeds margin requirements ☐
9. Tightly bound copy with print lost in spine ☐
10. Computer printout pages with indistinct print ☐
11. Page(s) ________ lacking when material received, and not available from school or author.
12. Page(s) ________ seem to be missing in numbering only as text follows.
13. Two pages numbered __________. Text follows.
14. Curling and wrinkled pages ☐
15. Other ____________________________
COHESION PATTERNS
IN ENGLISH EXPOSITORY PARAGRAPHS

DISSertation

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By
Robin Bell Markels, B.A., M.A.

******

The Ohio State University
1981

Reading Committee:
Professor Christian K. Zacher
Professor Arnold M. Zwicky
Professor Sara Garnes

Approved By

Christian K. Zacher
Co-Adviser
Department of English

Arnold M. Zwicky
Co-Adviser
Department of Linguistics
This is for Dudley Hascall
And because of Julian Markels.
VITA

ROBIN BELL MARKELS

DATE OF BIRTH: July 23, 1947

EDUCATION

M.A., 1972, The Ohio State University, Columbus, Ohio
Thesis: Form and Content in Virginia Woolf's
Between the Acts

B.A., 1969, Wittenberg University, Springfield, Ohio, magna cum laude

EXPERIENCE

Lecturer and Teaching Associate in the Department
of English, Fall, 1970 to Present, The Ohio State University

AREA OF INTEREST

Discourse Analysis and Rhetoric
# TABLE OF CONTENTS

**DEDICATION**  
11

**VITA**  
iii

**TABLE OF CONTENTS**  
iv

**INTRODUCTION**  
2

**Chapter**

1. Unity and Cohesion  
8

2. The Reader and Cohesion  
39

3. Cohesion in the Single Term Paragraph  
72

4. Multiple Chain Paragraphs  
107

5. Conclusions: Implications and Applications  
146

**BIBLIOGRAPHY**  
159
INTRODUCTION

All users of language have an intuitive understanding about what constitutes cohesive discourse. Whether listening to a lecture or reading a text, a person looks for meaningful relationships between sentences and will accept or reject a sentence sequence depending on the visibility and adequacy of those relationships. The purpose of my dissertation is to formalize some of this innate knowledge about discourse by describing some of the textual cues that contribute to cohesion in particular types of English paragraphs. The principal area of investigation is the explicit and implicit relationships between nouns and noun chains in a text. The investigation does not examine the contributions made to cohesion by verbs, tense, and aspect.

Past and present research into cohesion falls into one of two categories, rhetorical or linguistic. Rhetoricians since Aristotle have identified and extolled the effects of cohesion on a text: that it becomes unified, logical, dynamically continuous; but they have not attempted to identify and describe systematically the textual constituents of those qualities. Linguists and psycholinguists, especially in the 1970's and 80's, have been attempting such descriptions. The most conspicuous example is Halliday and Hasan's
Cohesion in English, which posits five categories of cohesion devices: reference, ellipsis, substitution, conjunction, and lexical. On a smaller scale, the chapter on "Sentence Connection," in Quirk et al., A Grammar of Contemporary English, proposes several additional devices. Both these works, however, offer dictionaries of possible rather than necessary cohesive devices. Both deal all but exclusively with pairs of sentences rather than genuine texts, and both classify their cohesive devises according to grammatical form rather than textual function. Neither offers a means for evaluating sentence sequences with shared "ties," in Halliday and Hansan's terminology, but nevertheless without cohesion, as in the following example:

? John eats apples in the summertime at his grandmother's farm. Ralph hates John.

Thus neither work offers a means for specifying the necessary conditions for textual cohesion.

Another group of linguistically oriented approaches is based on Charles Fillmore's case grammar, and this includes the work of Bonnie F. Meyer, E. J. Crothers, C. H. Frederickson, and Walter Kintsch. While I will not try to summarize these scholars' monumental efforts to categorize the possible relationships among sentences, one example will indicate the basic inadequacy of these approaches. Bonnie F. Meyer, in The Organization of Prose and Its Effect on Memory, lists under the category of sentence relationships which she
calls "hypotactical rhetorical predicates" the following sub-categories: "attribute," "equivalent," "specific," "explanation," "evidence," "analogy," "manner," "adversative," "setting time," "setting location," "setting trajectory," "representative identification," "replacement identification," "constituency identification." But there are no clear theoretical criteria for the identification of these categories; they are ultimately ad hoc improvisations, and not reproducible by other researchers. Other case grammar approaches produce similar lists, all self-contained and incommensurate with one another.

Yet even if it were possible to overcome this methodological confusion, the theoretical motivation of the case grammar approaches would still remain suspect. All employ a sentence-based deep structure grammar, and all assume that deep structure sentence relationships are the defining relationships of all discourse. They offer no means by which to synthesize the elements of a text, to describe textual totality, or to account for our intuitive sense of a textual whole that is greater than the sum of its parts.

The textual whole that will form the basis for my analysis is the English paragraph, and I assume that in describing some of the relationships between nouns and noun chains in paragraphs, I am also describing some of the basic discourse structures from which paragraphs are made. My description leads to a new, structural classification of
paragraph types, and it offers additional evidence that the paragraph is an identifiable structure within a shared, public grammar of discourse, rather than a mere agglomerate produced by artificial convention or authorial whimsy. Despite evidence to the contrary, the conception of the paragraph as an arbitrary indentation still dominates the work of rhetoricians and composition teachers, and this conception remains an obstacle both to further research and to effective teaching.

My analysis begins with a justification for using the term "cohesion" as the concept that describes most fully the qualities we associate with a coherent text. "Cohesion" is presented as a superordinate term embracing both "unity" and "coherence," and is defined as the presence of a dominant term, either directly or inferentially, in each sentence of a paragraph. Chapter One concludes with the hypothesis that the textual patterning of a dominant term creates in rudimentary form a structural whole or totality. Whereas Chapter One examines cohesion as inherent in the text, Chapter Two examines cohesion from the reader's perspective and summarizes recent psycholinguistic research that supports the hypothesis developed in Chapter One. In Chapters Three and Four, the definition of cohesion as a dominant term present in each sentence of a paragraph and whose recurrences manifest a structural totality is developed through a variety of paragraphs, specifically paragraphs whose structure is
defined by a single-term chain, a double-term chain, and a mixed chain. The dissertation concludes by suggesting briefly the implications of my analysis both for the theory of discourse and the teaching of composition.
NOTES TO INTRODUCTION


6The **Representation of Meaning in Memory** (New York: Wiley, 1974).

7Meyer, pp. 34-35.
CHAPTER ONE: UNITY AND COHESION

On an ordinary, day-to-day level, people equate cohesion with the simple and sustaining fact that some sentence sequences make sense and others do not. Cohesion elevates a random collection of sentences to the status of a text, and in the process imparts meaning, insight and purpose to those sentences. Without cohesion, the text can hardly be said to exist at all, for cohesion provides the textual means for initiating comprehension or sense.

Since that description attributes to cohesion the survival of civilization, the maintenance of friends, lovers, marriage, and peace, not to mention the successful preparation of a Stouffer's spinach souffle or a simple paragraph, some clarification of the term, as distinct from the traditional terms, unity and coherence, seems appropriate.

An ordinary reader would, I think, label the following passage as incoherent:

A fresh batch of cheese is received at Mjuller's store. The best adornment for a maid is modesty and a transparent dress. At sunset wild ducks flew over the little cradle. They're waiting for you at the meeting of the municipal government, Mr. Lancelot.

The passage does not make sense: in the absence of a single subject, it lacks unity; in the absence of logical order
among statements, or of individual words linking statements together, it lacks coherence; and in its overall failure to become a comprehensible text, it lacks what I am calling cohesion. At this level cohesion represents the first and most basic condition of any text as it differentiates a meaningful sequence of sentences from a mere agglomerate. This usage is compatible with our intuitive understanding of the term but does differ from the rhetoricians' "coherence." In their usage, coherence is one of three equally important qualities—along with unity and emphasis—which together constitute a text. I am using cohesion as a superordinate term embracing all three qualities and in the following section will try to justify that usage. I review the main characteristics of unity and coherence as traditionally understood, and try to show how both in fact are components of cohesion. The overall purpose of this chapter is to establish a framework for analyzing the cohesion patterns in English expository paragraphs.

The division of labor between coherence and unity probably had no specific historical origin. Rhetoricians distinguished the two, of course, but the real heritage of the distinction derives from Alexander Bain's influential book, English Composition and Rhetoric, 1877. In his famous chapter on the paragraph, Bain prescribed a set of six rules for achieving good paragraphs. The first required that "the bearing of each sentence upon what went before be
explicit and unmistakable," or, in traditional terms, exhibit coherence.\(^2\) The fifth rule required a paragraph to possess unity, which "implies a definite purpose, and forbid digressions and irrelevant matter."\(^3\) (Bain's other rules call for a topic sentence, parallel construction, appropriate subordination, and "consecutiveness.")

As generally understood, Bain's rules ascribe singleness of subject matter to unity, and to coherence the orderly arrangement of the unified material; more recent scholars maintain Bain's distinction:

A thing is unified if its parts relate to some over-arching principle informing the whole. It is coherent if its parts—especially adjacent parts—are fixed in their relationships to one another. A handful of marbles has unity; but they have no coherence, as one may easily prove by trying to bounce them as one piece across the floor. On the other hand, a piece of sticky bubble gum and a marble may show coherence, but they won't show unity, at least not if they were brought together without principle. . . .

Unity and coherence are thus independent variables.\(^4\)

If we can temporarily accept unity as the development of a central or single idea, we can look more closely at the attributes of coherence, a considerably more difficult concept to pin down, even descriptively. Some definitions offer a starting point. James McCrimmon writes: "A paragraph is said to have coherence when its sentences are woven together or flow into one another."\(^5\) Echoing McCrimmon, Hodges and Whitten advise: "Give coherence to the paragraph
by so interlinking the sentences that the thought may flow from one sentence to the next. Others interpret coherent, "flowing" sentences as orderly sentences. Charles Ruhl: "I will consider discourse to be coherent if there are logical relation between its parts." H. B. Lathrop: "... orderly sequence is coherence. ..." And finally, Lybbert and Cummings add a new dimension to McCrimmon's quality of "woven together:" "A minimal definition of coherence might be 'fixed interdependence and non-autonomy of parts.'"

What emerges from these descriptions and definitions is the simple idea that linking sentences together generates coherence and that to attain that linkage, authors need to repeat words or establish equivalence chains of meaning:

Link sentences together by means of pronouns referring to antecedents in preceding sentences.  
Link sentences by repeating words or ideas used in preceding sentences.  
By repeating key words and phrases, a writer can keep the dominant subject in the reader's mind and maintain the kind of continuity necessary for a smooth flow of logical thought.

Implicit in this advice to repeat words to create coherence is really the advice to create unity. Despite the premise that the two are separate qualities, at least one aspect of coherence involves unity. If there is no "dominant idea," if there are no "key phrases to repeat," then unity as well as coherence is missing. The distinction between unity and
coherence is not so easily maintained; they do not turn out to be independent variables, as Lybbert and Cummings argued; rather coherence is a function of unity, as seen in the following primitive, albeit cohesive paragraph, where the mere repetition of "fluorite" establishes simultaneously both a "dominant idea" and cohesion:

Fluorite in commercial quantities occurs in both sedimentary and igneous rocks. Veins of fluorite with quartz or calcite sometimes contain lead, copper, and zinc materials. We are the world's largest producers of fluorite, most of it in Illinois, Kentucky, Colorado, and New Mexico. Fluorite is used to produce a fluid slag in steel-making and in smelting ores. It is used in making high-test gasoline, Freon, and many other chemical products. Fluorite is a most attractive mineral of varied colors—white, blue, green, and violet.13

Given the way real texts actually hang together with little more than repeated nouns or themes, the concept of unity becomes the basis for a minimal but at least somewhat operational definition of cohesion, as seen in the following:

... a necessary (though obviously not sufficient) condition of the coherence of a text consists, roughly speaking, in repetitions. The... structure of each utterance $S_1$ of a text $S_1, \ldots, S_n$ is such that at least one lexical item which occurs in it, or one proposition which can be inferred from it, can also be found within the utterances of the sequences $S_1, \ldots, S_{i-1}$, or within the propositions which can be inferred from those.14

We suggest a text base is cohesive if it is connected by argument repetition.15

The first level of cohesion, then, is the recurrence of the same lexical item. The simplest form of this recurrence
is the repetition of the same term in which a strict identity relationship is upheld, as in:

The boys climbed the trees. The trees weren't too tall for them.

Frequently, however, strict identity of the terms is not required, even when the terms are morphologically the same, as in:

I wanted to buy some apples. But they were sold out.

Take care of your wallet. Peter always tries to pick it from his closest friends.

In these cases, the apples to be purchased are not identical to those sold out nor is John's wallet identical to the ones picked. For these instances, one must posit a kind of existential class of apples or wallets in which both terms represent subsets of their respective classes; instead of the relation of identity, the terms exhibit one of class membership.

A third kind of recurrence is synonymy: items which share the same referent and which can be substituted for one another without significantly changing the meaning, as in:

The boy climbed the fence. The youth is a member of a gang.

This general and commonplace notion of synonymy would outrage philosophers; in their view, no two words will be absolutely identical in all contexts and thus synonymy is impossible. Yet since almost all texts consistently exploit synonymy for stylistic variation, any practical discourse
analysis must either rely on this subjective definition or must be abandoned.

Additionally, a text uses items that are equivalent--similar--rather than identical to signal repetitions. In some instances, there is no clear cut distinction between synonymy and equivalence, causing some authors to group them together. I am separating them on the basis that synonymy involves the same referent, but equivalence involves a class-member referent like the apples and wallets examples. Like synonymy, equivalence is an intuitive concept that resists formal definition; having no sufficiently rigorous semantic theory for equivalence, we must again rely on rough and ready and usually subjective notions of similarity or equivalence. Some potential kinds of equivalence would be the following suggested by Enkvist:

contracting hyponymy: People got on and off. At the newsstand Frenchmen, returning to Paris, bought that day's paper.

expanding hyponymy: Tulips are cheap even in January. But then flowers seem to be necessary to Scandinavians during the darkest season.

sustained metaphor: The sun sagged yellow over the grass plots and bruised itself on the clotted cotton fields. The fertile countryside that grew things in other seasons spread flat from the roads and lay prone in ribbed fans of broken discouragement.

co-membership of the same world field: Tulips are cheap. Roses are expensive.\textsuperscript{17}

As this classification suggests, categorizing the semantic relation between words is a very problematic
venture. Besides the examples cited above, one also needs a system that describes the cohesive effect of a sequence like:

The girls played house. The boys spit on them. In all the children were feisty that day.

Here boys and girls, since they are opposites, cannot share the same referent, but as members of the class of children do share the textual antecedent "children." Given the breadth and complexity of these relations, equivalence can best be defined as items exhibiting set relations with one another. This evasive definition echoes those reached by other researchers. Halliday in "The Linguistic Study of Texts" defines equivalence as "two or more occurrences, in close proximity, of the same lexical item, or of items paradigmatically related in the sense that they may belong to the same lexical set." Halliday uses the example "I took leave and turned to the ascent of the peak. The climb was perfectly easy." Similarly, Winburne, in his analysis of the Gettysburg Address, defines equivalence as items belonging to the "same semantic class" and groups into "one semantic class the items here and battlefield." Despite lacking a full and detailed articulation, these three categories of 1) item repetition (with either complete or partial identity), 2) synonymy, and 3) equivalence do enable us to trace Bellert's repetitions (or what I am calling item recurrences) in patterns through a text,
even when the recurrences are not literal ones. A text's ordinary use of repetition/substitution/variation as it develops and refines its topic requires such an approach. Additionally, using these three categories of an item's relation to its referent enables a researcher to calculate the cohesiveness of a recurrence chain by asking whether the repeated items all stand in the same relation to their shared antecedent/referent. Thus the cohesiveness of the following sequence can be attributed to tulips' and daisies' shared relationship to flowers as members of a class:

Tulips are cheap in January. Daisies are plentiful also. But then flowers seem to be necessary to Swedes.

On the other hand, the following sequence lacks cohesiveness because "dogs" and "tulips" do not share an identifiable relationship to "flowers" as their common antecedent/referent:

Tulips are cheap in January. Dogs are plentiful, too. But then flowers seem to be necessary to Swedes.

Secondly, these three categories provide a means for determining the cohesiveness of recurrences that are implied rather than stated. If no antecedent/referent is stated, but if one can be inferred to which all recurring items stand in the same relation, then the sequence can be perceived as cohesive:

The gearbox was smoking. The windshield wipers were stuck.
Here "gearbox" and "wipers" both stand in part-whole relation to the implied referent, "car."

Finally, the categories provide a means for distinguishing between strongly and weakly cohesive passages. The more narrowly the implied antecedent/referent is specified, the more strongly cohesive the sequence will be. Thus the following passage would receive low marks for cohesion:

Tulips are cheap in January. Shrubs are plentiful, too. But then flowers seem necessary to Swedes.

The implied referent has to be some highly general term like "plants," which designates too broad a class to establish clear equivalence among tulips, flowers, and shrubs.

In summary, then, recurrences establish patterns of unity where two or more recurring items—what I am calling a recurrence chain—enable a text to develop its basic topic by a series of substitutions in which there may be an interplay between lexical repetition, synonymy, and equivalence.

Besides these three semantic means for creating recurrences, texts also employ grammatical means: substitution, reference, and ellipsis. All basically abbreviations for the just discussed lexical recurrences, these methods are quite straightforward and considerably less problematic than the lexical recurrences. Halliday and Hanson, in their book *Cohesion in English*, exhaustively describe and analyze these methods; in brief summary, they are as follows.
Substitution is "the replacement of one item by another,"\(^{20}\) as in "My axe is too blunt. I must get a sharper one."\(^{21}\) Reference involves items that cannot be understood in their own right and must be interpreted by their antecedents:\(^{22}\) personal references indicated by he, she, him, and the like; demonstrative references indicated by this, these, those, and the like; and comparative references indicated by better, best, and the like. Ellipsis is "the omission of an item,"\(^{23}\) as in "Would you like to hear another verse? I know twelve more."\(^{24}\) Since these are grammatical methods, and to that extent placeholders for lexical items, these methods are really ancillary to semantic recurrences in a cohesion study.

As a means for tracking the topic of a text, recurrence chains--whether primarily lexical or grammatical--represent a first step in describing cohesion. But recurrence chains are merely word lists, and to attain the status of discourse they must be incorporated into sentences. At this juncture a second set of relationships must be introduced: the syntactic relations of sentences and sentence constituents to one another. These relations lie in the traditional domain of coherence--the relationships between parts, with the individual sentence assigned the role of "part." Halliday and Hansan typify the traditional view: ". . . we can interpret cohesion. . . as the set of semantic resources for linking a SENTENCE with what has gone before"\(^{25}\) and
there are many examples where cohesion is achieved by linking entire sentences to each other:

The door slammed shut. The window blew open. He showed no pleasure at hearing the news. Instead, he looked even gloomier.

But there are also instances where we sense that only parts of sentences are being linked with only parts of other sentences:

In joining at organization, the individual accepted a set of fixed obligations for a specified set of rewards. These obligations and rewards remained the same over a relatively long span of time.

With a little tinkering, we can accentuate this sense of parts linked to parts:

In joining an organization, the individual accepted a set of obligations in return for a set of rewards. The obligations were fixed. The rewards were specified. These obligations and rewards remained the same over a relatively long span of time.

In the same way, we can often condense into one sentence the previously linked parts of two sentences:

The hustlers have co-opted every salable affliction. This bothers me.

The hustlers' co-opting of every salable affliction bothers me.

Because of a kind of semantic drag created by item repetition, whole sentences are often linked together as traditionally claimed; but these samples also remind us, first, that the sentence, once minimal requirements of subject and predicate are met, is an arbitrary and elastic unit, with no syntactic limits on its length or complexity,
and second, that the syntactic make-up of a sentence affects our sense of coherence. We are dealing, then, with two perspectives on sentences: the sentence as a unit, and the sentence as a structured array of internal elements. English is a language of synthesis and a language of position; meaning, emphasis, and theme all depend upon position and only the dual perspective of a sentence as a unit and a structured array of elements captures these important discourse qualities. The question for a cohesion study is what method of describing sentences best maintains this double perspective.

The most immediate choice lies between surface and deep description. The usual distinction between surface and deep constituents is that the surface structure represents the actual textual shape and content of a sentence, and the deep structure its formal and fully realized originating kernels. Thus the sentence "John bought a ball and the ball was red" and the sentence "John bought a red ball" have the same deep structure but different surface structures. The deep structure is revealed by analysis of the sentence into its primary arguments, its actual or potential recurrence chain items. Then a series of transformations affecting these kernels links and transforms the kernels' deep structures into their surface form.

Many transformations, however, are optional; whereas a given surface structure will always yield the same deep
structure, a given deep structure will not always yield the same surface structure. Because of this incommensurability between deep and surface structures, and because the surface structures represent the author's actual decision about what transformations to employ, and because these decisions reflect what is actually said, what is omitted, what is emphasized, the description of surface forms seems most appropriate to cohesion study. Put another way, the surface forms enable us to talk about the sentence as a unit, and also provides information about the syntactic make-up of the sentence, specifically its length and kind.

Surface forms not only allow us to maintain a double perspective on the sentence as a unit and as a structured array of elements; they also enable us to introduce into discourse analysis the traditional set of grammatical concepts: subject, predicate, and modifier. These concepts, especially subject and predicate, offer a fairly straightforward means for tracking thematic movement—or recurrence chain items—through a text. They help us identify the overall pattern established by a recurring item as it moves from subject position to object position, back to subject, and so forth. Even more important, the interplay between these traditional grammatical concepts and a recurring item help explain the differences between more cohesive and less cohesive sequences (by enabling us to determine when syntax is congruent with the recurring item), as in the following samples:
Stronger - cohesion: John likes peanuts. He goes to ballgames just to buy them. He also rides a ten-speed bike and takes his sister to the movies.

Weaker - cohesion: John likes peanuts. Mary doesn't like John. APC gave him a research grant. John's parents think he should see a dentist.

The major drawback to surface structure analysis lies in its assignment of subordinate and superordinate status to constituents. Calling "John bought a red ball" more coherent than "John bought a ball and the ball was red" presumes that the surface structure accurately reveals the hierarchy among constituents. Generally it does, as in this example, but not always: certain prepositional phrases are persistent exceptions. In the surface structure prepositions are always subordinate; in the deep structure they are superordinate and dominate other constituents. Moreover, the head-words in these prepositional phrases often become dominant recurring items in a text, as in the implied recurrence of "park" in the following sequence:

John walked in the park. The trees were just turning color. People sat on benches and fed squirrels.

In summary, then, surface forms are obvious choices for describing the sentence as a unit and for providing classificatory information about the sentence; in addition, surface forms introduce grammatical functions that can be and are exploited for purposes of coherence. On the other hand, deep structure analysis provides a more consistent hierarchy of
constituents, and a hierarchy that is often reflected in the overall discourse structure but not in the surface structure. Given these parameters, my analysis will proceed on the assumption that surface forms are valid units for analysis, and also that deep structures at certain points must be dealt with; I will use the surface forms as far as they carry us, and shift to deep structures when the analysis seems to require it. Since these methods represent two demonstrable stages of sentence analysis, it should not be surprising that one or the other should be more revealing or relevant in particular situations.

Finally, it must be noted that in cohesion, the interaction between recurring lexical items and syntax is not an interaction between equals. Where a recurrence chain exists, there is cohesion; without a chain, no cohesion. It is a rocklike condition of cohesion that any sentence sequence containing a recurrence chain, no matter how simple, raw, crude, or convoluted the sentences, will still be minimally or at least suggestively cohesive. Syntax can maim, twist, or obscure the relations among recurring items, but whether it can completely thwart their collective power is debatable. Does a reader "understand" "colorless green ideas" because he uses the phrase's semantics or because he interprets the phrase through the paradigm of syntax? In the analysis of cohesion, the importance or non-importance of syntax is a very wide open question.
Assuming that coherence, as distinct from unity, involves both sentences' deep and surface syntactic relationships, we might now seek to establish the range of these relationships and to identify their particular contribution to overall cohesion. To this end, we will compare two texts: an "original" that is unified but weakly coherent, and a revised version with greater coherence:

The opossum has survived in definitely hostile surroundings for seventy million years. The opossum is small; it can easily find a little food, while big animals starve. The individual opossum is not very delicate; it can stand severe punishment, It "plays 'possum" when it gets into trouble. It can go without food for a long time. Many different things are food to an opossum. Traits of the opossum have a high survival value. The opossum is a survivor from the Age of Reptiles.

The reasons our opossum has survived in definitely hostile surroundings for 70 million years are evident. One is his small size: small animals always find hiding places, they always find a little food, where the big ones starve. Another of its assets was its astounding fecundity: if local catastrophes left only a few survivors, it did not take long to re-establish a thriving population. Also the individual opossum is not exactly delicate: it can stand severe punishment--during which it "play 'possum" and then scampers away--and it can go without food for a considerable time. Finally, a great many things are "food" to an opossum. Each of these traits has a high survival value, and their combination has presented the United States with a survivor from the Age of Reptiles.

Perhaps the most striking difference lies in the sentence constructions. In the original, the sentences are primarily kernel sentences--short, choppy, grammatically simple--containing single propositions. In the revision,
the sentences vary in grammatical complexity and contain several propositions per sentence; they "flow." Thus the sentence's surface structure—the form in which propositions, particularly recurrence chain items—contributes to coherence in a significant way. For example, the revision combines two kernel sentences from the first text—"The opossum has survived for 70 million years" and "The reasons are evident"—into a single, complex sentence that signals the relationship between the two kernel sentences. Quite simply, the syntactic capacity of sentences to show relationships—complementarity, as in this example, or subordination, or coordination—can be exploited to increase coherence.

A second difference lies in the number of recurrence chains in the two versions. In the original only one chain—the opossum—is overtly signaled. In the revision two chains are signaled—the opossum, and the reasons for its survival. This number, two, is a critical if somewhat obvious criterion, since it is relationships that make for coherence, and relationships require at least two elements. As we shall see later, while the requirement of "two" can be fulfilled by two recurrence chains, it can also be fulfilled by a recurrence chain plus another "thing."

Recurrences make for coherence not only by their frequency but also by their physical location, and ultimately by their function. Consider the following examples:
I was accepted and started work. My experience had been chiefly derived from books. I was not prepared for the difficult period of adjustment. I soon became discouraged with myself and so dissatisfied with my job that I was on the point of quitting. My employer must have sensed this. He called me into his office and talked to me about the duties of my position and the opportunities for advancement. I realized that there was nothing wrong with me or the job and I decided to stay.

I was accepted and started work. Until that time my experience had been chiefly derived from books, and unfortunately those books had not prepared me for the difficult period of adjustment that every inexperienced secretary must face in a new position. Consequently, I soon became so discouraged with myself and so dissatisfied with the job that I was on the point of quitting. I think my employer must have sensed this, for he called me into his office and talked to me about both the duties of my position and the opportunities it offered for advancement. That talk helped me considerably. From then on, I realized that there was nothing wrong with me or the job that experience could not cure, and I decided to stay.

The revision differs from the original in its greater number of localized recurrences, that is, recurrences that are positioned in adjacent sentences, make only that single or double appearance, and do not form chains.

The grammatical manifestations of a recurrence chain--reference, ellipsis, and substitution--are all forms of partial repetition. In the original text, the primary recurrence was the repeated word "opossum;" the revision used reference, ellipsis, and substitution as well. Functionally, these devices do two things: they maintain an
unbroken chain of recurrences and thereby establish some degree of cohesion through unity; they also function—at least psychologically—to subordinate information already known or recoverable by reducing the autonomy of sentences containing these devices and forcing the reader back to preceding sentences for the antecedents or other substitutions.

These grammatical means for signaling recurrence chains do not convey new semantic information; the lexical means, especially equivalence, do and it is through their addition of semantic information that further relationships between parts are established:

**Grammatical:** John and Bill pick peaches. They live next door.
**Lexical:** John and Bill pick peaches. These troublemakers live next door.

As indicated earlier, the potential relationships between items can be seen most usefully as set relations. In this instance, John and Bill are members of the set of troublemakers; when further recurrences are added to the chain in a cohesive text, the chain itself will instantiate a pattern of relationships such as part/whole or member/class.

Thus far we have looked at the ways in which relationships between parts manifest themselves in sentence structure, the number of recurrences, the kind of recurrences, and finally the location of recurrences. We must now look at the effect of transition words in producing coherence. This group of cohesive devices has probably attracted the greatest attention because clearly they do link whole sentences, and
elaborate classifications have been made of the sentence relationships signaled by transition words. Milic, for instance, has proposed a system of eight relations between sentences: additive, initial, adversative, alternative, explanatory, illustrative, and causal.\textsuperscript{30}

Winterowd, in his article, "The Grammar of Coherence," posits seven and only seven relationships. These are:

1) coordinate; 2) obversative; 3) causative; 4) conclusive; 5) alternative; 6) inclusive; and 7) sequential.\textsuperscript{31} No doubt these seven exist, as Winterowd's examples and our own common sense indicate; yet there are several problems, and Winterowd's analysis carries us only so far. For one thing, his list of relationships is not exhaustive after all, as the following sequences indicate:

1) Fluorite in commercial quantities occurs in both sedimentary and igneous rocks. Veins of fluorite, with quartz or calcite sometimes contain lead, copper, and zinc materials.

2) It is necessary to define intellectuals. They are all those who create, distribute, and apply culture.

3) On the station platform were Negro soldiers. They wore brown uniforms and were tall, and their faces shone.

What seems basically and sometimes inexplicably to hold these sentences together is simply repeated reference, either through actual repetition or through pronouns. Of Winterowd's seven relationships, only coordination seems a possibility, but when one inserts his coordination words, the results are quite disarming:
It is necessary to define intellectuals. (Furthermore) they are all those who create, (And) distribute, and apply culture.

The second and more fundamental difficulty with Winterowd's system is that where the sentence relationships are implicit—which he recognizes as possible—they can only be implied by the lexical content of the sentences, i.e. the recurrences I have already discussed. Transitional words do not in themselves create relationships; they simply mark the relationships already existing among the lexical items. As the following example indicates, if there is not already some lexical compatibility between the words of different sentences, the mere presence of transitional words cannot establish relationships among those sentences:

The albacore are beginning to run. However in Montana it often snows in August. That is why Freud created a revolution in the way we view the human mind. Therefore, grammar seems to be a dull subject.  

While the logical relationships among sentences described by Winterowd and others certainly do exist, they must be seen as secondary to a more basic set of relations, namely "equivalence" or the sharing of a single antecedent.

A final distinction between texts that are minimally unified and texts that are truly cohesive involves the means by which the pattern of relationships is instantiated, namely the order of appearance of the recurring items. Contrast the following two texts:
1) Troublemakers lower the attractiveness of a neighborhood and bring police cars. Bill and Peter pick peaches. These troublemakers live next door.

2) Bill and Peter pick peaches. These troublemakers live next door. Troublemakers lower the attractiveness of a neighborhood and bring police cars.

The second version is more cohesive than the first because of the way in which its recurrences are ordered. In some texts such continuity involves entire propositions rather than just the recurrence chain items themselves, as in:

1) After the death of Saul, David ruled Israel for forty years. Once he incurred the king's anger and was driven ignominiously from court. As a shepherd lad he had lived in the hills of Judea. He had vanquished the mighty Philistine with his slingshot. The sad-faced Saul was charmed with his songs. He was the sweetest singer in all Israel.

2) David, the shepherd lad who lived in the hills of Judea, was the sweetest singer in all Israel. It was he who charmed the sad-faced Saul with his songs. It was he, too, who vanquished the mighty Philistine with his slingshot. Later he incurred the anger of Saul and was driven from court. But upon Saul's death David came back and ruled Israel for forty years.

In both sets of examples, ordering is contingent on the particular pattern or patterns of relationships being instantiated by the overall text.

In summary, then, cohesion consists of unity, as manifested by a recurrence chain, and coherence, as manifested by the semantic and syntactic relations among links in that chain. Unity and coherence are manifested in a variety of ways: sentence construction, number of recurrences, location
of recurrences; transition words; kinds of recurrences; and order of recurrences. Of these six, only two--sentence construction and transition words--belong to the traditional category of coherence; all the others involve recurrences and belong to unity. The implication seems to be that while unity is absolutely necessary to cohesion, the status of coherence is less clear; it may constitute essentially a refinement of unity.

This view of coherence as a refinement of unity was strongly articulated by H. B. Lathrop in 1918, when he coined the phrase, "unity of connection," a phrase that eminently suits our perception of cohesion. Lathrop writes:

In a word, the unity of theme, directive unity, the unity of a line, is not more truly unity than unity of connection, the unity of a chain or the unity of a woven fabric. . . . Coherence, orderliness, is simply arrangement in a systematic way, as by deduction or induction, from cause to effect, from top to bottom or from bottom to top. There is no one plan that is preeminently coherent; a coherent result is produced by any systematic principle of guidance appropriate to the case. 34

Modern researchers echo Lathrop with similar definitions of cohesion, most of them less operational than those of Bellert and Kintsch, but definitions which attempt nevertheless to capture the interaction between unity and coherence. Geoffrey Leech writes: "Cohesion. . . is the way in which independent choices in different points of a text correspond with or presuppose one another, forming a network of sequential relations." 35
Ross Winterowd: "If one perceives form in discourse, he also perceives coherence, for form in discourse is the internal set of consistent relationships in any stretch of discourse, whether poem, play, essay, oration, or whatever." In using such phrases as "network of sequential relations" and "form in discourse," these writers too are trying to identify a larger, more formal kind of unity than simple unity of content, a unity that also satisfies our sense of totality or structure. Or as Jonathan Culler writes: "Some Christmas trees are more successful than others, and we are inclined to think that symmetry and harmonious arrangement of ornaments makes some contribution towards success."37

In an effort to clarify the notion of "unity of connection," the remainder of the dissertation will explore and develop the following hypothesis: cohesion consists of unity, which in turn consists of both a) a dominating item recurrence present in or inferrable from all sentences in a paragraph, and b) a pattern or totality manifested by that item recurrence. Coherence, the orderly arrangement of recurrences, is an aspect of unity and functions to make more explicit the already present and implied relations among recurrences.

The basic difference between my definition and Bellert's lies in my additional criterion that the recurring item also "dominate" in the sentences of the paragraph. Bellert herself
notes that her definition of cohesion, as basically a term shared by all sentences in a sequence, is inadequate, as seen in a passage like the following:

John likes oranges. Oranges grow in California. Roscoe always throws oranges at the referees of basketball games. Oranges and apples are fruits. My grandmother makes orange preserves and dried orange dolls.

Although the passage contains a recurring item, oranges, it is still not cohesive, because conceptually it is not about anything; it lacks the unity of connection that can only be provided by a dominating recurrence. The concept of dominance is crucial here, and operationally defining it is the main theme of chapters three and four.

A second major aspect of my hypothesis is that dominating item recurrences create or manifest a totality. I adapt the term "totality" from Jonathan Culler, who uses it to designate a form or pattern that creates a sense of wholeness: "Their (texts') unity is produced not so much by intrinsic features of their parts as by the intent at totality of the interpretive process; the strength of the expectations which lead readers to look for certain forms of organization in a text and to find them." Culler briefly suggests six possible "forms of organization": binary opposition, the dialectical resolution of a binary opposition, the displacement of an unresolved opposition by a third term, the four term homology, the series united by a common denominator, and the series with a transcendent or summarizing final term.
Culler's observations parallel Kenneth Burke's when Burke says, "Form in literature is an arousing and fulfillment of desires. A work has form in so far as one part of it leads a reader to anticipate another part, to be gratified by the sequence." Burke suggests five kinds of form--progressive (subdivided into syllogistic and qualitative progression), repetitive, conventional, and minor or incidental form. The latter two need not concern us since they cover forms like the sonnet, metaphor, and apostrophe that are not directly associated with the organization of texts. Burke describes syllogistic progression as follows: "To go from A to E through B, C, and D is to obtain such form. We call it syllogistic because, given certain things, certain things must follow, the premises force the conclusion."

Whereas in syllogistic progression it is a particular event, in qualitative progression it is a particular state of mind that leads through stages to another state of mind. Repetitive form, the one that most concerns us here, Burke defines as "the consistent maintaining of a principle under new guises" or the restatement of a theme by new details. The extent to which these broad conceptions of form apply to paragraphs and paragraphing is a major focus of chapters three and four.

These chapters will also test the hypothesis that totalities exhibiting cohesion also exhibit particular structures that can be identified and described, so that a
theory of cohesion will also yield a description of paragraph structure. If an analysis of cohesion reveals the relationships among discourse parts, then it may also be expected to reveal the structure of paragraphs.

Finally, my hypothesis assumes that these large notions of totality and form are to some degree a priori, and that readers bring to the text various prior experiences that enable them to form a concept of wholeness which they expect the text to satisfy. The enormous capabilities of the reader, and his contribution to cohesion, will be the subject of chapter two.
NOTES TO CHAPTER ONE


3 Bain, p. 116.


8 H. B. Lathrop, "Unity, Coherence, and Emphasis," University of Wisconsin Studies in Language and Literature, 2 (1918), 32.

9 Lybbert, 35.

10 Hodges, p. 332.

11 Hodges, p. 333.


34 Lathrop, 82.

35 Geoffrey Leech, "'This Bread I Break' -- Language and Interpretation," A Review of English Literature, 6 (1965), 66.


38 Culler, p. 91.

38a Culler, p. 174.


40 Burke, p. 184.

41 Burke, p. 184.
CHAPTER TWO: THE READER AND COHESION

The first chapter presented cohesion from a basically text-oriented perspective and relegated the reader to a minor role. But that is in many ways wishful thinking, since cohesion is not a static, immanent quality of the text but rather a judgement made ultimately by the reader. The cohesiveness of a text is a matter of degree, if not taste, and what is acceptably cohesive for one reader may very well not satisfy another. While the text is distinctly more accessible and manageable than the reader, some note must be taken of the reader's role and capacities. It is thus my aim in this chapter to outline some of the facts and theories about reader cognition and information processing as these affect a text's cohesion, and to suggest some means for incorporating the reader's role into an analysis of cohesion. The second purpose of this chapter is to suggest that the analysis of cohesion outlined in chapter one seems to coincide with current theories of listener-reader information processing.

"The reader" is probably the most fashionable research topic around these days, one which offers work to philosophers, psychologists, linguists, artificial intelligence programmers, and literary critics; we have Freudian,
pluralistic, ideal, informed, and Norton readers, each with different and sometimes contradictory traits, abilities, and intelligences. Even so, all reader response theories share one very basic and informing assumption—that the reader is an active rather than passive processor, sharing with the author the pragmatic assumption that a text is really a text and that it will make sense to the reader who expects cohesion and takes for granted at the outset that there are connections between sentences. As a result, finding cohesion becomes a telelogical process in which cohesion is first assumed and then, so to speak, attained by the reader who explores the text for all possible ways in which it might be manifested.\(^1\) The desire for cohesion is so powerful that even apparently contradictory sentences can be smoothly and logically connected by readers. Given the sequence,

\[\text{The party was a success. Bill and Mary fought.}\]

most readers, jarred at first, quickly construct acceptable connections:

\[\text{The party was a success, even though Bill and Mary fought.}\]

\[\text{The party was a success. Bill and Mary fought; they were funny.}\]

In a deliberate way, the preceding example illustrates a totally commonplace phenomenon, as much taken for granted as regular breathing. No text supplies all the information necessary for its understanding; authors consistently leave out information that they feel can be inferred by the reader,
so that comprehension and cohesion depend not only on linguistic cues underlying the message but also on general knowledge—"The layman's belief system concerning human behavior in Western culture"—as if the text itself were only the tip of the iceberg. Freedle and Carroll illustrate this with the following example:

Suppose Frank and Joe are walking down the street and are about to pass a hamburger stand. Frank says, "Would you like to eat here?" Such a question contains assumptions of at least time and place: The time is not "right now" but "pretty soon, after we get served," and the place is not "right here on the street" but "at the counter of the hamburger stand." Yet if Frank and Joe are walking in the park and carrying a picnic basket the question "Would you like to eat here?" would assume that the time is "right now, as soon as we can unload the basket" and the place could be "right here." In either case, the total situation supplies the context which evokes the semantic assumptions that are most consistent with the utterance.

Freedle and Carroll's analysis implies a modest version of what I shall call the iceberg theory of language, in which words and sentences become tokens, lumps of coal or pieces of wampum, fragments that can be interpreted in only much larger contexts. Besides the reader's belief system about Western culture, that context also includes: the actual visible scene; the linguistic medium; the relationships between parties, and the specific purpose of communication. These are all aspects of the larger contexts of language, or what Karl Bühler designated as "fields." Blumenthal describes Bühler's theory:
Buhler's field concept was the most important. Given two speakers of the same language, no matter how well one of them structures a sentence his utterance will fail if both parties do not share the same field to some degree. . . . There are inner aspects of the field, such as an area of knowledge, or outer aspects, such as objects in the environment. Indeed, the field can be analyzed into many aspects. The total field (Umfeld) consists not only of the practical situation (Ziegelfeld) in which the utterance occurs, but also the symbol field (Symbolfeld) which is the context of language segments preceding the segment under consideration. . . . The structure of any particular language is largely field-independent, being determined by its own particular conventional rules, but the field determines how the rules are applied . . . with a "rich" external field less needs to be specified in the sentence.5

Buhler's field theory outlines the several perspectives guiding discourse analysis today, and prefigures the steady movement away from the conception of language as an autonomous formal entity that can be studied in vacuo. Whether as intellectual backlash against Chomskian rationalism or as a merely predictable pendulum swing, discourse analysis at all levels now seeks to incorporate and account for the discourse pragmatics affecting language use and comprehension. Bonnie Lynn Webber, in "A Formal Approach to Discourse Anaphora," temperamentally concludes that: "None of the three types of anaphoric expressions that I have studied--definite anaphor, 'one'-anaphora and verb phrase anaphora--can be understood in purely linguistic terms. That is, none of them can be explained without stepping out of the language into the conceptual model each participant is synthesizing
from the discourse." Others, less temperate, escalate the role of pragmatics. After a study of sentence memory, Bransford, Barclay, and Franks write: "In a broader sense the constructive approach argues against the tacit assumption that sentences 'carry meaning.' People carry meanings, and linguistic inputs merely act as cues which people use to recreate and modify previous knowledge of the world."

Language is a tool, a medium, a complicated interface between people and their world views; as an interface, a text will always include potentially more information that is explicit in its linguistic components. Yet while a text cannot be understood by attending only to its sentences, neither can it be understood without attending to those sentences. If a sentence does not carry meaning, what does it carry?

Given the capacious abilities of readers to make connections between sentences and their world knowledge, and given our equally capacious ignorance about how these connections are facilitated, my discussion of reader inferences will aim at only some limited and simple generalizations for certain types of situations; more specifically, it will speak exclusively to how reader inferences often supply the recurrence chains that chapter one described, and it will suggest a means for incorporating reader inference into an analysis of cohesion.
In my study of inferencing and cohesion, I have assumed Buhler's field theory—that successful communication depends on a certain amount of field overlap between author and reader—to be essentially accurate. On that assumption, the first stage of the analysis is to delimit more usefully these overlapping fields. We can begin with the following example:

John walked into the room. The chandeliers were beautiful.

On reaching the second sentence, the reader must decide which chandeliers. Since there has been no mention of chandeliers, there is no antecedent for "the chandeliers." To fill this gap, the reader builds "bridging assumptions," a term proposed by Haviland and Clark. Whenever a reader cannot directly identify an antecedent, the reader "co-operates with the author and supplies the antecedent. In this example, when the reader makes the bridging assumption that the room has chandeliers, then an inference chain or quasi syllogism connecting the two sentences might go something like this:

John came into the room. The room contained some chandeliers. John thought the chandeliers were beautiful.

The fields necessary for the reader to construct the inference chain seem to be of three kinds, all mutually supporting each other. The reader must first know the lexical meaning of "chandelier." Second, the reader must have world knowledge that rooms may contain chandeliers. And finally, the reader
must have discourse knowledge about the significance of the definite article. The definite article before a noun can signal at least two things: most simply, that the noun presupposes some antecedent, and correlative, that a noun accompanied by a definite article can signal a relationship between this noun and another, preceding noun, as in:

The room was filled with people. The women were. . . .

John bought a new house. The front door was of massive steel.10

Van Dijk describes this kind of inferencing as follows: "... any individual having a specifiable relationship with an already introduced individual is PARTICULARIZED by this relation and hence definite in surface structure."11

Although no single relationship is signaled by the definite article, the fact that some relationship is signaled is relevant to an analysis of cohesion. When reaching the cue for the definite article, apparently the reader attempts to construct a quasi syllogism, an Aristotelian enthymeme, both from available textual information and his stored world knowledge. Even though applying discourse knowledge will not guarantee the correct inference, it does apparently direct and instruct a reader's use of world and lexical knowledge.

Yekovich and Walker, in "Identifying and Using Referents in Sentence Comprehension," report on their experiments in which subjects were tested to determine how much the use of
definite articles instead of indefinite articles increased comprehension. Their results indicate clearly that definite articles facilitated the identification and use of referents. Their work basically confirms observations made by Haviland and Clark, who also argued that the definiteness of a referent affected the extent to which an antecedent was established in memory and used to integrate new, incoming material. 

In some instances, discourse knowledge, besides directing and instructing a reader how to use both world and lexical knowledge, can also create new world knowledge; in this respect, only two of the three fields need overlap, and not all three, as Bellert in "Solutions of the Problem of Presuppositions" has argued. Given the following sequence,

Picasso left. The painter went to the coast.

the reader can infer that "the painter" refers to "Picasso," even without knowing that Picasso is a painter, because of both the cuing presence of the definite article and the lexical knowledge that "Picasso" and "painter" are compatible terms that share the feature of humanness. Indeed, authors frequently exploit this condition for purposes of transmitting new information. Consider the almost endless list of words that could reasonably follow "Picasso left:"

The octogenarian . . .
The prize winner . . .
The screwball . . .
There can be little doubt that readers make and use such inferences constantly during comprehension, and naturally comprehension studies confirm this; but more interestingly, they also show that inferred propositions attain the same status in memory as explicit ones. Kintsch tested subjects who were given the following sequence:

A burning cigarette was carelessly discarded. The fire destroyed many acres of virgin forest.

After a fifteen minute delay, the subjects could recall the implicit information (that the cigarette started the fire) as fast as the stated information. It is interesting and useful to note, however, that the delay in testing a subject's recall was necessary for subjects to remember the implicit information as quickly as the explicit; in recall tests done immediately after subjects read a passage, they remembered the explicit information better and faster.  

There is obviously some virtue in being direct and explicit.

If one grants the status and function of inferred propositions, then a number of simple observations can be made about the relation of inferences to cohesion. First, if the reader is required to construct only a single premise for the completion of the enthymeme, then the sequence will be more cohesive than if the reader must supply several premises. Thus the chandelier example is more cohesive than the following:

Barbara is seventeen, and Wendy is old enough to have a driver's license too.
For this sequence to be cohesive the reader would have to construct something like the following:

Barbara is seventeen. Seventeen is old enough to have a driver's license. Barbara is old enough to have a driver's license. Wendy is seventeen. Wendy is old enough to have a driver's license too.

Second, as the example suggests, if the missing premises can be constructed from terms within the text, as in the chandelier example, rather than from information outside the text, as in the driver's license example, then the sequence will be considered more cohesive.

Finally, the closer the textual antecedent is to the particularized noun, the stronger the cohesion will be. For example:

John walked into the room. He sat down and read a magazine. Then he listened to a record. The chandelier was beautiful.

Even if the reader has accurate lexical knowledge and world knowledge about "chandelier," and does not misinterpret the sequence by assuming "chandelier" to have the same status as, say, "piano concerto" or "song," she would still not find the passage as cohesive as:

John walked into the room. The chandelier was beautiful. He read a magazine. Then he listened to a record.

Apparently because the completion of the inference chain requires the reader to leap backward over an intervening and interfering antecedent, the first sequence is less cohesive.

Several studies confirm the importance of ordering. In "Maintenance and Control in the Acquisition of Knowledge
from Written Materials," by Lawrence Frase, subjects given scrambled texts were more prone to recall error than subjects given ordered texts. As good-order texts, Frase used A's are B's and B's are C's-- as in the chandelier example cited just above--as bad-order texts, he used B's are C's and A's are B's. The good-order sequence produced almost twice as many correct inferences as the bad-order.18

In another study by Frase, information about chess pieces was ordered in three ways; around names, i.e.

The bishop is worth . . .

The bishop moves . . .

around attributes, i.e.

The pawn moves . . .

The bishop moves . . .

and finally in a mixed order where neither attribute nor name was exhausted before the other was introduced. In recall tests, subjects performed best with the name order, then the attribute order, and finally the mixed order.19

Several hypotheses may be drawn from these findings. First, since in no case did order completely block comprehension, order must be relegated to secondary importance. This is consistent with chapter one's hypothesis that unity is more important than coherence--or order--in establishing cohesion. Second, because two kinds of order were explicitly compared in the second Frase experiment, it appears that different orders can be equally cohesive and
that there is no Platonic order. In this respect, another experimenter, Crothers, concluded that: "It does not appear that sentence order matters--with the important qualification that only more acceptable orders are being compared." Order is simply a manifestation of the particular whole or totality.

Briefly, then, we have a crude but objective means by which to measure the function of some types of inference in promoting cohesion. If a reader is required at all to construct an enthymeme from the text in order to maintain a chain or referents, then a text which requires the addition of only one premise, constructed from information within the text, and placed sequentially adjacent to the antecedent, will be more cohesive than texts not meeting those requirements.

Obviously these parameters only scratch the surface of inferencing and speak only to those inferences involving Is-A and Has-A relationships. Readers easily construct other kinds of rich inference patterns, as illustrated by Charniak's famous discussion of the following:

Jane was invited to Jack's birthday party. She wondered if he would like a kite. She went to her room and shook her piggy bank. It made no sound.

As Charniak points out, even young readers understand that Jane wanted to buy Jack a kite for a birthday present but that she has no money to buy one--even though the words "present", "money" or any of their synonyms are not mentioned.
These kinds of synthetic inference patterns are common in discourse, and their number and kind do not necessarily affect cohesion or comprehension, although they certainly, by their number and kind, limit any attempt to formalize their relations. Commenting on colleagues' attempts to formalize multiple inference patterns, I. M. Schlesinger poses the simple "John hit Mary," and asks:

If John is a baby, it certainly does not follow that Mary is hurt, and if John is three, five, or seven years old the probability of her getting hurt increases correspondingly. Now, differences of probability are taken into account by Schank and Reiger, who make provisions for different "strengths" of inference. But the really serious problem, which they do not discuss, is the amount of information that has to be stored. They do mention that the "measure of the amount of injury done is a function of the hardness, heaviness, sharpness, etc. of the propelled object (with which one hits), and of the particular body part hit." But by granting this they conjure up a host of data, and one fails to see how they are going to cope with it.22

Besides its "simple" sentence-by-sentence application, inferencing also applies to whole passages. It is this role of inferences that we now need to look at: the way they create "frames," "scripts," "paradigms," "semantic representations"—terms for the listener/reader's equivalent to textual totality. Sentences and their implications are not comprehended in isolation, but as parts of wholes which are formed in part by inferences drawn by the reader. The reader supplies inferences to determine the meaning of each
individual sentence, and then to integrate that sentence with information she already knows. This integration produces a unifying context or situation into which further pieces of information may then be fit and understood. Thus a "dozen eggs, a loaf of bread, and tomato soup" make sense and exhibit cohesion when the reader constructs an appropriate frame for the information—a grocery list. As Rumelhart says, "The process of understanding a passage consists in finding a schema which will account for it."

To illustrate the power, if not the necessity, of the frame, we can look at an experiment done by Bransford and Johnson in which people were read the following passage entitled "Watching a Peace March from the Fortieth Floor:"

The view was breathtaking. From the window one could see the crowd below. Everything looked extremely small from such a distance, but the colorful costumes could still be seen. Everyone seemed to be moving in one direction in an orderly fashion and there seemed to be little children as well as adults. The landing was gentle, and lucky the atmosphere was such that no special suits had to be worn. At first there was a great deal of activity. Later, when the speeches started, the crowd quieted down. The man with the television camera took many shots of the setting and the crowd. Everyone seemed glad when the music started. Given the context created by the title, most people were able to understand the passage until the sentence, "The landing was gentle, and luckily the atmosphere was such that no special suits had to be worn;" at that point their frame
would not accommodate the new information. On the other hand, those given the same passage with the title "A Space Trip to an Inhabited Planet" had no special trouble with the landing sentence.

In this example, the reader constructs a temporary framework of interpretation which makes some meanings possible while excluding others. Each new piece of information reduces the possibilities until a quasi-deterministic structure emerges, or, in other words, until a comprehension context emerges. This concept of the frame is straightforward and noncontroversial; its more impenetrable aspects emerge, however, when we ask the traditional structuralist question, "Where does the frame come from?" One school of research, the text-based school represented by Charniak, Hobbs, and Reiger, places the burden on the text to supply enough cues for a reader to invent a frame and

stresses the notion that the inference process looks far meaningful relationships between different propositions in the text.25

In contrast, the model-based school of Lehnert, Rumelhart and Ortony and Schank and Abelson presumes a priori structures that the reader brings to the text and that the text merely activates. This school

argues that a central purpose of inference is to synthesize an underlying model which organizes and augments the surface structure fragments in the text. In this view, inference is controlled by a target structure that specifies the a priori constraints on the kind of model to be synthesized. This
target structure acts as an organizational principle for guiding a set of inference procedures. 26

In light of language's dependence on world knowledge and in light of our own common sense experiences with language, the model based approach seems intuitively to describe textual understanding best. In On Message Structure, Rommetveit cites the situation of the middle-aged, married couple where the wife, "in response to the visibly exposed gloomy tension of her husband at that moment may serve as a prelude to the--to the outsider--very cryptic remark 'Pot.' And the wife 'understands perfectly:' what is worrying her husband at the particular moment is the possibility that their son Sam may start smoking pot." 27 Only the concept of an a priori frame composed either of a perfectly shared world view, as in this example, or of an imperfectly shared world view, as in texts, can explain language use.

Even so, accepting the existence of frames and their a priori basis advance us only so far; given the "less than perfect synchronization of intentions and thought" between most people, the next question must be, What are the textual constituents of a frame? 28 Unfortunately, applied research in this area is minimal, although test-tube theories abound. Under the influence of recent linguistic advances, theoreticians have adapted transformational and case grammars' distinction between deep and surface structure as a
means of dealing with the fact that people do not remember texts \textit{verbatim}, and thus of resolving one of the cruxes involved in defining a frame: "As for storage in memory, there is experimental evidence that the sentence content or deep structure tended to be retained, and surface form to be discarded."\textsuperscript{29} They have relegated surface structure to the status of a nuisance, an irrelevancy that has to be endured to get to the omnipotent deep structure: "Phrase structure boundaries may thus serve as cues to deep structure relationships. Note, however, that the surface structure is not accorded an independent status in the comprehension model."\textsuperscript{30} These statements typify an almost wholesale bypassing of surface structure by recent scholars, in favor of some form of deep structure. And while the components of the deep structure may vary according to particular theories— from propositions to cognitive structure to I-markers to nebulous nonlinguistic "meaning"— all these theories share the same basic premise that single sentence deep structure analysis is the first step in the analysis of total discourse structure. The operating correlative is always some variation of transformational analysis. This presumption that the deep structure of a sentence, however defined, is analogous to the structure of discourse produces serious inconsistencies arising from a tenuous initial hypothesis. Transformational grammar seeks to explain what structures a sentence; discourse and
comprehension analysis seek to explain what structures a discourse; one predicates the sentence as the unit of study; the other, the text. There is no reason in general to assume that an answer to one set of questions will be appropriate or transferrable to another set of questions; there is no reason in particular to assume that the transient, semantic/syntactic structures built during sentence construction are similar or identical to the larger structures of texts. For instance, transformational grammars are verb based, with parts of sentences seen as instantiating verb frames. But in discourse analysis one can easily imagine verbs being subordinated to nouns or topics. Similarly, these transformational theories, preoccupied as they are with the deepest of deep structures and with semantics in general, do not account for the fact that features such as articles, which for good reason are given low priority in sentence grammar, seem to be of consummate importance in discourse structure. In effect, comprehension theories modeled after transformational grammar assume discourse to be one giant sentence made up of a list of ordered propositions, and thus provide neither analytic nor synthetic means for going beyond simple sentence-by-sentence understanding to any kind of larger, assimilative structure. Such theories do not really allow for the frame.

Even frankly granting the catch-all attractiveness of deep structure sentence analysis and granting some
correlation between the deep structure of sentences and the total structure of discourse, one feels that the preoccupation with deep structure has seriously colored the few actual attempts to identify textual cues to comprehension. When faced with various possible conclusions from experiments, researchers today opt for the conclusion favoring deep structure. Kintsch, for instance, arguing for the validity of propositional analysis, concluded that his propositions are "real units" because deep-structure propositions, written as VA and VAO and representing sentences such as "The sentry yawned." and "The travelers noticed a restaurant." were remembered as units by his subjects. Ross Winterowd writing in 1980 cites Kintsch's study and says: "Fillmore (1968) explored the grammar of propositions, and Kintsch (1974) demonstrated their psychological reality." But one could just as easily conclude from Kintsch's findings that the subjects remembered surface forms, because his propositional analysis is in fact congruent with traditional surface structure analysis. One feels that something congruent with an already existing form requires more than a name change to establish independent existence, and the Kintsch is doing no more than calling janitors "custodians."

On the other hand, researchers deliberately testing the reality of surface forms have demonstrated that these
structures affect reader comprehension in both positive and negative ways. As the division of a sentence into its surface constituents is made more salient, so, too, is the sentence's comprehensibility increased. Trabasso found that sentence subjects were identified faster in active sentences than passive sentences. Multiple embedded sentences, with discontinuity of phrases, seriously impede comprehension.

Given the evidence suggesting the integrity, even if not the Platonic reality, of surface forms, an analysis of their interactive effect on comprehension and cohesion is very much in order. At the very least, surface structure analysis can develop empirical data to help refine our present models of discourse comprehension. Instead of viewing surface features as irrelevant nuisances, such an analysis assumes that they represent an intermediate stage between sentence deep structure and discourse structure. In order to discern the structure and form of a house, one does not need to decompose its bricks into their elements; rather one takes the bricks as given and tries to determine how their arrangement and order contribute to the house's stability and structure.

As I suggested a few pages earlier, although deep structure theories about comprehension regularly fill the pages of journals and books and although they claim to describe reader comprehension, few have exposed themselves
to the hard reality of empirical testing. This final section will review the few reader-tested attempts to identify various textual constituents of a frame.

The most extensive investigation in this area is derived from Kintsch's propositional analysis. Kintsch and Janet Keenan, in "Reading Rate and the Retention of the Number of Propositions," develop a concept of "superordinate" based on simple sequence. Proposition A is superordinate to proposition B if 1) proposition A precedes proposition B, and 2) if both A and B share a common term. In recall tests based on this assumption, subjects were found consistently to recall superordinate propositions better than subordinate ones. The generality of this conception of superordinate and subordinate propositions was confirmed by Waters, who tested third-grade, sixth-grade, and college students, all of whom recalled superordinate propositions better than subordinate ones.

In Kintsch's memory model, this superordinate-subordinate relationship between propositions is not equivalent to a text's frame, or "macrostructure," as he calls it; this relationship between propositions represents only one part of the macrostructure. The other set of relationships which make up a text's frame is that of proposition "importance." Gail McKoon, in "Organization of Information in Text Memory," used the following criterion for identifying "important" propositions: "So for the paragraphs used
in the experiments, intuition determined the most important proposition." Contributing to the researcher's intuition were: 1) surface structure cues, involving such devices as syntactic foregrounding and "surface order relations such as placing the topic sentence first;" and 2) "the relationship of the information in the text to the reader's knowledge of the world." Combining this definition of important with the concept of superordinate, McKoon found that subjects best remembered propositions satisfying both of those criteria.

Two researchers, Bransford and Franks, have used a non-Kintschian conception of superordinate based explicitly on syntax. As precursors to sentence-combining, they conceive of "subordinate" as "holistic semantic ideas," the nature of which is best understood, I think, by an example. "The rock which rolled down the mountain crushed the tiny hut at the edge of the woods" represents the holistic unit for the individual sentences: "The rock rolled down the mountain;" "The rock crushed the hut;" "The hut was tiny;" "The hut was at the edge of the woods." Bransford and Franks found that subjects given such a series of four simple sentences did not identify as "new" any complex sentence containing three of the four simple sentences.

The final attempt to identify textual cues for a frame comes from E. J. Crothers. In a more ambitious approach, Crothers, in "Memory and the Recall of Discourse," tried to
develop frames that incorporated both the implied rhetoric and the thematic content of a passage.\textsuperscript{40} In his system the rhetoric of a passage represented the superordinate node or frame, and was symbolized by logical connectives such as IS, WHY, OR, AND and IF. These framing words were not stated in the text but deduced by Crothers. The actual content of the passage, and the words or concepts representing that content, were treated as subordinate to these connectives. Within this framework, Crother's expectations that superordinates would be recalled more often than subordinates was not confirmed; nor was the corollary hypothesis confirmed, that secondary subtrees would be recalled less often than primary subtrees.

In discussing his negative results, Crothers concedes that factors not measured by his system, particularly term repetition, probably contribute to a reader's apprehension of a frame. Thematization, or the establishment of a referent through associated noun phrases, has in fact been demonstrated to contribute to the reader's sense of totality. Charles Perfetti and Susan Goldman tested subjects for their recall of thematized referents as against non-thematized referents. They note that in their experiment it is not simply a word that is thematized by a referent: a passage that identifies Dr. Jones and "relates that he is a surgeon, is thematizing a referent that is realized by the expressions Dr. Jones, the surgeon, the doctor, he and other noun
phrases identifying the referent. Under these circumstances, the thematized referent became a more reliable recall cue than non-thematized referents.

One of the teasing implications of Crother's negative results is the never discussed question of the upper limit to a frame. Just as the frame must be broad enough to encompass and handle all relevant aspects of a text, likewise that frame must be concrete and real enough not to evaporate into useless generalities. To illustrate this problem, we can look at the following passage which is based on but does not literally describe a picture. The passage, but not the picture on which it was based, was given to subjects who were then tested for passage recall; not surprisingly, they did quite poorly even though the passage is cohesive, i.e. has Crother's IF-THEN connectives, term repetition, and the superordinate concept "problem:")

If the balloons popped the sound wouldn't be able to carry since everything would be too far away from the correct floor. A closed window would also present (sic) the sound from carrying, since most buildings tend to be well insulated. Since the whole operation depends on a steady flow of electricity, a break in the middle of the wire would also cause problems. Of course, the fellow could shout, but the human voice is not loud enough to carry that far. An additional problem is that a string could break on the instrument. Then there would be no accompaniment to the message. It is clear that the best situation would involve less distance. Then there would be fewer potential problems. With face to face contact, the least number of things could go wrong.
In retrospect, we can see that the shared constant in these experiments is term repetition. Although only Perfetti and Goldman in their thematization study explicitly tested the effect of repetition, item recurrence contributed to the success of Kintsch's work, to that of his disciples', and to Bransford and Franks', whereas the failure of Crothers' taxonomy to provide for item recurrence probably accounted for his negative results. Even though the experiments ostensibly tested term order and synthesis, the obvious prerequisite to such measurement is the actual presence of a repeated term(s). Limited as these results are, they confirm chapter one's hypothesis that cohesion consists primarily of unity, the presence of a repeated term.

Furthermore, these experiments, and others cited throughout the chapter, confirm the hypothesis that unity dominates coherence. A minimum and necessary threshold of cohesion is attained through the unity created by the repeated term, so that comprehension can be impeded but never fully blocked by "bad" or inappropriate order.

Because of the theoretical assumptions involved in these experiments, only one kind of order was tested, the basic chain-like transitivity established by sequences in the form A is B; B is C and so on—or what Frase labeled in his experiment as "good order." As the results of the experiments indicate, such an order clearly does exist, and does aid in comprehension and cohesion. In an independent
textual study, Danes identified this same order, labeling it as "simple linear progression," in which the rhyme of one sentence (basically the sentence's predicate) becomes the theme (subject) of the next. However, he also identified three other common recurrence patterns:

1) passages with run-through themes (a sequence of sentences with the same theme but different rhyme);
2) progression of derived theme and
3) the development of a split-rheme (the themes of successive sentences are co-members of a concept forming the rhyme of the initial sentence).

Not only do such additional patterns of item recurrence exist, but the actual contribution of the A is B order to cohesion must be qualified. Of the following sentence sequences, only the first one is cohesive:

The process of learning is essential to our lives. All higher animals seek it deliberately. They are inquisitive and they experiment. An experiment is a sort of harmless trial run of some action we shall have to make in the real world.

(J. Bronowski, The Common Sense of Science, p. III)

I don't like you. You are a breadboard. Breadboards are made of wood. Wood is grown in forests.

As the sequences illustrate, at least part of the cohesive power of the A is B order derives from the semantic relations of the item recurrences themselves, thus the explanation for any success in the cited experiments must be modified: if some semantic compatibility or logic already exists between item recurrences, the A is B order will then facilitate cohesion. The implication of this necessary
modification is that while cohesion consists of both unity and coherence, not only does unity dominate coherence, but coherence is a function of unity which was another hypothesis of chapter one.

Finally, even though all the experiments tested only the A is B, B is C order, different researchers, working with different texts, were able to abstract a recurrence pattern--the core of cohesion--so that chapter one's hypothesis that cohesion patterns manifest structural and not simply semantic totalities is not only confirmed, but partially instantiated.

As a concluding reflection on the psycholinguistic work reviewed in this chapter, one must feel simultaneously a meagerness and a plentitude: a plentitude because obviously the review selected only the most germane articles from literally hundreds; a meagerness because the articles tend only to confirm one's intuitive ideas about reader comprehension. (It is this meagerness that I suspect contributes to the profusion of comprehension theories.) I think, too, one must feel a certain self-irony about attempts to formalize the tremendous synthetic and analogic capacities of the human mind in one of its most particularly subjective and personal experiences:

A text is timeless, universal, objective, speaks to anyone who is literate, and reveals to him a self-contained meaning. Yet this objective and universal character can be realized only through
the subjectivity of some reader; thus the burden of interpretation. Though a text is always about a world or a possible world, it does not refer to a world outside itself, for its world is created in the dialectic between itself and the reader...
Note

The distinction between lexical and world knowledge in this division is admittedly a shadowy one. Bolinger questions the existence of any distinction between "knowledge of language" formalized by semantic markers and "knowledge of the world." Wilson comments: "There is no sharp line between what properly belongs in a dictionary and what properly belongs in an encyclopedia." Undoubtedly, categorizing knowledge is a complicated affair, but examples do exist that suggest the validity of two separate categories. Katz provides one such example with his analysis of the following three sentences:

- Bachelors are male.
- Bachelors are not married.
- Bachelors are over one inch tall.

As Katz points out, readers reject the negations of sentences one and two based on the lexical definition of bachelor, while rejecting sentence three's negation has nothing to do with one's sense of bachelorhood.
NOTES TO CHAPTER TWO

1Jonathan Culler, op. cit., p. 171.


3Freedle and Carroll, p. 361.


10Van Dijk, op. cit. p. 50.

11Van Dijk, op. cit. p. 50.


13Haviland and Clark.


16Ruhl, op. cit. 15.

17Ruhl, op. cit. 15.


20Crothers, op. cit. 273


28 Rommetveit, p. 29.
29 Schlesinger, p. 139.
30 Schlesinger, p. 138.
31 Kintsch, 1974, op. cit.


34 Schlesinger, 1975, pp. 53-66.

35 Walter Kintsch and Janet Keenan, "Reading Rate and Retention as a Function of the Number of Propositions in the Base Structure of Sentences." Cognitive Psychology, 5 (1973).


38 McKoon, 248.


40 Crothers, 247-283.


42 Bransford and Johnson, 393.

43 Danes in Enkvist, op. cit. pp. 120-21.


CHAPTER THREE: COHESION IN THE SINGLE TERM PARAGRAPH

As the previous two chapters have indicated, there is both experimental evidence and theoretical motivation for positing the existence of some kind of global structure for discourse units—something akin to the totalities of chapter one and the frames of chapter two. The paragraph, by virtue of its traditional definition as a discourse unit that develops a single idea, offers itself as a logical choice for exploring, testing, and articulating the concept of the totality. Thus the aim of this chapter is to test the hypothesis that the paragraph is a totality, and to identify the linguistic cues that produce cohesion patterns, or totalities. A variety of simple to complex single-term paragraphs will be analyzed to establish a range of possibilities for cohesion. These will include both paragraphs in which the dominating term is explicit and those in which it is implicit. Such an analysis should further our understanding of the nature of paragraphs and their formal requirements.

To provide a context for the actual paragraph analyses, this section will briefly review paragraph research that has attempted to identify factors contributing to a paragraph's wholeness. As a formal structure, the
English paragraph received little attention until 1877, when Alexander Bain published his *English Composition and Rhetoric* and included a forty-five page chapter on the paragraph, making Bain the first rhetorician to treat the unit formally. Although rhetoricians in his own time largely ignored his work, later rhetoric and composition books canonized Bain's definition and description of the paragraph. Bain's forty-five pages molded and shaped a view of the paragraph that has lasted essentially unchanged for a hundred years.

As Bain defined it, the paragraph is "a collection of sentences with unity of purpose . . . that handles and exhausts a complete topic." Among his six paragraph rules, the third recommended that the opening sentence "indicate with prominence the subject of the paragraph," thus giving rise to the topic sentence. Over the years, Bain's definition has become synonymous with the textbook paragraph and the staple of composition and rhetoric courses; "later rhetoricians tinkered with the language of this (Bain's) definition, but without improving upon or substantially changing the original insight."

Yet like the concept of the totality or the frame, the concepts of the paragraph and the topic sentence, so intuitively obvious, defy unambiguous definition. What constitutes unity of purpose? What constitutes a complete idea? The most basic insights are also the most resistant to
definition. One finds paragraphs with apparently more than one main idea, paragraphs with implied main ideas, multiple paragraphs that can be collapsed into one, paragraphs of one sentence and paragraphs of pages. Faced with this onslaught of qualifications and appendages, rhetoricians have begun to question, if not reject, the paragraph as a viable unit:

Since every paragraph of the essay is part of the general flow, it is difficult to find in many paragraphs anything so static that it can be isolated as the single idea, or topic, of that paragraph. The notion that every paragraph must have a topic sentence is hence misleading.

(The paragraph) is simply a convenient grouping of sentences. In a progression of sentences a few places will be more suited to indentation than others, but you can justify indentations before almost any sentence of sophisticated prose.

Reporting on an informal experiment in "When Is A Paragraph?" Arthur Stern describes the results when he asked his students to divide an excerpt from Brook's and Warren's Fundamentals of Good Writing. The results were erratic: "... some students divided the passage into two paragraphs, others into three, still others into four or five ... only five students out of more than 100 who have tried the experiment have paragraphed the passage precisely as Brooks and Warren originally did.""5

Confirming these results were the findings of a research project by Richard Braddock, "The Frequency and Placement of Topic Sentences in Expository Prose." Braddock
examined a corpus of twenty-five essays published in magazines such as The Atlantic Monthly, Harper's, and The New Yorker for four kinds of topic sentences: the simple direct statement; the delayed completion, in which the topic sentence is initiated by one T-unit but completed by another; the assembled topic sentence, in which the topic sentence is built from phrases and clauses through the text; and finally, the inferred topic sentence in which the topic sentence is implied even though it cannot be directly assembled from the original passage. He concluded: "It is just not true that most expository paragraphs have topic sentences in that (simple direct) sense. Even when simple and delayed-completion topic sentences are combined into the category 'explicit topic sentences'... the frequency reaches only 55% of all entries..."\(^6\)

Yet counterbalancing Braddock's results is the positive one obtained by Frank Koen, Alton Becker and Richard Young. In a study similar to Stern's but much more statistically and rigorously administered, they asked students to paragraph uninterrupted texts. From the results, they concluded: "It is apparent that the Ss (subjects) agree with each other in their judgements of paragraphing boundaries in both English and nonsense passages. Paragraphing, then, is a reliable phenomenon..."\(^7\)

Given these sometimes conflicting views and results, most current researchers working on paragraphs have discarded
any Platonic, if not Procrustean, view of THE paragraph and instead seek to identify and describe the various different kinds of paragraphs. The remainder of this chapter continues that approach.

As outlined in chapter one, the minimal requirement for cohesion is unity—a sentence sequence dominated by a single term. Besides embodying one of Culler's totalities, the series united by a common term, unity also constitutes the fundamental characteristic of a paragraph, that it be about one thing. The rest of this chapter will examine both simple and complex patterns of single term dominance in paragraphs, and seek to explain how this dominance is attained. While such an analysis will not reveal all the mysteries of the paragraph, it will provide concrete information about the patterns of structure and cohesion in one basic and very common paragraph type.

The series united by the common term represents the simplest form of intransitive stative cohesion, attainable through the most rudimentary kind of unity, and also the simplest paragraph structure. This series is created by a sequence of sentences sharing a single term so as to create a pervasive chain of referents and antecedents. The following example is patently dull and boorish:

The Char-Bar is a bar on High Street. The Char-Bar swings. It permits dancing. The bar specializes in foreign beers. The Char-Bar attracts weirdos. It seats 198 people.
But it fulfills the basic requirements of unity and hence cohesion and in fact constitutes the basic structural paradigm for many more sophisticated paragraphs.

Dominance is attained not simply by the mere presence of the term Char-Bar or its pronoun substitute in each sentence but by its appearing consistently in the subject position. This condition—that of the repeated term appearing in the subject position—is the simplest fulfillment of the concept of two described in chapter one: that two somethings are required for a relationship. If a term is repeated in each sentence but does not appear in the subject position, the concept of two is not fulfilled, dominance is not attained, and the sequence is not cohesive:

Alfred likes peaches. Oregon doesn't grow peaches. Peaches contain nitrogen. We have a peach tree in our backyard. No one throws rotten peaches at politicians or ball players. Cut five peaches and sprinkle sugar on them. Do you think peach melba would be a good dessert?

Once a repeated term appears in the predicate position, it forfeits the inherently limiting power of the subject position and is itself "subjected" to at least five other topics: Alfred, Oregon, we, no one, you.

This is a very clear example of syntax being more important, or at least as important, as semantics in creating cohesion. Halliday and Hasan's sweeping statement that cohesion is only a semantic relation is wrong. 8
Since we are pursuing the hypothesis that cohesion involves the relationships between parts and simultaneously that a description of cohesion will yield a description of a paragraph's structure—intuitively understood as also involving relationships between parts—we can begin to develop from the preceding analysis the notion of paragraph structure.

The phrase "paragraph structure" has a very winsome ring and appears in composition texts and discourse analysis articles as frequently and casually as intoxicated flirting at cocktail parties. As a concept, paragraph structure is itself very intoxicating, and also very difficult to define. Here is a typical definition:

The paragraph may be defined as a sequence of structurally related sentences. By a sequence of structurally related sentences I mean a group of sentences related to one another by coordination and subordination. If the first sentence of a paragraph is the topic sentence, the second is quite likely to be a comment on it, a development of it, and therefore subordinate to it. . . . A sentence that is not coordinate with any sentence above it or subordinate to the next above it breaks the sequence.9

Christensen's definition, one feels, is essentially correct and even exciting, but frustratingly indefinite in its descriptions of subordinate and coordinate. My own preceding analysis evolves from Christensen's basic perception and can, I think, lead to a more specific understanding of "structurally related sentences" if we identify the assumptions made in that analysis.
First, the repeated term was identified within the sentence's syntax. Specifically, in the Char-Bar paragraph, the repeated term appeared in the sentence's primary noun-phrase which functions as the sentence's subject. Second, the relations between the repeated items themselves—in this case simple identity and reference—were taken into consideration. As noted in chapter one, however, simple identity is not the only relation that will establish an equivalence chain; equivalence, synonyms, and general set relations may also create a chain. The interaction of these two sets of relations, one semantic and the other syntactic, establishes term dominance and cohesion. The semantic set of relations, or the equivalence chain, narrows, broadens, or refines the actual discourse topic, while the syntactic information contextualizes and conjoins that topic with other terms, arguments, or concepts within the paragraph. When, as in the Char-Bar paragraph, the repeated term appears consistently in the subject position, or the dominant noun phrase slot, and when the repetition constitutes a straightforward equivalence chain, we have the simple but necessary paradigm for the single term paragraph wherein the syntactical information concurs with and reenforces the semantic. That is, the sheer repetition of a single term signals semantic thematization, as documented in chapter two; the term is then syntactically thematized by appearing consistently in the subject position or the dominant noun phrase slot. When these two
thematization methods concur, the repeated term evolves into the discourse topic.

If one accepts the premise that cohesion, as the relationships between discourse parts, is the avenue through which we can approach paragraph structure, the preceding analysis provides a framework for describing paragraph structure in other than rhetorical terms. It identifies the paragraph's parts, both semantic and syntactic, along with their multiple relationships. From a more general perspective we can view the Char-Bar paragraph as an exemplar of Culler's totality of a series united by a common term in which the common term is represented by the equivalence chain and the series by the material predicated to that term in each succeeding sentence. The appearance of the shared term in the dominant noun phrase position is critical, since it creates the series.

Obviously not all paragraphs read as crudely and abruptly as the Char-Bar text. The next few pages will examine variations of this basic pattern, paragraphs that move beyond the minimal requirements into more complex and interesting manifestations of the paradigm structure.

At this stage, we can look at a paragraph in which the dominant term is not literally repeated but is inferrable from the text. The use of ellipsis, which in turn requires the reader to make inferences, is very common to description paragraphs such as the following by Truman Capote:
The truck was of the Ford pick-up type.
Its interior smelled strongly of sun
warmed leather and gasoline. The broken
speedometer registered a petrified twenty.
Rainstreaks and crushed insects blurred
the windshield of which one section was
shattered on a bursting star pattern. A
toy skull ornamented the gear shift. The
wheels bump-bumped over the rising,
dipping, curving Paradise Chapel Highway.

Written without discourse ellipsis, the passage would read
something like this:

The truck was of the Ford pick-up type.
The truck had an interior. The interior
smelled strongly of gasoline. The truck
had a speedometer. The broken speedometer
registered a petrified twenty. The truck
had a windshield. Rainstreaks and crushed
insects blurred the windshield of which one
section was shattered in a bursting star
pattern. The truck had a gear shift. A
toy skull ornamented the gear shift. The
truck had wheels. The wheels bump-bumped
over the rising, dipping, curving Paradise
Chapel Highway.

Even in the expanded text, no single term appears in
every sentence, although "truck," makes a strong bid for
attention by its presence in six of the eleven sentences.
The absence of a repeated term immediately entails a
weakening of the basic paradigm: if no term appears in each
sentence, the term repeated most frequently becomes the
basis for the text's equivalence chain. While such a qualifi-
cation may have to be invoked eventually, it is not yet
the only recourse because a transformational analysis pro-
vides us with the recurring term "truck" in each sentence.
If we assume that definite articles are reduced possessives,
then we can reconstruct sentences such as: "The truck's interior smelled strongly of sun warmed leather and gasoline," "The truck's speedometer registered a petrified twenty." and so on. This analysis is actually consistent with the earlier one since the same bridging assumptions are assumed, but in this instance they are presumed to be parts of the total sentences, albeit in reduced forms, instead of individual sentences.

Having established a repeated term in each sentence, we need to determine the degree of concord between that semantically important element and its syntactic prominence. Continuing with the transformational analysis of the passage's sentences, we have no trouble: the designated bridging assumption is in the highest, hence dominating sentence node. Thus the sentences, "Rainstreaks and crushed insects blurred the windshield, of which one section was shattered in a bursting star pattern," and "A toy skull ornamented the gear shift," where the implied "truck" is in the object position, nevertheless exhibit concord between the syntactically dominant position and the repeated term in the deep structure, although not in the surface structure. "Although not in the surface structure" needs some qualification: since the truck parts are in direct object positions, they could be passivized and thus move into the subject positions:
The windshield, of which one section was shattered in a bursting star pattern, was blurred with rainstreaks and crushed insects.

The gear shift was ornamented by a toy skull.

Before looking at other pattern variations, I want to pause and reflect on some implications of the preceding analysis. Happily, Capote did not write the reconstructed version my analysis was based on. If he had, the result would have been another ponderous Char-Bar paragraph. Instead he wrote a self-sufficient but abbreviated version of the deep structure version, reminding us that all discourse is ellipsis. What Capote also did was to maintain the delicate and dynamic balance between suggestive repetition—the core of unity and cohesion—and boring repetition. And the paragraph was transformed from a skeletal paradigm to an artful and stylish text.

A second point is that in neither the Capote paragraph nor the Char-Bar paragraph did any of the intermediate sentences contain direct links with their immediate predecessors. A well wrought, unified and stylish paragraph did not lock each sentence into position through some immediate repetition or transition word; indeed, individual sentences could easily be reordered; finally, the paragraph did not fulfill Hasan and Halliday's prediction that sentence links, repetitions basically, would cluster at a text's center, so
that the number of sentence links might be graphically represented thusly:

\[
\begin{array}{c}
\text{more} \\
\text{ties} \\
\text{fewer}
\end{array}
\begin{array}{ccc}
\text{paragraph} & \text{paragraph} & \text{paragraph}
\end{array}
\]

Instead we find fulfilled Lathrop's prediction of unity of connection emerging from and shaped by the total structure. Cohesion is a general effect for which there is no simple prescription. Different kinds of textual totalities will have different cohesion patterns, and, as Christensen noted, we will need different yardsticks for measuring and identifying these patterns, which also involve different kinds of paragraph structures.

Both of the preceding examples were paragraphs of description, but the paradigm of a series with a common term can also describe paragraphs traditionally labeled as paragraphs of example, suggesting that we are dealing with a fundamental method for describing paragraph structure and classification. The following paragraph manifests the single term pattern in which the predicated content is examples:

Admen and packagers, of course, are not the only euphemizers. Almost any way of earning a salary above the level of ditchdigging is known as a profession rather than a job. Janitors for several years have been elevated by image conscious unions to the status of "custodians;" nowadays, a teen-rock
guitarist with three chords to his credit can class himself with Horowitz as a "recording artist." Cadillac dealers refer to autos as "preowned" rather than "secondhand." Government researchers concerned with old people call them "senior citizens." Ads for credit and department stores refer to "convenient terms"--meaning 18% annual interest payable at the convenience of the creditor.13

As in the Capote paragraph, here the repeated term is implied rather than stated. Through some back reconstruction, we can identify it as some form of euphemizers. Every reader, I think, will agree that sentences two through six give examples or instances of the generalization in sentence one, based on our dictionary knowledge of "euphemizers," and she will mentally insert "for example" or "for instance" before each of those sentences. Such terms as "for instance" or "for example" are transition words and as such operate as whole sentence modifiers; transformationally, they are members of the sentence's highest syntactic constituent, just as possessives were in the preceding example. However, if there is to be concord between the highest syntactical constituent and the repeated term, the repeated term must occupy that constituent. The implied "for example" or "for instance" presently holding that slot are not "euphemizers." The bridging assumption that first posited a simple "for example" must now be reexamined: "for example" is not a complete assertion, that is, it is not a complete predication or argument, so we can presume it to be a
condensed form of the complete bridging assumption. For the analysis to conform with and reflect our intuitive understanding of the text, we need a complete bridging assumption that responds to the question "example of what?" with the answer "That there are other euphemizers." This complete bridging assumption, or some version of it, prefixes and contexts sentences two through six, with "for example" or "for instance" acting as that assumption's condensed placeholder. (Instead of viewing phrases like "for example" or "for instance" as elliptical forms of phrases, alternatively one might view such phrases as discourse substitutes along the lines of "one" in sentence analysis.) Sentence two needs some further qualification through an additional inferencing step since no "euphemizer" is stated—as in the other sentences which name unions, guitarist, cadillac dealers, government researchers—and the reader must infer "society" or "people in general" as the actual euphemizer. While I think that the actual bridging assumptions used by readers may vary with different readers and that one can debate the nature of the actual steps in tracking "for example" to "euphemizer," each individual's analysis will always lead back to "euphemizers." This common conclusion suggests that the euphemizer paragraph is yet another version of the series united by the common term, even though textbooks would classify it as exemplification rather than description.
In preceding analysis, no acknowledgement was made that "euphemizers" first occurs in the predicate position, in contrast to the earlier samples in which the repeated term consistently appeared in the subject position and/or in the highest syntactical position. In this paragraph, the repeated term first appears in the predicate position and then consistently in the highest syntactical position. This is not simply an ad hoc exception to the concept of dominance earlier defined as concord between a repeated term's location and the syntactically highest constituent, only a refinement of it. As the peaches non-text demonstrated, in a single term sentence sequence, given sentences in which the subject and predicates are respectively in the highest noun phrase and verb phrase slots, the repeated term cannot consistently appear in the predicate position because it forfeits its limiting power of sentence subject. However, since a sentence's subject and predicate are syntactically co-equal, the euphemizer paragraph did not violate the definition of dominance but merely refined it.

One should remember that the first sentence in any text does not predict what will follow; our sense of pattern or structure develops as we proceed through the text. We could have encountered a sequence like:

Admen and packagers are not the only euphemizers. But they are the most influential and ubiquitous. They attack our senses and form our opinions about "products" that once were not products, like presidential candidates.
Barthes describes this assembling process thusly:

Whoever reads a text collects bits of information under the generic names of actions (Walk, Assassination, Rendez-vous), and it is this name which creates the sequence. The sequence comes to exist only at the moment when and because one can name it; it develops according to the rhythm of this naming process, which seeks and confirms.14

Certainly, some paragraphs do exhibit inconsistent placement of the repeated term as it shifts between the subject and predicate position. The degree to which readers would find such paragraphs cohesive is ultimately a function of quantity: the more appearances in the subject position, the more likely the passage will be accepted as cohesive. Presumably a statistical study of such shifts would yield a set of formulas: a paragraph of five sentences requires the repeated term to occupy the subject position three times and so on. Such a study would presumably reflect the skew from the paradigm and help distinguish borderline cases of cohesion.

A second argument and perspective for viewing the predicate initial position as a refinement of the basic paradigm lies with the sentence's discourse function. Discourse involves thematic progression (although the samples thus far examined are not primarily defined through their transitivity) and progression is attained through the conjoining of arguments. At the sentence level, the grammatical functions of subject and predicate accomplish this progression with the subject understood as what is being talked about and the
predicate as what is being said about the subject. Discourse analysts have developed various terminologies for this condition: theme/rheme; known/unknown; given/new. (However, since the defining of these states for actual sentences has proved difficult, I will continue using the subject/predicate distinction.) Despite the differences in terminologies, each schema seeks to capture a sentence's message structure as it progresses from the given, the known, the understood to the new and the unknown. John Erskine describes this fundamental characteristic of communication this way:

What you wish to say is found not in the noun but in what you add to qualify the noun. The noun is only a grappling iron to hitch your mind to the reader's. The noun by itself adds nothing to the reader's information; it is the name of something he knows already, and if he does not know it, you cannot do business with him.*5

Yet just as sentences move from the known to the unknown, the unknown becomes the known and shared, and a new progression is initiated. The particular sequence of Subject₁/Predicate₁ to (Subject₂/Predicate₂) or A:B; B:C provides the ideal pattern for this movement; it duplicates Frase's documented conception of "good order" and rhetorically it conforms with basic information psychology of introducing the unknown through the known. This particular shift from subject to predicate to subject is, as we shall see shortly, one of the most common forms of paragraph juncture for this
paragraph structure. Besides opening paragraphs, this sequence just as frequently closes them, providing in both cases content transitions with adjacent paragraphs.

Thus far the sample paragraphs exhibiting the series united by the common term have not been characterized by the usual devices of coherence, namely direct links between sentences; in each sample, sentences could be moved without affecting the reader's sense of cohesion. This particular structure does not, however, prohibit immediate links between sentences, as suggested by the following paragraph, one that probably represents structurally the most interesting and complex variation of this paragraph pattern. We return to euphemisms:

From a Greek word meaning "to use words of good omen," "euphemism" is the substitution of a pleasant term for a blunt one--telling it like it isn't. Euphemism has probably existed since the beginning of language. As long as there have been things of which man thought the less said the better, there have been better ways of saying less. In everyday conversation, the euphemism is, at worst, a necessary evil; at its best, it is a handy verbal tool to avoid making enemies needlessly, or shocking friends. Language purists and the blunt spoken may wince when a young woman at a party coyly asks for directions to the "powder room," but to most people this kind of familiar euphemism is probably no more harmful or annoying than, say, the split infinitive.16

In this paragraph, we have something a bit more satisfying to most readers than the earlier samples; instead of a text in which the bare bones of the structure rattle about for
our attention, here we find the structure fleshed out, more robust but also less conspicuous. Even a cursory examination reveals a repeated term--"euphemism"--in each sentence, although sentence three substitutes a phrase synonym, "better ways of saying less." "Euphemism" occupies the subject slot in each sentence except sentence three which again has the empty "there" in the subject position.

Two qualities distinguish this paragraph from its predecessors: it conveys more information through more complex and heavily modified sentences and it uses that additional information to link each sentence to its immediate neighbor. These sentences cannot be moved around as easily as in the earlier examples, if, indeed, they can be moved at all.

Chapter one predicted increased coherence through more complex sentence constructions, so it won't be elaborated on here. Instead, we will look at the form and effect of the immediate sentence links. In any of the earlier paragraphs, the authors could have chosen to head each sentence with an enumeration transition word, such as "first," "second," "next," and so on. There is little reason and less desire to believe these insertions would have enhanced the passages. They would have been superfluous to reader comprehension, and with their wide and general applicability convey little information in-and-of themselves. They would have been stuck in mechanically as vestiges of a Freshman Composition lesson. However, inserting any of the traditional transition
words before the sentences of this example (with the exception of "for example" or "for instance" before the last sentence) would be semantically incongruous and illogical. Here the immediate sentence links are attained through message-contributing adverbials, such as "since the beginning of time," "as long as," "in everyday conversation," "at worst," and "at best." Besides conveying additional information, these adverbials lock individual sentences in place and create mini-contexts within the larger series pattern for the adjacent sentences, even though the adverbials themselves do not enter directly into the main series pattern of "euphemism." Barthes observed this phenomenon in larger discourse units, labeling the two patterns, "satellites" and "kernels." He distinguished "between 'kernels' which link up with one another to form plot and 'catalysts' or 'satellites,' which are attached to kernels but do not themselves establish sequences."¹⁷

Of further significance is the fact that these links modify the entire sentence unit, "euphemism" predicate, rather than appearing as forms incorporated with that unit. For this reason, they do not interfere with reader comprehension of that basic unit, in contrast to embedded modification modes. Further study is needed to reveal the extent to which there are limits to modification incorporation, but this particular method does seem to be common among successful authors. The following Orwell paragraph, built on the series pattern, uses the same modification placement:
Mrs. Simpson was nicknamed Bingo, and I shall call her by that name for I seldom think of her by any other. (Officially, however, she was addressed as "Mum", probably a corruption of the "Ma'am" used by public school boys to their housemasters' wives.) She was a stocky, square-built woman with hard red cheeks, a flat top to her head, prominent brows and deepset, suspicious eyes. Although a great deal of the time she was full of false heartiness, jollying one along with mannish slang ("Buck up, old chap!" and so forth), and even using one's Christian name, her eyes never lost their anxious, accusing look. It was very difficult to look her in the fact without feeling guilty, even at moments when one was not guilty of anything in particular.

As readers can easily see for themselves, once again we have the series pattern, but we also have the modification pattern in a descriptive paragraph that was just described for the preceding expository paragraph. Although further study would be necessary before drawing any final conclusions, this duplication (repeated in many other paragraphs as well) suggests that the basic paradigm of the series united by a common term informs and directs appropriate modification structures, so that the alignment between form and function extends to aspects of sentence structure beyond subject placement and choice. Clearly, stylish and cohesive texts do exist that employ embedded modification structures, but inevitably any such embedding would affect the scope and nature of the repeated term and hence the nature of the equivalence chain and ultimately the paradigm type.
These final pages will consider the effectiveness and the validity of the preceding concepts by applying them as an evaluative standard to paragraphs of mediocre cohesion but which are nevertheless comprehensible.

As a test paragraph we can consider the following Momaday text, one that may strike readers as attractive on a first reading because of its descriptive and allusive qualities, but which finally, I think, leaves the reader with a sense of looseness and unraveling, of weak cohesion:

1) A single knoll rises out of the plain in Oklahoma, north and west of the Wichita Range. 2) For my people, the Kiowas, it is an old landmark, and they give it the name of Rainy Mountain. 3) The hardest weather in the world is there. 4) Winter brings blizzards, hot tornadic winds arise in the spring, and in summer the prairie is an anvil's edge. 5) The grass turns brittle and brown, and it cracks beneath your feet. 6) There are green belts along the rivers and creeks, linear groves of hickory and pecan, willow and witch hazel. 7) At a distance in July or August the steaming foliage seems almost to writhe in fire. 8) Great green and yellow grasshoppers are everywhere in the tall grass, popping up like corn to sting the flesh, and tortoises crawl about on the red earth, going nowhere in plenty of time. 9) Loneliness is an aspect of the land. 10) All things in the plain are isolate; there is no confusion of objects in the eye, but one hill or one tree or one man. 11) To look upon that landscape in the early morning, with the sun at your back, is to lose the sense of proportion. 12) Your imagination comes to life, and this, you think, is where Creation was begun.

The paragraph opens with a sputter and indecision, as the first candidate for the repeated term, "knoll," appears
as subject, again subject, and then predicate, meanwhile vying with "Kiowas," which appears in a superordinate position and then the subject position of the compound sentence (2). Sentence (3) partially resolves this confusion with its anaphoric "there," which is inferred at first as referring to "knoll." That inference is sustained until the appearance of "prairie" in sentence (4), which necessitates a re-inferencing either to "plain" in sentence one, or to some general concept such as locale including both the knoll and the plain.

Sentences five through eight run the reader through another disconcerting string of inferences: the "grass" of sentence five is apparently the prairie's grass; sentence six implies that the prairie also has rivers and trees, an ecological and therefore semantic incongruity; sentence seven's bridging assumption returns the reader to the "grass" of sentence five—violating the ordering principle of chapter two on inferences—and then implies a further bridging assumption, that "the prairie has red earth." This is a disparate collection rather than a consistent series of bridging assumptions: The prairie has grass, a defining characteristic of prairies; the prairie also has trees and rivers and creeks, an apparently anomalous feature of this particular prairie; and the prairie has red earth, a characteristic of this particular prairie. This collection contrasts sharply with the bridging assumptions in the Capote
and euphemism paragraphs, which were not only consistent with one another but which manifested a second set of relations by means of that consistency: parts of a truck and types of euphemizers. Consistency in the bridging assumptions seems a very conspicuous quality, if not a necessity, for strongly cohesive implied repeated term paragraphs. One can contrast the Momaday paragraph's bridging assumptions with those in the following Tuchman paragraph which consistently reflect "effects of the plague:"

The plague raged at terrifying speed, increasing the impression of horror. In a given locality it accomplished its kill within four to six months, except in the larger cities, where it struck again in the spring after lying dormant in winter. The death rate in Avignon was said to have claimed half the population, of whom ten thousand were buried in the first six weeks in a massive grave. The mortality was in fact erratic. Some communities whose last survivors fled in despair were simply wiped out and disappeared from the map forever, leaving only a grassed-over hump as their mortal trace.20

In contrast to the Momaday text, the Tuchman paragraph contains a clear sequence of bridging assumptions: the death rate of the plague, the mortality from the plague, survivors who fled in despair from the plague--whose very constancy add to the paragraph's cohesion.

The Momaday text's other inconsistency--the failure of the repeated term to appear regularly in the superordinate syntactic position--is probably the most important single
A weasel word is "a word used in order to evade or retreat from a direct or forthright statement or position." (Webster). In other words, if we can't say it, we'll weasel it. And, in fact, a weasel word has become more than just an evasion or retreat. We've trained our weasels. They can do anything. They can make you hear things that aren't being said, accept as truths things that have only been implied, and believe things that have only been suggested. Come to think of it, not only do we have our weasels trained, but they, in turn, have got you trained. When you hear a weasel word, you automatically hear the implication. Not the real meaning, but the meaning it wants you to hear. So if you're ready for a little re-education, let's take a good look under a strong light at the two kinds of weasel words.21

Even though "weasel word" appears in all the sentences, as a repeated term, it never quite attains dominance because of its competition from "we" and "you" for subject position in the various sentences. A less obvious competitor is the author who inserts himself through transition word phrases: "in fact" and "come to think of it." These transitions fall into the category of attitudinal disjuncts, common devices for conveying authorial interventions. As a result, "in fact" and "come to think of it" dominate two out of ten sentences in which the author places a self-reflexive and self-highlighting phrase in the sentences' highest syntactic position, effectively blocking "weasel word" from fulfilling
the second criterion for term dominance: appearing in the subject position and/or highest syntactic position.

The structural and then semantic inappropriateness of these transition words add credence to the earlier hypothesis that paragraph structure, no less than paragraph content, controls and predicts appropriate transition words.

The shifting and intermingling of terms in a sentence's highest syntactic slots ultimately affects a reader's sense of paragraph juncture, or closure, a concept vital to any overall view of paragraph structure. Where and how does a paragraph conclude? Some pages ago I suggested that shifting a repeated term from the subject position to the predicate position and vice versa frequently signals closure for this kind of paragraph structure. In the weasel word paragraph, as just noted, it was the author's mishandling of this shift that contributed to the paragraph's weak cohesion. But it is important to recognize the power of this pattern, even when it is used medially rather than terminally; regardless of the length of pattern development, that is, the number of actual sentences involved before the shift from subject to predicate position, a unit is still created:

We've trained our weasels. They can do anything. They can make you hear things that aren't being said, accept as truths things that have only been implied, and believe things that have only been suggested.

Indeed, because of its power and because of its potential to either introduce a term or to develop a term, this pattern
enables one to predict paragraph junctures in single term paragraphs, such that given the following paragraph, readers would either accept it as a whole unit or else paragraph it at "In so far as we are..." or at "This is the essence of rational behavior:"

A sign is anything that announces the existence or the imminence of some event, the presence of a thing or a person, or a change in the state of affairs. There are signs of the weather, signs of danger, signs of future good or evil, signs of what the past has been. In every case a sign is closely bound up with something to be noted or expected in experience. It is always part of the situation to which it refers, though the reference may be remote in time and space. In so far as we are led to note or expect the signified event we are making correct use of a sign. This is the essence of rational behavior, which animals show in varying degrees. It is entirely realistic, being closely bound up with the actual objective course of history—learned by experience, and cashed in or voided by future experience.22

Just as the first sentences define a sign, with congruent placement of the term in the subject position, the latter sentences signal a new perspective and appropriately move predicate material to the subject position; in both instances, syntax instantiates structure and rhetoric.

Furthermore, because this pattern of a series united by a common term can be instantiated through any quantity of sentences, it not only defines complete paragraphs, but in its smaller versions defines subunits of more complex,
double and triple term paragraphs, as seen in the following paragraph with its subunit on "cheap cases:"

It is commonly believed that a reader's interest is attracted by a case with which he can identify himself--there but for the grace of God, et cetera. But if the average tabloid reader were murdered, his misfortune would not receive much coverage in the average tabloid. He would be a "cheap case." Essentially, a cheap case involves what tabloid editors consider to be cheap people. This includes all working-class people, such as factory hands, waitresses and the unemployed. It also includes farmers, usually brushed aside as "hillbilly stuff." Alcoholics, whose antics are sometimes extremely entertaining, come under the same ban. So do Negroes, Mexicans, Puerto Ricans and other "lesser breeds without law." This causes some difficulty for the wire services since the current fashion is to delete any references to a criminal's race as "irrelevant." Thus an editor who might begin by showing great interest in a murder would cut the story down to a few paragraphs after learning that it involved a "Jig," but he would not publicly divulge the dread word that motivated his editing--and, of course, his editorial columns would continue to clamor for civil rights. 23

As the next chapter will develop, the series united by the common term seems to one of the most basic structural parts of paragraphs, and it presents a means by which to begin analyzing the structural hierarchy of a text. Indeed, this pattern exemplifies one of the characteristics of Barthes' satellites and kernels, where the satellites attach themselves to the kernels but do not themselves form a sequence. While in single term paragraphs, sentence modifiers were the satellites, in more complex paragraphs, the basic
kernel pattern of a series united by a common term becomes the satellite, as it did in the preceding example. As Culler describes it: "... kernel and satellite are purely relational terms: what is a kernel at one level of plot structure will become a satellite at another, and a sequence of kernels may itself be taken up by a thematic unit."^24

Another issue to be addressed in the next chapter is the topic sentence. In the present chapter, a whole family of happy, well-adjusted paragraphs did not share a single topic sentence between them. They all had topics, the repeated term, but none had a topic sentence. The euphemizer paragraph's opening line, "Admen and packagers are not the only euphemizers," was the only possible candidate. In contrast to the other opening sentences, this one contains some argument conjunction: "not the only euphemizers." However, because this conjunction involves adverb, adjective, and noun, it does not constitute a topic sentence of noun plus predicate. One can substitute "Parents and teachers are not the only euphemizers." and still keep the original sentences and the paragraph's cohesion. What the euphemizer paragraph does have is a more refined topic than the other paragraphs.

That paragraphs exist without topic sentences has already been documented. This chapter basically extends that observation and suggests that the appropriateness or inappropriateness of a topic sentence depends on the paragraph
totality or structure. If one assumes that a topic sentence necessitates the development and support of a statement—or an argument made up of at least a noun plus a predicate—the textual manifestations of that development would necessarily involve at least two recurrence chains, and we would thus be dealing with a different kind of structure from the single-term totality.

If one carries the issue of the topic sentence one step further and asks: should paragraphs exist without topic sentences? the foregoing analysis in this chapter suggests quite simply: it depends. It depends on the kind of paragraph structure. By definition, a series united by a common term excludes the topic sentence; additionally, the samples examined here have all been comprehensible, even stylish and elegant paragraphs, without the aid of a topic sentence.

On the assumption that cohesion defines relations and that relations define structure, this chapter referred to and developed the concept of paragraph structure. What can be said about the nature of this structure? First, structure may well be too powerful a concept to apply to discourse, if by it one means structure in the sense of atomic structure where all elements are interdependent and any addition or change produces a new totality, a new chemical. On the other hand, if one interprets structure in the sense of pattern in which certain minimal requirements must be met, then the concept seems to apply and is analogous to the
"structure" of an alphabet letter, a circle, or a box, in which certain defining patterns must be established but for which the size is not consequential.

As just suggested, paragraph size, or more accurately paragraph length, and eventually paragraph closure, seem to be the limiting factors in any attempts to describe paragraph structure. There is obviously some truth to the claim, cited at the opening of this chapter, that people paragraph at whim. Besides that, paragraphs are obviously affected by physical and genre conventions and limitations. Newspaper paragraphs differ in length from textbook paragraphs, letter paragraphs from typed paragraphs and so on. However, the methodology outlined in this chapter, combined with the concept of kernels and satellites, seems to provide a means by which we can predict where an author will not paragraph. If one assumes paragraphs to have some structural basis, any paragraphing that obscures or contradicts that structure would be considered inadequate. In this respect, the preceding analysis and suggested methodology extend previous paragraph research. Becker and Christensen, the modern pioneers in this area, posited analyses based primarily on rhetorical relations: topic, comment, restriction; coordination and subordination. 25, 26

And while these relations describe paragraph characteristics, they also describe characteristics of good prose in general, from paragraphs to chapters to full texts. The
preceding analysis, since it is primarily grammatical, offers a means for describing the linguistic cues and their interactions that produce those characteristics within single paragraphs. At the same time, it also allows for finer discriminations between text units insofar as it captures the potential structural similarities between traditionally disparate forms as the descriptive, exemplificatory, expository, and definitional paragraphs just analyzed.
NOTES TO CHAPTER THREE


5 Arthur Stern, "When is a Paragraph?" College Composition and Communication, 27 (1976), 253.


8 Halliday and Hasan, pp. 10-11.


11 Halliday and Hasan, p. 297.

12 Christensen, p. 205.

14 Roland Barthes cited in Culler, p. 220.


16 The, p. 17.

17 Culler, p. 219.


24 Culler, p. 220.


26 Christensen in Corbett, pp. 200-16.
CHAPTER FOUR: MULTIPLE CHAIN PARAGRAPHS

While chapter three developed the paradigm of the series united by a common term as a self-sufficient unit, this chapter will explore the ways in which equivalence chains are conjoined or embedded within one another to form more complex paragraph structures.

The immediate ramifications for a cohesion study are several. First, the conjoinment of two units entails two instead of one equivalence chain, and the number of chains will increase with the number of units conjoined or embedded within one another. Thus, not only must the relations of items within each chain be accounted for, but also the relations between the chains themselves. Also, the presence of more than one equivalence chain necessitates an extension of the definition of dominance developed in chapter three, since the definition was predicated on only a single term chain. As will be shown shortly, the presence of multiple chains can and does create a topic sentence, a condition prohibited by the single series. Finally, in paragraphs with topic sentences, the relationships between items within a single chain will necessarily be more complex and varied than the simple identity and reference relationships of the single term paragraph. If one imagines the
topic sentence as the most general and the summarizing sentence of a paragraph, then individual items in chains will perforce reflect that general-to-specific relationship.

To illustrate the concept of conjoined chains, we can begin with a paragraph which represents its simplest manifestation:

John likes Columbus. He likes its low inflation rate and its very low crime rate. He enjoys its many foreign food stores. He likes the city's extensive park system. Finally, he likes the cultural and sporting events that the city provides.

The first equivalence chain, the "John" chain, is easily and immediately identified within the framework described in chapter three. "John," or a pronoun for "John," appears in each sentence and occupies the dominant noun phrase position. The second chain, "Columbus," is instantiated through "its" and "city" and also occurs in each sentence. A third possible chain would be the verb sequence, but it does not constitute the same kind of equivalence chain as a noun sequence, lacking antecedents and referents. Verbs obviously perform the actual conjoining of equivalence chains, and thus contribute to cohesion, but this chapter will focus on the noun equivalences.

In illustrating the basic paradigm for two conjoined chains, the preceding paragraph also illustrates a straightforward, noncontroversial topic sentence, the kind Braddock in his study labeled "simple:" a stated and complete T-unit
representing the main idea of the paragraph. In this instance, the topic sentence also establishes an extended definition of dominance. As documented in chapter three, the mere presence of a repeated term in each sentence will not insure unity or cohesion; neither will the presence of two repeated terms in each sentence:

John likes Columbus. He likes its low crime rate and its low inflation rate. Someone told John that Columbus is located in the geographic middle of Ohio. While visiting friends in Columbus, John saw a Clippers' game last summer. Woody Hayes bumped into John at the Columbus Civic Center.

To insure term dominance and hence unity and cohesion, 1) the two terms must appear in, or be inferrable from, each sentence, and 2) one term must appear consistently in the sentences' subject position and the other in the sentences' object position, i.e. the dominant noun phrase constituent for the predicate. (In a deep structure analysis, "its" dominates "low crime rate" and "low inflation rate." ) These conditions merely extend the single term definition of dominance by incorporating the second and syntactically equal noun phrase in an SVO construction when that noun phrase participates in an equivalence chain. Schematically, the pattern might be represented as:

\[ A \text{ verb } B \]
\[ A^2 \text{ verb } B^2 \]
\[ A^3 \text{ verb } B^3 \]
\[ A^n \text{ verb } B^n \]
The schema graphically depicts the fact that the sample paragraph is a pure double chain paragraph. That is, the first or "John" chain fulfills all the requirements for a single series paragraph. And although occurring in the object position, the second or "Columbus" chain also appears in each sentence and dominates the subsequent predicate material and thus fulfills the same criteria except from the object position; by itself, the "Columbus" chain would form a single series paragraph.

An important ramification of this extended definition of dominance is that it reveals more evidence for cohesion as a condition attained by the merger of syntactic and semantic elements rather than as an exclusively semantic condition. If cohesion were a purely semantic manifestation, then the second version of the "John likes Columbus" paragraph would have been cohesive simply because each sentence contains the two repeated items. Thus any study that presents cohesion as merely equal to or definable by recurrences will be counting recurrences but not describing cohesion.

Additionally because the definition of dominance incorporates syntax and syntactical relationships, the first sentence of the original version can be seen as dominating the support sentences syntactically, as well as semantically through the set relations between "Columbus" and "characteristics of Columbus." This condition offers a syntactic explanation for why the reader interprets "John likes
Columbus," rather than one of the support sentences, as the paragraph's topic sentence.

Rarely, however, do authors construct such over-cued texts; usually some form of discourse ellipsis is applied to one of the chains, so that the example paragraph might appear as:

John likes Columbus. One reason is its low crime rate and very low inflation rate. Another is its many foreign food stores. It also has an extensive city park system. Finally, it offers many cultural and sporting events.

An even more elliptic version might be:

John likes Columbus. It has a low crime rate and a very low inflation rate. It has many foreign food stores. It also has an extensive city park system. Finally, it provides many cultural and sporting events.

In the first version, the transition words, "one reason," "another," "also," and "finally," operate again much like the substitute "one" at the sentence level and serve as placeholders for "John likes Columbus." so that the equivalence chain is maintained. Comprehension is not impeded since each sentence is prefaced by a word or phrase that can be interpreted through back anaphora to "John likes Columbus." However, in the second version, reader comprehension depends less on the text itself than it does on the reader's inferencing abilities and general world knowledge. With the "John" chain essentially erased, the author is depending on the reader to interpret each support sentence
with an inferential preface of "One reason John likes Columbus." The content necessary for the inference is not supplied by the text but by the reader's world knowledge—knowledge which the author is assuming as shared between the two of them. Of the author estimates inaccurately the kind or amount of shared world knowledge and builds a text around this misconception, readers will not find such a passage satisfactorily cohesive:

John likes Columbus. One reason is its low crime rate and very low inflation rate. Another is that it's located in the geographic center of Ohio. Finally, it has a population of 500,000.

As this passage illustrates, besides fulfilling the criteria for dominance just described, as this version does, a double series text must meet another criterion in order to be completely cohesive: it must support an inference chain that links, in this example, all the cited characteristics of Columbus under the general category of "likeable characteristics." More broadly formed, information predicated to the second equivalence chain must either directly or inferentially satisfy the reader's conception of the argument created by the conjoined chains, "John likes Columbus," or in other words, the topic sentence. It is at this final and inevitable inference chain inherent in any text that language operates truly and ultimately as an interface: a path of connection between discrete and unique individuals. At this juncture the author gives the text over to the reader.
As suggested earlier, "John likes Columbus," through its relations with the paragraph's subsequent sentences, represents what Braddock considers a simple topic sentence. Although he does not develop subcategories in this class, a second kind of simple topic sentence does exist: the set-relation statement illustrated by "Eldon Miller is a poor basketball coach" versus the argument statement illustrated by "John likes Columbus." While functionally identical in the sense that both can summarize a paragraph's theme, the set-relation topic sentence exacts different cohesion and inference requirements. Illustrated below is a sample set-relation paragraph, followed by its corresponding argument manifestation:

Eldon Miller is a poor basketball coach. One reason is that he is consistently out-coached in important games. Another is that he doesn't use his players to their full potentials. Also, he can't teach offense. Finally, he relies strictly on man-to-man defense.

John likes Columbus. One reason is its low crime rate and very low inflation rate. Another is its many foreign food stores. It also has an extensive city park system. Finally, it offers many cultural and sporting events.

In both versions, two equivalence chains are established; in each the transition words, "one reason," "another," "also," and "finally," form the first chain, but in the Miller paragraph the second chain, "Eldon Miller," comes from the subject position, whereas in the John paragraph, it comes from the object position, "Columbus." The predicate
material in the Miller paragraph, "poor basketball coach," "players," "offense," and "man-to-man defense," does not, in fact, form a straightforward equivalence chain but rather an inference chain. As a result, the paragraph is not a pure double series paragraph; the inferentially related rather than equivalent nouns in the predicate could not by themselves form a single series paragraph.

The role of the inference chain manifests itself most clearly in the following, skeletal version of the Miller paragraph; as the sample illustrates, the set-relation totality is defined by an equivalence chain and an inference chain, in contrast to the argument totality defined by two equivalence chains:

Eldon Miller is a poor basketball coach. He is consistently out-coached in important games. He doesn't use his players to their full potentials. He can't teach offense. Finally, he relies strictly on man-to-man defense.

Without any conscious awareness of the process, a reader who accepts the Eldon Miller paragraph as cohesive brings a considerable amount of knowledge to bear on the passage and transforms that knowledge into proposition relating inferences. No doubt many potential inference sequences exist, such as "Eldon Miller is a poor basketball coach. Basketball is a game. Eldon Miller is consistently outcoached in important games." It would be revealing to test different readers to determine what inferences they
actually employ, but without such evidence, one might posit "poor basketball coach" as the core of the most probable inference chain:

Eldon Miller is a poor basketball coach. A poor basketball coach is consistently outcoached in important games. Eldon Miller is consistently outcoached in important games. A poor basketball coach doesn't use his players to their full potentials. Eldon Miller doesn't use his players to their full potentials. And so on.

It is this multiple step process of drawing on world knowledge, shaping it into an inference, and accepting that inference as true, that finally gives cohesion to the passage. If the necessary inference is not true in terms of a reader's world knowledge, the passage will not be completely cohesive:

Eldon Miller is a poor basketball coach. He is consistently outcoached in important games. He doesn't use his players to their full potentials. He likes ice cream. He is a Mormon.

Basketball fans, and, one hopes, readers in general, will not accept this as a cohesive passage, foremost because the last two necessary inferences are not true; it is not true that liking ice cream or being a Mormon makes a poor basketball coach. The willingness of the reader to accept the necessary inferences as true, or at least as plausible, represents the final stage of a cohesive text. Although fundamentally a question of pragmatics, a reader's decision can be affected by one additional feature of the text itself. What
textually seems to predict the acceptability of inferences is whether or not a second inference chain can be formed from the reader's dictionary knowledge of terms in the predicate. In this instance one's knowing the dictionary definition of coach:

A person who trains athletes or athletic teams.

and of basketball:

A game played between two teams of five players each, the object being to throw the ball through an elevated basket on the opponent's side of the rectangular court.

enables the reader to set up a second inference chain connecting nouns in the sentence predicates; for example, "Basketball is a game," linking sentence one with sentence two: "Basketball is played with five members per team," linking sentence one with sentence three: "A coach trains players," again linking sentence one with sentence three. And so on. In contrast, noncohesive statements fail to yield acceptable inferences. Neither dictionary knowledge of ice cream nor of Mormons provides the reader with a term shared by "Mormon" or "ice cream" and "basketball coach." That individual reader's inferences may vary seems less important than the fact that the text itself helps shape these connections, that they are not exclusively governed by the reader and the world of pragmatics.

That this second inference chain can be established through dictionary knowledge of "basketball coach" helps
explain why the reader interprets the first sentence, "Eldon Miller is a poor basketball coach," as the topic sentence. Although "basketball coach" is not repeated in each sentence, the phrase's definition provides the implied premises that connect it with other predicate nominatives. This kind of implied repetition and the explicit repetition of "Eldon Miller," establish both as the two semantically dominant terms in the paragraph, while their occurrence in either the subject or the predicate nominative position establishes their syntactic dominance. This congruence between syntactic and semantic importance shown by both terms fulfills the criteria for cohesion and operationally defines the conjoined terms as the paragraph's topic sentence.

As both the Miller and the John paragraphs have illustrated, inference chains represent the last and most critical links between author, text, and reader. In those samples, at least two kinds of inferential relationships are involved. There are, first, the presumed logical relationships between entire sentences, as in the relationship between "Eldon Miller is a poor basketball coach" and "He is consistently outcoached in important games." Cohesion requires that the second sentence be interpreted as a "reason." Second, there is the truth value of the first set of relationships, that is, whether or not being outcoached in important games characterizes poor coaching in general and Eldon Miller's coaching in particular. Since evoking these appropriate
inferences establishes both a text's general communicative success and its cohesion, authors regularly incorporate textual constraints into a paragraph's basic structure. Reproduced below is a paragraph illustrating the incorporation of those textual constraints:

The reasons our opossum has survived in definitely hostile surroundings for 70 million years are evident. One is his small size: small animals always find hiding places, they always find a little food, where the big ones starve. Another of its assets was its astounding fecundity: if local catastrophes left only a few survivors, it did not take long to re-establish a thriving population. Also the individual opossum is not exactly delicate: it can stand severe punishment--during which it "plays 'possum" and then scampers away--and it can go without food for a considerable time. Finally, a great many different things are "food" to an opossum. Each of these traits have a high survival value, and their combination has presented the United States with a survivor from the Age of Reptiles.

In terms of kind and number of equivalence change, this is a set relation paragraph; its basic structure consists of one substantive equivalence change, "opossum," and one inference change, "survival traits." The second equivalence change, "reason," alluded to in earlier pages as a simple equivalence change, is actually an interpretative rather than substantive equivalence change, composed of logical connectors rather than purely lexical items. Syntactically dominant through its appearances in either a sentence's subject position or its sentence modifier
position, this interpretative change dominates the total unit comprised by the "opossum" chain conjoined with the "survival trait" chain. It instructs the reader explicitly how to interpret the conjoined chains.

To emphasize the veracity of this set of relations, the author buttresses the implicit, dictionary-based inferences connecting the trait to survival with explicit enthymemes: "One is his small size. Small animals can always find a little food. . . ." and so on.

As a relevant sidenote, one of these support segments manifests itself as a version of the double-series paradigm:

Also the individual opossum is not exactly delicate: it can stand severe punishment--during which it "plays possum" and then scampers away-- and it can go without food for a considerable time.

In the middle of the paragraph, the author inserts a mini-topic sentence. (Since the sequence is headed by the dominating "also," which links its sentence back to the paragraph's topic sentence, the reader is blocked from interpreting the sentence as a topic sentence.) Just as the single series paradigm can be a unit within more complex paragraphs, so, too, can the double series paradigm. It is thus another instance of Barthes' concept of kernels and satellites which assumes certain defining structures, but also assumes that the size of those structures (number of sentences in this application) affects whether or not
they are perceived as major or minor text units.

Schematically, the opossum paragraph might be represented as below, with the proviso that the basic structure can be expanded through embeddings like those discussed above:

\[ A \quad (B \ C) \]
\[ A_2 \quad (B_2 \ C_2) \]
\[ A_n \quad (B_n \ C_n) \]

The degree to which this description and perspective of cohesion and paragraph structure is useful, as well as valid, can be seen when the opossum paragraph is contrasted with a student attempt that follows the same kind of classic topic sentence:

"Mrs. Smith was my worst teacher in high school because she was a computer-like person. Myself, being a very studious person, though not a genius, like most other people find themselves frustrated when they cannot find the answer to a study problem after a reasonable length of time. I really appreciate a teacher who will take the time and discuss a student's problems with them on a person to person basis. The student is able to relax and probably understand his problem better. My teacher, Mrs. Smith, was all the time spitting out facts and theorems about this and that. Any student who could not understand these facts was bound to experience trouble. I feel that she could have been a much better teacher if only she were a little more personal with her statements. A more relaxed atmosphere in which to study could possibly have resulted."
Instead of appearing as an upsetting jumble, the student paragraph—if viewed through the set relation paradigm—reveals a kind of off target logic. Although laden with local recurrences, the paragraph does not establish a primary equivalence chain. More interesting, I think, is that the paradigm offers an explanation as to why the necessary chain was not instantiated. In trying to anticipate and bridge the differences between his world knowledge (what he expects teachers to be like) and the reader's on that same topic, the student entangles himself in explicating the inference chain that develops why Mrs. Smith's characteristics make her a poor teacher even before clearly establishing what are her actual characteristics. That is, the author tries to develop the inference chain before establishing the equivalence chain.

The opossum paragraph illustrates the explanatory power of the set relation paradigm in one way. In fact the opossum paragraph is an imperfect, but probably allowable, instantiation of the paradigm. In the base sentence introducing the three relevant chains, the main verb phrase operates merely as a grammatical placeholder: "The reasons our opossum has survived in definitely hostile surroundings for 70 million years are evident." No chains originate from the verb phrase which is, after the subject, a sentence's syntactically most important position. On the other hand, two of the three chains, the opossum and the
survival traits, are embedded in the subject phrase, and thus originate from syntactically subordinate positions. This unequal distribution of the chains prevents the topic sentence from attaining a complete dominance, lacking, as it does, a chain originating from the main verb phrase. The paragraph's actual topic sentence may be contrasted with the following, which not only fulfills the criteria for dominance, but in so doing more accurately reflects the paragraph's thesis: "Our opossum has many survival traits," because of its structurally co-equal introduction of the primary chains.

In some cases when the conjoined primary chains are embedded in a syntactic structure such as the subject, the sentence containing that conjoinment may lose its status as the topic sentence:

The arbitrariness of the structuring of things and thoughts is familiar to anyone who has ever learned a foreign language. That continuum of reality which English separates into two units through the labels arm and hand is treated as a single unit by the Russian label ruka. That area of action which English divides into two units through the labels carry and bear is handled as a single unit by the French porter. That area of thought which English treats as a single unit through the label remember is divided in German into two units through the one-word label behalten ('remember' = 'keep in mind') and the three-word label sich errinnern an (‘remember’ = 'recall to mind')?

If a reader views the first sentence as a topic sentence, it is clearly of a different nature from "John likes
Columbus;" it does not operate as overarching sentence for
subsequent sentences; it operates rather as a base sentence
from which the others rise. Indeed, one can substantially
change the first sentence: "The arbitrariness of the
structuring of things and thoughts is the basis for Whorf's
theory of language." The reason one can change that
sentence is that no chains track back to the verb phrase,
so that the main idea of the paragraph is captured in the
noun phrase only, and not in the entire sentence.

Whether the conjoined chains are instantiated as an
entire paragraph or as a paragraph subunit, the grammatical
form is the simple SVO sentence/clause. Psycholinguists
have accumulated impressive evidence demonstrating that the
clause is the primary perceptual unit because it is the
minimal unit with semantic determinancy. That is: "The
constituent words of a clause become perceptually deter-
minate only when they are perceived as functional elements
within the clause, and these definite functions are fully
determined only after the whole clause is perceived."^4

This chapter's analysis of conjoined chains suggests
that the sentence/clause SVO creates a similar determinancy
at the discourse or paragraph level: sentences following the
SVO sentence/clause become maximally determinate only when
perceived as being dominated by the SVO sentence/clause and
fulfilling the criteria for dominance. The conjoined
chains in the SVO form are the textual equivalent to the
psycholinguists' frame or the rhetoricians' topic sentence, or the structuralists' totality. It is the base from which everything else is contextualized and given meaning. It not only contains the head words for the chains, but also presents them as a unit; it is the whole into which the parts are fit and understood when all chains lead away from the conjoined unit and all lead back to it. In this respect, the double series paradigm as defined through dominance represents a convergence of theories from structuralism, psycholinguistics, and linguistics. Sentences or sentence constituents which are not incorporated into the recurrence chain disrupt the cohesion of a paragraph; conversely, the degree to which sentences and sentence constituents are incorporated into the recurrence chain reflect the paragraph's cohesion and also its structural pattern. As Culler wrote on Christmas trees: "Some Christmas trees are more successful than others, and we are inclined to think that symmetry and harmonious arrangement of ornaments makes some contribution towards success."\(^5\)

The concept of dominance can also be grounded in information theory. It has long been noted that in English there is a general tendency to place shared information—information presumed common to both author and reader—in the sentence initial position, and to place new information in the end position. The conjoined series primary SVO sentence/clause seems to perform this function at the
paragraph level, providing in one conjoined unit all the initiating words in the two recurrence chains.

Information theory and also inference overloading seem necessary to explain the parameters affecting the relations of items within a single recurrence chain. Thus far all the recurrence chains we have examined have been based either on simple coreferentiality instantiated through pronouns or on the set-relations of HAS:A or IS:A. Basically neither of these instantiations add new information. However, consider the following hierarchy of more-to-less acceptable instantiations of "Eldon Miller" in a recurrence chain:

Eldon Miller is a poor basketball coach.
He is always outcoached in important games.
The jerk is always outcoached.
The man is always outcoached.
The Wittenberg alumnus is always outcoached.
The husband is always outcoached.
The Rotarian is always outcoached.

Since there is only one available referent, Eldon Miller, for "jerk," "man," "Wittenberg alumnus," "husband," and "Rotarian," any unacceptableness is not caused by competing or confusing antecedents. The increasing threat to cohesion in the later items seems to result from the amount of new information they introduce; whereas "jerk" is all but implicit in "poor basketball coach," and "man" is implicit in the name "Eldon Miller," terms like "Wittenberg alumnus," "husband," and "Rotarian" must be connected with "Eldon Miller" by inferences based on the reader's world knowledge.
As the reader is required to make these connections, the recurrence chain loses the stability derived from unequivocally shared information; the equivalence chain gets overloaded by the inferences necessary to accommodate new information.

Chapter three examined the pure single chain paragraph, and thus far this chapter has examined the pure double series paragraph and the conjoined equivalence and inference chain paragraph. These paragraph types are very limited and limiting, consisting of fairly simple chains; neither the chains themselves nor their various combinations accommodate much information. Indeed, probably because of this limited information capacity, most paragraphs combine these structural forms and consist of one primary chain and two or three secondary chains that appear in many but not all of the paragraph's sentences. This next section reviews paragraphs of this mixed type.

For cohesion, the mixed form manifests a classic communication problem. In contrast to the earlier types, the mixed form is more informationally rich and full, characterized by a greater sense of thematic progression, without, however, sacrificing its integrity as a unit; it produces a strong sense of back anaphora, or vertical structuring, as well as an advancing theme, or linear and horizontal structuring. The dynamic juxtaposing and balancing between two potentially conflicting conditions characterizes this
paragraph type. It is also this dynamic juxtaposing that has impeded efforts to describe paragraph cohesion and structure. In a text characterized by thematic progression, the inherent control imposed by a repeated subject no longer exists, since the subjects themselves are continually being replaced:

John likes cheese. Cheese is made in Wisconsin. Wisconsin is the home of my best friend.

Although patently lacking cohesion, the above example fulfills Danes' discourse category of "simple linear progression, in which the rheme of one sentence becomes the theme of the next."^6

While Danes' pattern may describe recurrences, it obviously does not describe cohesion; identifying some of the other necessary textual requirements for this linear pattern in the goal of the following section.

The following is an example of a mixed structure paragraph that employs linear progression:

Imagine, if you will, a minuscule "chip" shaved from a crystal of silicon. On this chip are all the components of an entire information storage and programming system—a full-fledged computer, in other words, that takes up less space than the first four letters of this paragraph. That is the microcomputer of tomorrow and its prototype already exists. The information it stores could come from anywhere—The Library of Congress, The New York Times, the personal banking records of thousands of taxpaying citizens. It could come from the tape cassettes of psychiatrists, the daydreams of novelists, the logbooks of birdwatchers.
Soon information of this sort will be available to anyone and everyone at the flick of a switch. You'll be able to plug into it. So will the government. So will the guy next door. Your kid will be able to plug into that information—your doctor, your thesis advisory, your guru, your garbage man. And when everyone is all plugged in and accessing, in unison, this monumental new universe of data, the existential situation is going to change.

In order to identify the underlying patterns of recurrences in this paragraph, the following schema lists the subjects and objects of the paragraph's sentences; such an analysis is consistent with the preceding discussion which designated those two positions as the informationally most important ones in a sentence:

<table>
<thead>
<tr>
<th>SUBJECT</th>
<th>OBJECT OR PREDICATE NOMINATIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>(You) implied</td>
<td>chip</td>
</tr>
<tr>
<td>Chip</td>
<td>components for an entire information storage and programming system</td>
</tr>
<tr>
<td>That</td>
<td>microcomputer</td>
</tr>
<tr>
<td>Its prototype exists</td>
<td>anywhere</td>
</tr>
<tr>
<td>Information</td>
<td>tape, cassettes, daydreams, etc.</td>
</tr>
<tr>
<td>It (information)</td>
<td>anyone</td>
</tr>
<tr>
<td>Information</td>
<td>it (information)</td>
</tr>
<tr>
<td>You</td>
<td>(information)</td>
</tr>
<tr>
<td>Government</td>
<td>(information)</td>
</tr>
<tr>
<td>Guy</td>
<td>information</td>
</tr>
<tr>
<td>Kid</td>
<td></td>
</tr>
<tr>
<td>Situation going to change</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The simple list reveals a perhaps startling fact: the basic, dominating recurrence chain, based not only on repetition, but also on its occurring in the most syntactically important positions of subject or object is "information."
While "information" does not occur in each sentence, it does occur in eight of the twelve sentences (including in the total the compound sentence as two) and it consistently appears in either the subject or the object position. It is thus the only candidate for the primary recurrence chain, fulfilling the criteria for dominance with its congruence between semantic and syntactic importance. The establishment of this primary recurrence chain is the first requirement for the successful mixed structure paragraph, just as it is the basic requirement for the single series paragraph.

The second requirement of the mixed structure paragraph involves the secondary recurrence chains--their syntactic positions and the means by which they are introduced. These secondary chains and their relationships to the primary chain determine whether or not a paragraph is cohesive.

Any sequence opening with an item other than "information," or in other words, any opening item other than that of the primary recurrence chain, qualifies for membership in a secondary chain. The most obvious secondary recurrence chain, because of its length, is the "people" chain:

You'll be able to plus into it. So will the government. So will the guy next door. Your kid will be able to plug into that information--your creditor, your thesis advisor, your guru, your garbage man.

This four-sentence sequence fulfills the basic conditions of a secondary chain, that is, a chain sufficiently different
from the primary chain to introduce new information, yet related to the primary chain in such a way that the new information, as it in turn is incorporated in its own recurrence chain, can be interpreted as an extension of the primary recurrence chain rather than as a completely new and different chain. To this end, the secondary recurrence chain should be introduced into the text by the primary chain, as in: "Soon information of this sort will be available to anyone and everyone at the flick of a switch." Second, when an instantiation of the secondary recurrence chain appears in the subject position, an instantiation of the primary recurrence chain should appear in the predicate position, as in: "You'll be able to plug into it (information)." Schematically, the pattern is as follows:

```
A verb B
B verb A
```

Both of these conditions on the introduction of secondary chains depend on and reflect the necessity of back anaphora in preventing a noncohesive, linear progression sequence like the Wisconsin cheese sentences. From an information theory viewpoint, these conditions reveal how new, unknown information—the secondary recurrence chain—is introduced and developed through old, known information without its signaling a completely new topic or paragraph structure.

In the "chip" paragraph, another candidate for a secondary recurrence chain is found in the following sequence:
On this chip are the components for an entire information storage and programming system—a full-fledged computer, in other words, that takes up less space than the first four letters of this paragraph. That is the microcomputer of tomorrow and its prototype already exists.

The recurrence pattern for this sequence can be represented as follows:

A verb B
B verb C

To the extent that the new information about the computer is introduced via old information, the sequence fulfills one of previously outlined conditions for new information introduction. But because the new information, "microcomputer of tomorrow" and "prototype already exists," is not then incorporated into other chains, and has only one recurrence link, the sentence is of marginal importance informationally and structurally. A sentence such as this with only the single recurrence link can easily, without disrupting either cohesion or sense, be pruned. It is an allowable informational aside within the structural and cohesion patterns, but not critical to either.

There are two final issues to be addressed with regard to secondary chains and their relationships to the primary chain: the status of the secondary recurrence of "chip" and the actual initiation of "information" into the text. Both of these events occur in the paragraph's first two sentences:

Imagine, if you will, a minuscule "chip" shaved from a crystal of silicon. On
this chip are all the components for an entire information storage and programming system—a full-fledged computer, in other words, that takes up less space than the first four letters of this paragraph.

The pattern of the "chip" recurrence chain duplicates both the computer pattern just examined (A verb B; B verb C) and also one of the opening patterns for the single series paradigm discussed in chapter three. Informationally, the second sentence introduces both "computer" and "information" to the text, besides other items not chained at all ("programming system," "components," etc.) The sentence is informationally very busy and rhetorically, structurally full of authorial opportunities and choices. Besides setting up the possible development of "chip" in the form of a single series, it also sets up the possibility of a "computer" chain, as well as several other potential chains just noted. The actualized chain is, of course, the "information" chain and it emerges from the position of adjective, a minor sentence constituent in any surface structure analysis, rather than from the major object position.

However, in a deep structure analysis, adjectives are viewed as NP's, so that the originating structure for "information storage system" is the basic SVO clause, "system stores information," in which "information" occupies the object position in one of the deep structure sentences. Given the various deep structure constituent sentences of
the actual sentence, "On this chip are all the components of an entire information storage and programming system—a full-fledged computer, in other words, that takes up less space than the first four letters of this paragraph," the author, among several other choices, could have written this sequence: "On the chip are components. The components are for a system. The system stores information. The system programs information." Thus, although in the surface structure, "information," is subordinated to the adjective position, by deep structure definition, it has the potential to be actualized as an object. And it is precisely because "information" can be actualized as an object that its eventual role as a primary recurrence chain is predictable.

Finally, among texts which synthesize information by combining SVO clauses into complex sentences, those that maintain the primacy of this unit will be more cohesive than those which do not. That is, texts whose recurrence chain items can be derived from a deep structure SVO sentence/clause and which, furthermore, preserve and maintain that SVO derivational history in its surface manifestations will be more cohesive than those whose recurrences either lack such a derivation or have it obscured by inconsistent surface manifestations.

One can contrast the preceding cohesive use of primary and secondary chains with the following weakly cohesive
student text that violates the rules just outlined:

The book *Who's Who in Engineering* is an index of engineers and engineering societies in the United States. The American Association of Engineering Societies (AAES) compiled the data used in the book. The book starts with the criteria for being placed in the index. The names and positions of the governing body of the AAES are presented and the member societies, associate societies, and regional societies are listed. The different engineering societies and organizations and engineering awards are indexed separately. The major engineering societies are listed with the officers' names, the award they give, and the past recipients of the awards. The engineers are listed in alphabetical order with a small profile on each. At the end of the book engineers are indexed geographically and by area of specialization.

In skeletal, SVO form, the paragraph would be as follows:

**SUBJECTS**  **VERBS**  **OBJECTS**

book  is  index
AAES  compiled  data
book  starts with  criteria
names and  are presented
positions
member societies,  are listed
associate societies,
regional societies
different engineering  are indexed
societies and organiza-
tions and engineering
awards
engineering societies  are listed
engineers  are listed
engineers  are indexed

Although "book" (or "index") would seem an obvious choice for the primary recurrence chain, it never gains textual instantiation because of the many other competing items in
the subject positions: "AAES," "names and positions," "engineering societies," and "engineers." Besides this competition, "book" also loses thematic and cohesive status by its inconsistent placement in sentence two. After first being introduced as a subject in sentence one, "book" is inconsistently demoted to a non-SVO prepositional phrase in sentence two, "in the book."

With such a weak and doubtful primary chain, the burden for cohesion would seem to fall to the secondary chains. But here again the recurrences are basically helter-skelter: sentence subjects, and thus potential recurrence chains, are not introduced by preceding sentences and thus lack that immediate back anaphora inherent in the A:B; B:A pattern. And a very likely prospect for a secondary chain, "AAES," introduced in the subject position of the second sentence, is simply banished from the text entirely. Like the demotion of "book," the disappearance of "AAES" reflects the text's failure to respect and maintain the integrity of an SVO component and thus its cohesive potential. The following passage illustrates the latent cohesive power of "AAES" when it fulfills its potential as a secondary recurrence chain:

The names and positions of the governing body of the AAES are presented and its member societies, associate societies, and regional societies are listed. Non-AAES engineering societies and organization and their awards are listed separately.
While the sample passage invites more discussion (and more editing), I will make only one final observation. Although not as inherently evil as most composition texts suggest, the passive sentence construction, which this passage used extensively in an elliptic form, is probably the primary cause of the overall weak cohesion. In an elliptic form with the presumed agents deleted, the passive constructions effectively stripped deep structure SVO units of their subjects and produced the following sentences:

Names and positions . . . are presented and member societies . . . are listed.
The different engineering societies . . . are indexed separately.

Lacking surface objects, the sentences distort their deep structure SVO patterns and prevent the instantiation of the A:B; B:A pattern simply by deleting both the potential recurrence chain item itself and its syntactic position. Once again, not only is the importance of syntax to cohesion reconfirmed, but also the importance of surface syntax.

In summary, the last few pages have examined the interactions between recurrence chains. As the texts themselves have become more complex, so, too, has the definition of dominance, but in a logical, not ad hoc, manner. The criterion for dominance is always congruence between semantic importance and syntactic importance. In the single series, dominance is attained by the chain's appearing in the subject position; in the double series,
by the two chains' appearing repeatedly in the subject and object positions. Finally, in the mixed paradigm, first primary chain must regularly appear in either the subject or object position; secondary chains must be introduced through a sentence containing the primary chain and finally, item recurrences must maintain and reflect a derivational history from an SVO sentence/clause.

The "information" paragraph previously examined is an example of the mixed sequence paragraph without a topic sentence; our sense that it lacks a topic sentence is generated by the means through which the secondary chains were introduced. Their introduction followed the pattern of A:B; B:A, of "chip" to "information" to "people." As a result, because all the chains were introduced through discrete and separate sentences, there was no single "topic sentence." Our sense of a topic sentence is fulfilled, however, when all secondary chains are introduced through reduced SVO clauses that are combined with the primary recurrence chain into a single unit, a single sentence. We can see that in the following paragraph:

In regard to the use of space, it is possible to observe a basic and sometimes inexplicable dichotomy in the animal world. Some species huddle together and require contact with each other. Others completely avoid touching. No apparent logic governs the category into which a species falls. Contact creatures include the walrus, the hippopotamus, the pig, the brown bat, the parakeet, and the hedgehog
among many other species. The horse, the dog, the cat, the rat, the muskrat, the hawk, and the black-headed gull are non-contact species. Curiously enough, closely related animals may belong to different categories. The great Emperor penguin is a contact species. It conserves heat through contact with its fellows by huddling together in large groups and thus increases its adaptability to cold. Its range extends over many parts of Antarctica. The smaller Adelie penguin is a non-contact species. Thus is is somewhat less adaptable to cold than the Emperor, and its range is apparently more limited.

Instead of introducing secondary chains through the object position as the information paragraph did, this text presents them through reduced SVO forms which are then immediately combined into the first, complex sentence of the paragraph. Instead of: "The dichotomy is basic. The dichotomy is also inexplicable." these two basic SVO's are combined into "a basic and inexplicable dichotomy."

Yet just as the SVO clauses are reduced and synthesized into a single sentence, the remainder of the paragraph plucks out each reduced SVO from the first sentence, restores it to full SVO status, and proceeds to develop it individually. Thus from the perspective of reduced-to-restored SVO's, the core sentences of the paragraph are:

No apparent logic governs the category into which an animal falls. Curiously enough, closely related animals may belong to different categories.

As this list suggests, the shared recurrence is the chain of "dichotomy" from the first sentence and "category" from
these two core sentences which together form the single "dichotomy: category" chain. The chain's semantic importance, attained through this repetition, is re-enforced and congruent with its syntactic importance, attained through the chain's appearances in the sentences' object positions. This condition of limited dominance, in conjunction with the manner in which these expended SVO's operate in the text and the means by which they are developed, make the chain, "dichotomy: category," the primary recurrence chain in the paragraph.

First, each expanded SVO operates as a mini-topic sentence as illustrated by the sequence:

No apparent logic governs the category into which a species falls. Contact creatures include the walrus, the hippopotamus, the pig, the brown bat, the parakeet, and the hedgehog among many other species. The horse, the dog, the cat, the rat, the muskrat, the hawk, and the black-headed gull are non-contact creatures.

In this and the other SVO sequence, "category" and its subject--either "no apparent logic" or "closely related animals"--contextually frame or dominate their succeeding sentences. In turn, these core sentences, under a reduced-to-restored analysis, are dominated by the first sentence of the paragraph. Finally, because of this hierarchy of dominance and because "category" occupies the object position of each core sentence and "dichotomy" the object position of the first sentence, the "category:dichotomy"
chain emerges as the primary recurrence chain and the first sentence as the topic sentence.

Also contributing to and supporting these two perceptions is the manner in which the mini-topic sentence segments are developed. Instead of relying solely on single term repetition, the chains in this paragraph move from general to less general instantiations: dichotomy to category; animal world to species to particular animals. With a text employing this kind of chain instantiation, our sense of topic sentence and primary recurrence chain is shaped by a semantic relation other than simple repetition; it is formed by the set relations of class/member, so that the "class" items are seen as encompassing the member instantiations or items. Here, once again, it is the paragraph's initial sentence that contains the most general class instantiations of the recurrence chains, which also leads it to be interpreted as the paragraph's topic sentence.

In conclusion, this chapter has examined how recurrence chains, first in the double-series paradigm, and then in the mixed, interact with one another to form more informationally rich and complex paragraphs. In doing that, it has implicitly described the paragraph structure for the analyzed texts. Developing and employing the definition of dominance as congruence between semantic and syntactic importance, the chapter outlined the parameters and conditions under which recurrence chains may be combined and conjoined. Once a
sentence sequence meets these criteria, a pattern has been formed and a structure delineated. In this respect paragraph juncture, based on structure, is possible and predictable whenever such a pattern exists. Likewise, the analysis accommodates the phenomenon of readers' paragraphing texts at different junctures; since the patterns are not exclusively dependent on length, they may be instantiated through three, four, five or ten sentences. (Presumably there is some limit but once the pattern or subpattern is established, genre would probably influence a paragraph's length.) Put another way, the analysis shows where people will not paragraph; where they will not see a new unit being initiated.

Although not empirically tested (something I hope sometime to accomplish), the two most stable patterns seem to be the series united by the common term and the mixed sequence with reduced SVO's. The series united by a common term seems to be the building block of many paragraphs as seen in texts like the information paragraph which conjoins these units to create more complex paragraphs. The degree to which that conjoined paradigm is acceptable as a unit deserves empirical testing. One might predict that those units are most susceptible to segmenting, paragraphing, into single series units. The strongest constraint on such segmenting would probably be the number of recurrence chains
involved, or the use of text-dependent phrases such as "this...", "others...", which must be interpreted through back anaphora.

Empirical investigation might show the mixed sequence based on the double series to be the next most stable unit because of its use of reduced-to-restored SVO's, which tie subsequent sentences to the initial sentence containing the reduced SVO's. This pattern, however, would presumably be affected as well by the length of development of the restored SVO's or the mini-topic sentences; if the restored SVO's in the animal paragraph, for example, had been developed through ten or twelve sentences, its mere length might have signaled a paragraph juncture—especially if it appeared in a popular article.

Besides describing paragraph structure, the analysis accounts for both the horizontal thematic progression and the vertical deepening inherent in discourse. The two specific means examined here were the conjoined subject/predicate, or A:B (where B introduces new information), and the embedded reduced-to-restored SVO's. On the one hand, each is a single unit signaled by its syntactic conjoinment, and thus constitutes the psychological frame (or the rhetoricians' base sentence) from which other sentences arise. On the other hand, each can be combined with another unit to provide the means by which new information is introduced to create thematic progression.
Finally, the implications of the preceding analysis must be examined from the perspective of the much beleagured topic sentence. Like chapter three when it examined the single series paradigm, this chapter described a paragraph type that was cohesive, elegant, and also lacking a topic sentence. The functional, structural, and informational necessity of a topic sentence once again depend primarily on the message's informational purpose. Even among paragraphs with topic sentences, there are distinctions in both form and development, so that the simple argument topic sentence, "John likes Columbus," and the simple set-relation topic sentence, "Eldon Miller is a poor basketball coach," have different information and inference requirements and are manifested by different kinds of recurrence and inference chains. Furthermore, as the analysis of complex paragraphs showed, topic sentences, operationally defined through the concept of dominance, can and do appear within paragraphs. The overreaching conclusion to be drawn from this analysis is that the topic sentence is merely one means by which information is conjoined and combined. It does provide the reader with a contextual frame that allows for the integration of subsequent information, but so too does a base sentence sequence (A:B;B:A). The fundamental appeal of topic sentences lies with their structural tidiness; that is, paragraphs with topic sentences are characterized by more controlled, tighter recurrence chains that are "umbrellaed"
under the topic sentence.

Conversely, paragraphs with a base sentence sequence rather than a topic sentence lack this umbrella effect; all chains do not emanate from or lead back to a single sentence. Although one might argue that a topic sentence is intrinsically richer and more complex than a base sentence, it may not be better from a message point of view. By definition, the topic sentence paragraph develops at least two recurrence chains simultaneously, and thus requires more attention and processing ability from the reader. In technical or scientific writing dealing with esoteric or complex topics, the base sentence structure may indeed be more appropriate because it is characterized by the sequential joining of single term sequences. What the topic sentence presents as a single unit, the base sentence sequence breaks up into more easily understood segments, which are developed individually before they are linked with other segments. Thus the two types of paragraphs serve different informational needs, and the structural elegance of the topic sentence paragraph does not make it intrinsically "superior" to the base sentence sequence paragraph.
NOTES TO CHAPTER FOUR


CHAPTER FIVE:
CONCLUSIONS: IMPLICATIONS AND APPLICATIONS

This chapter briefly summarizes the preceding chapter and then discusses the applications of their conclusions to other discourse analysis systems and to composition teaching.

The preceding chapters have reached the following major conclusions about cohesion in English paragraphs. Cohesion consists of both unity and coherence. By far the more important of these two components, unity is produced by a dominant term present in, or inferrable from, all sentences in a paragraph. A term attains dominance when there is congruence between its semantic importance, attained either by its frequency of recurrence or by its referential breadth, and the term's syntactic importance, produced by its regular appearance in one of the most important syntactic position(s) of each sentence in a paragraph. The degree to which these recurrences manifest a formal pattern, or totality, reflects a paragraph's degree of cohesion. These conclusions are consistent with psycholinguistic studies showing the necessity for a text to instantiate a frame, or informing context, before a reader can accept it as cohesive. In this respect, the totalities operate as formal,
shared, definable interfaces between the reader's and the author's world knowledge.

Of all possible totalities that texts might generate, the dissertation examined four: the single series united by a common term; the double series united by two common terms; the double series united by one recurrence and one inference chain; and the mixed sequence based on combinations of the single and double series. In describing these cohesion patterns, the dissertation simultaneously identified paragraph structures based on the same semantic and syntactic relationships that produce cohesion, thus adding evidence that paragraphs are indeed structures, and not merely elective or idiosyncratic discourse phenomena.

To the extent that the analysis relied on current linguistic theory in its sentence parsing, the process of establishing the syntactic position of each recurrence item is straightforward and reproducible. The limitation of the analysis is that it cannot establish and track the recurrences themselves in an equally straightforward and reproducible manner. Paragraphs with highly visible and highly referential recurrence chains were used to illustrate the basic paradigms; but language is considerably more rich and poetic than the one-to-one referentiality of nouns and pronouns. As indicated in chapter one, current semantic theories offer little or no help in describing the rich allusiveness of language.
Despite its limitations, the analysis of cohesion presented here produces, I believe, a more accurate description of discourse than the two most currently popular approaches: Kintsch's propositional analysis and Francis Christensen's generative rhetoric. To illustrate the differences, below are text analyses from both Kintsch and Christensen, with Kintsch's approach discussed first.

As noted in chapter two, Kintsch segments discourse into what he calls propositions, which for all intents and purposes are surface structure constituents:

- John sleeps. (sleep, John)
- The man is sick. (sick, man)
- The old man smiled and left the room. (old, man) & (smile, man) & (Leave, man, room)

These propositions are then ordered by the (A,B) (B,C) chronology so that proposition (A,B) is superordinate to proposition (B,C) because A,B) precedes (B,C) and shares with it the common argument B. Thus in a Kintschian approach, a sample sentence is segmented as follows:

A great black and yellow V-2 rocket forty-six feet long stood in a New Mexico desert.

1 (GREAT, ROCKET)
2 (BLACK, ROCKET)
3 (YELLOW, ROCKET)
4 (V-2, ROCKET)
5 (LONG, ROCKET)
6 (FORTY-SIX FEET, 5)
7 (STAND, ROCKET)
8 (IN, 7, DESERT)
9 NEW MEXICO, DESERT)
These propositions are then schematized as follows:  

\[
\begin{array}{cccccccc}
7 & 1 \\
2 \\
3 \\
4 \\
5 & 6 \\
8 & 9 
\end{array}
\]

As the analysis illustrates, Kintsch's method is based on a contradiction; it uses surface structure constituents and at the same time completely masks and ignores their capacity to manifest the relationships involved in cohesion; all relationships are homogenized into the simple propositional relationship. This failure, combined with his ranking of propositions through shared propositions, disables his analysis from dealing with double chains, such as "John likes Columbus. He likes its park system." With the defined capacity of only one linear relationship, his system can neither identify nor describe these double recurrence chains of "John" and "Columbus." Conversely, Kintsch's system would parse as acceptable the sequence: "John likes cheese, Cheese is made in Wisconsin, Wisconsin is a cold state." simply because there are the shared recurrences of "cheese" and "Wisconsin." Kintsch offers no parameters other than simple repetition by which to evaluate information introduction.

Christensen's method suffers from essentially the same problems. Christensen posits two possible sentential relationships: subordination and coordination. These are signaled by and manifested within a single recurrence chain
so that his method, like Kintsch's, cannot deal with double recurrence chains, their relationships, or their contribution to cohesion. Below is a sample paragraph analyzed with his method; sentences that are coordinate with one another are ranked at the same level; sentences that are subordinate are ranked at a new level:

1. The process of learning is essential to our lives.
2. All higher animals seek it deliberately.
3. They are inquisitive and they experiment.
4. An experiment is a sort of harmless trial run of some action which we shall have to make in the real world; and this, whether it is made in the laboratory by scientists or by fox-cubs outside their earth.
5. The scientist experiments and the cub plays; both are learning to correct their errors of judgment in a setting in which the errors are not fatal.
6. Perhaps this is what gives them both their air of happiness and freedom in these activities.

In Christensen's analysis, the above paragraph is based solely on subordination; it thus ignores the strong back anaphora created by the chains of "learning" and "higher animals" that actually unifies the text.

The analysis presented in the preceding chapters, because it is based on the interactive contributions of both syntax and semantics, lays the groundwork for identifying and describing patterned double and triple recurrence chains. In terms of the analysis in chapter four, the "learning" paragraph just cited exhibits the same A:B;B:A pattern seen in
the "information" paragraph—a common pattern that curbs the runaway "Wisconsin cheese" sequences. The analysis does not presuppose one single set of relationships and does not ignore or mask complex relationships.

Finally, the preceding analysis is sensitive to one of the most basic tenets of discourse; that grammatically correct individual sentences may not be grammatically correct in particular contexts, and, as a corollary to that fact, that sentences derive their semantic and syntactic configurations from the context in which they appear. In this respect the analysis begins to account for some of the most interesting facets of discourse: that the whole is always larger than the sum of its parts, that a text is always an interface between worlds of knowledge, and that successful texts take that condition into account.

I would like to think that the applications to composition teaching are obvious. The most immediate is that the analysis provides a formal and reproducible model for teachers and students. Like current sentence combining techniques, the paragraph patterns suggest definable and discrete models for packaging sentence sequences into unified texts at the discourse level. The theory assumes only a rudimentary grasp of grammar, and its application requires neither elaborate techniques of analysis nor the mastery of obscure terminology. The only abilities required are the recognizing of subjects and objects and the counting of
their instantiations. In these respects the theory meets two widely accepted criteria of effective pedagogy: that students respond positively to models, and that teachers should meet students where they are, should begin with their present knowledge and then enlarge that knowledge into a more conscious awareness of discourse structure.

Depending on the students' particular needs, the models can function either as generative devices, or as evaluative, editing devices. As generative devices they offer the student patterns for combining sentence sequences, yet patterns that are at the same time free of text. I explain. Students may respond to models only too well: the weakest and least confident student writers—the ones most relevant to the teacher—tend all too often to imitate models slavishly rather than to assimilate them conceptually. If the possibility is left open to do so, such writers will turn composition exercises into mechanical activities equivalent to typing or penmanship practice. But text-free models, even in their most slavish and mechanical imitations, require thought and input from the students, i.e. composition.

The analysis also presents a spectrum of continuous and cumulative structures, from simple to complex, that enables the teacher for the first time to present discourse structures in a logically graduated fashion, as in the teaching of mathematics and some of the sciences. In the
traditional teaching of composition, the only progression is the basically rhetorical progression from description to narration to analysis to argumentation. But a theme must achieve cohesion whether it is descriptive or argumentative, so that the patterns analyzed here impart skills that precede, logically and pedagogically, the specialized skills necessary to produce the rhetorical modes of discourse.

In this way the patterns speak to one of the most common failures in student writing: the mechanically "coherent" but woefully disunified paragraph that the student wrote in The Ohio Union that is on the Olentangy River that flows through Columbus that really carried her away—the paragraph that begins in Poughkeepsie and end in Albuquerque because of the student's microscopic, deconstructed concept of the paragraph as a totality. As Piaget has written, one aspect of maturing involves the widening of perception from a personal and egocentric world view to a social world view. And as composition teachers have noted, one of the biggest hurdles for beginning writers to overcome is the qualitative difference between cohesive oral--flowing river--discourse and cohesive written discourse. These patterns provide one means for immature writers to develop that broader and more structured view necessary to written texts.

The usefulness of the models as evaluative or editing devices has already been suggested by the analyses of the
student texts on the poor teacher and AAES index in chapter four. The models can provide teachers and students with a specific and relatively non-threatening vocabulary for criticism and instruction. Instead of the teacher's criticizing a student paragraph in generalized, subjective terms for lack of cohesion—a criticism which can be taken personally and defensively—the teacher using these models has a means of reasoning with the student and raising the level of discussion from personal criticism to the sharing of a common and objective paradigm. Instead of objecting in seemingly idiosyncratic terms, the teacher can say, in effect, "Here's the reason, and it has nothing to do with you personally; on these specific grounds, you can see for yourself."

By the same token, the models begin to provide a means to forestall the deadening student rejoinder to teacher criticism, "Everyone gets it but you"—referring, of course to inferential connections among sentences that are obscure only to the pedantic English teacher. Not only do these models incorporate reader response in general and reader inferencing in particular; they identify specific conditions for allowable inference chains, specific strategies for appropriate placement of inference chains, and specific means for evaluating and controlling those chains.

As explicit editing devices, the patterns provide students with specific means for checking the cohesion of their work before turning it in. No student deliberately writes a noncohesive paper; by the time she finishes, she
thinks she has achieved cohesion. If she is given patterns as an editing check list, she has immediate feedback which in turn can replace doubt and uncertainty with the satisfaction of having mastered a precise skill.

In their several roles as generative, evaluative and editing devices, the patterns provide formal guidance to teachers and students while they provide at the same time a welcome breath of freedom. The fact that the patterns are derived from published texts in which there are many paragraphs without topic sentences can free the inexperienced teacher—especially the Ph.D. in literature condemned to forced labor in the composition camps—from the strangling grip and deadening pedagogy of the topic sentence. Students do read, and they notice lively, cohesive paragraphs without topic sentences, which can only leave them depressed by the lack of connection between their composition course and any real world writing they may be asked to do. The patterns provide the means for constructing and evaluating paragraphs without topic sentences.

After suggesting the various advantages that might accrue to those who use these models—and I think of them as advantages only, not the kind of panacea many teachers want to make out of sentence combining—let me suggest briefly here a three-stage introduction of the models into the classroom. The first task, obviously, is to identify them: in my own classes I use most of the sample paragraphs
discussed in the dissertation as examples of the various possible patterns and their requirements. The second task is to introduce additional samples for classroom discussion and exercises. The purpose here is to help the students see alternative arrangements of materials, along with the semantic and rhetorical implications of those arrangements, where before they saw none. The third task is to ask the students to rewrite inefficiently designed paragraphs and thus to demonstrate that they can not only recognize the patterns but can execute them as well.

Finally, I want to suggest that in whatever teaching capacity the models are used, they reflect the current—and, one believes, accurate—views of rhetoricians and discourse analysts. The models assume discourse to be a process not a product, an interaction between individuals that requires the alert and attentive application of shared and informing structures to bridge the gaps in world knowledge between the writer and the reader.

While most systems of discourse analysis use the metaphor of a tree hierarchy (whether explicitly or implicitly) to describe discourse structure, the analysis in this dissertation suggests the metaphor of a systems network. The network consists of the whole text, made up of dominant terms (or nodes) which are introduced, explicated, and connected with other terms (or nodes) through the relationships established by syntactic and semantic dominance. The
particular selection of these terms and relationships is governed by the communicative function of the text as a whole. Communicative function determines the complexity of the topic to be developed—whether by single-, double-, or combination recurrence chains—and also the amount of world knowledge that is assumed to be shared by author and reader. Seen as a network, discourse structure involves a system of interdependent components and "not merely ... a hierarchy of lower-level units or ... a one-dimensional linear concatenation," as implied by the metaphor of a tree hierarchy based exclusively on subordination.

Whether the metaphor of a network can withstand further scrutiny remains to be seen. But to the extent that it can explain certain demonstrably functional structures (without pretending that these structures fully account for cohesion), and to the extent that it can identify the necessary interface between people and world knowledge, it can perhaps take us a step further in understanding the patterns that underlie all discourse. It enables us to see these patterns as distillations rather than reductions—distillations analogous to a person's perception of a shape on a page that allows her to say that this shape is an "a" not an "o", but also allows her to say that this "a" is more elegant than that "a".
NOTES TO CHAPTER FIVE


BIBLIOGRAPHY


Frederickson, Carl H. "Representing Logical and Semantic Structure of Knowledge Acquired from Discourse." Cognitive Psychology. 7 (1975), 371-459.


Leech, Geoffrey. "'This Bread I Break'—Language and Interpretation," *A Review of English Literature*, 6 (1965), 66-75.


Schlesinger, I. M. "Why a Sentence in which a Sentence in which a Sentence is Embedded is Embedded is Difficult." *Journal of Psycholinguistics*, 4 (1975), 53-66.


