University Microfilms Inc.
300 N. Zeeb Road, Ann Arbor, MI 48106
INFORMATION TO USERS

This reproduction was made from a copy of a manuscript sent to us for publication and microfilming. While the most advanced technology has been used to photograph and reproduce this manuscript, the quality of the reproduction is heavily dependent upon the quality of the material submitted. Pages in any manuscript may have indistinct print. In all cases the best available copy has been filmed.

The following explanation of techniques is provided to help clarify notations which may appear on this reproduction.

1. Manuscripts may not always be complete. When it is not possible to obtain missing pages, a note appears to indicate this.

2. When copyrighted materials are removed from the manuscript, a note appears to indicate this.

3. Oversize materials (maps, drawings, and charts) are photographed by sectioning the original, beginning at the upper left hand corner and continuing from left to right in equal sections with small overlaps. Each oversize page is also filmed as one exposure and is available, for an additional charge, as a standard 35mm slide or in black and white paper format.*

4. Most photographs reproduce acceptably on positive microfilm or microfiche but lack clarity on xerographic copies made from the microfilm. For an additional charge, all photographs are available in black and white standard 35mm slide format.*

*For more information about black and white slides or enlarged paper reproductions, please contact the Dissertations Customer Services Department.
THE STRUCTURALIST ENTERPRISE AND ARISTOTLE'S POETICS

VOLUME I

DISSERTATION

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

By

Harry Mitchell Flower III, B.A., M.F.A.

*****

The Ohio State University
1986

Reading Committee:

Donald R. Glancy
Alan Hausman
Alan Woods

Approved By

[Signature]
Adviser
Department of Theatre
I dedicate this work to my loving parents to whom most of the credit belongs for any good I may do in this life.
ACKNOWLEDGEMENTS

I want to take this opportunity to gratefully acknowledge the invaluable help given me by Donald R. Glancy, Alan Hausman, Alan Woods, Lee Brown, and Robert Turnbull. They are responsible for much—if not most—of what is worthwhile in this work.
VITA

December 11, 1952

Born—Anoka, Minnesota

1973

Received B.A. in Theatre from The University of Texas at Austin

1975–76

Teaching Associate, Department of Theatre, The University of Minnesota, Minneapolis

1976

Received M.F.A. in Theatre from The University of Minnesota, Minneapolis

1976–80

Teaching Associate, Department of Theatre, The Ohio State University, Columbus

1980–81

Presidential Fellowship, The Ohio State University, Columbus

1981–85

Teaching Associate, Department of Philosophy, The Ohio State University, Columbus

FIELDS OF STUDY

Dramatic Theory and Criticism, Aesthetics
# TABLE OF CONTENTS

**DEDICATION** ................................................. ii  
**ACKNOWLEDGEMENTS** .................................. iii  
**VITA** ..................................................... iv  
**LIST OF FIGURES** .......................................... vi  

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. <strong>INTRODUCTION</strong></td>
<td>1</td>
</tr>
<tr>
<td>II. <strong>SYSTEMS, STRUCTURES, AND THE ANALOGY BETWEEN LITERATURE AND LANGUAGE</strong></td>
<td>21</td>
</tr>
<tr>
<td>Towards a Theory of Systems and Structures</td>
<td>28</td>
</tr>
<tr>
<td>The Nature of Systems</td>
<td>32</td>
</tr>
<tr>
<td>The Nature of Structures</td>
<td>52</td>
</tr>
<tr>
<td>The Analogy Between Literature and Language</td>
<td>55</td>
</tr>
<tr>
<td>More on Hierarchies</td>
<td>75</td>
</tr>
<tr>
<td>III. <strong>LINGUISTICS AND LITERATURE: THE TAXONOMIC APPROACH</strong></td>
<td>80</td>
</tr>
<tr>
<td>Taxonomic Linguistics</td>
<td>85</td>
</tr>
<tr>
<td>Langage, Langue, and Parole</td>
<td>86</td>
</tr>
<tr>
<td>Literary Application</td>
<td>101</td>
</tr>
<tr>
<td>Synchronic and Diachronic Approaches</td>
<td>120</td>
</tr>
<tr>
<td>Literary Application</td>
<td>128</td>
</tr>
<tr>
<td>The Units and Principles of Langue</td>
<td>140</td>
</tr>
<tr>
<td>Implications and Extensions of Saussure's Analysis of Langue</td>
<td>163</td>
</tr>
<tr>
<td>Literary Application</td>
<td>177</td>
</tr>
<tr>
<td>IV. <strong>LINGUISTICS AND LITERATURE: THE GENERATIVE APPROACH</strong></td>
<td>233</td>
</tr>
<tr>
<td>Literary Application</td>
<td>265</td>
</tr>
</tbody>
</table>
**LIST OF FIGURES**

<table>
<thead>
<tr>
<th>FIGURES</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Sample Tree Diagram of a Sentence's Phrase Marker</td>
<td>248</td>
</tr>
</tbody>
</table>
CHAPTER I
INTRODUCTION

During the first decade of the twentieth-century, the Swiss linguist Ferdinand de Saussure envisioned a science of "semiology" that would utilize modern linguistic concepts and techniques in its study of all human signs. The notion that such concepts and techniques might profitably be utilized outside the domain of linguists remained more or less dormant, however, until the late 1950's when Claude Lévi-Strauss began to advocate their use within the field of cultural anthropology. Shortly thereafter, linguistic

---

1 Ferdinand de Saussure, Course in General Linguistics, eds. Charles Bally and Albert Sechehaye, trans. Wade Baskin (New York: McGraw-Hill Book Company, 1966), pp. 16-7. Although Saussure is listed as the author of the Course, the book was actually assembled by Bally and Sechehaye—primarily on the basis of notes taken by students who attended classes Saussure taught between the years of 1906 and 1911; see Bally and Sechehaye's preface to the Course, pp. xvii-xviii.

In accordance with the MLA Handbook, the present study will enclose in double quotation marks all words to which attention is being directed, including "words purposely misused or used in a special sense"—MLA Handbook (New York: Modern Language Association, 1977), p. 12.

concepts and techniques began to be employed in fields throughout the social sciences, humanities, and the fine arts, and intellectuals began to speak of something called "structuralism," a word whose uses over the course of the intervening years have been so varied that some discussion of them seems necessary.3

One group of writers maintains that the term "structuralism" ought to be associated exclusively with the sort of methodological approach envisioned by Saussure and popularized by Lévi-Strauss. Thus, for example, in 1967 Roland Barthes characterized structuralism as "a certain mode of analysis of cultural artefacts, in so far as this mode originates in the methods of contemporary linguistics."4 Connected with the preceding, primary use of the term, there is, according to such writers, a secondary, derivative use; the term can also be used to refer to an intellectual movement that is associated with the

---

3 The following remarks are by no means intended to constitute an exhaustive survey of the myriad ways in which the term "structuralism" has been used and characterized; only three or four of what are probably the most prominent characterizations of the term will be mentioned. For a more complete list of uses, see Jonathan Culler's book, On Deconstruction: Theory and Criticism after Structuralism (Ithaca: Cornell University Press, 1982), pp. 18-28.

application and promotion of the approach just described—a movement whose rise, according to Fredric Jameson, "may be conveniently dated from the publication of Lévi-Strauss' *Tristes tropiques* in 1955."\(^5\)

Another group of writers believe that the term "structuralism" ought to be associated with a broader methodological approach that encompasses the aforementioned one as a particular instance or manifestation—an approach that is not dependent on the use of linguistic concepts and techniques and has applications not only within the social sciences, the humanities, and the fine arts, but also within mathematics, the physical sciences, and the biological sciences. Such a characterization is, for example, advanced by Jean Piaget in his book *Structuralism*.\(^6\) Such writers do not usually speak of a clearly datable movement called "structuralism" for, in the words of Piaget, "though a good many applications of the method are new, structuralism itself is not. It has a long history, which forms part of the history of the sciences . . . ."\(^7\)

---


\(^7\) Piaget, p. 136.
Of the two preceding characterizations, the former, more restrictive one comes closer to capturing the notion of structuralism that will be of concern in the present work; yet it is still not restrictive enough—for the concepts and techniques of modern linguistics have provided a basis for two quite different approaches to the study of literary phenomena. In 1976 J. Hillis Miller observed that

Already a clear distinction can be drawn . . . between what might be called . . . Socratic, theoretical, or canny critics, on the one hand, and Apollonian/Dionysian, tragic, or uncanny critics, on the other. Socratic critics are those who are lulled by the promise of a rational ordering of literary study on the basis of solid advances in scientific knowledge about language. . . . Included [as examples of this type of criticism] would be aspects of the work of Gérard Genette, Roland Barthes, and Roman Jacobson, as well as that of scholars like A. J. Greimas, Tzvetan Todorov, Cesare Brandi, and Jean-Claude Coquet. Jonathan Culler's *Structuralist Poetics* is a canny . . . introduction to the work of such critics. . . .

Opposed to these are the critics who might be called "uncanny." Though they have been inspired by the same climate of thought as the Socratic critics and though their work would also be impossible without modern linguistics, the "feel" or atmosphere of their writing is quite different from that of a critic like Culler, with his brisk common sense and his reassuring notions of "literary competence" and the acquisition of "conventions," his hope that all right-thinking people might agree on the meaning of a lyric or a novel, or at any rate share a "universe of discourse" in which they could talk about it. "Uncanny" critics would include, each in his own way, a new group of critics gathered
at Yale: Harold Bloom, Paul de Man, and Geoffrey Hartman. Jacques Derrida teaches a seminar early each fall at Yale and so may be included among the Yale group.8

As Miller observes in an earlier passage, such "uncanny" critics not only question the "canny" supposition that literary phenomena can be characterized and explained in a systematic way, they question even "the apparently solid basis of new linguistic theory."9 In fact, their skeptical analyses of literary phenomena are usually based on their similarly skeptical analyses of linguistic phenomena. Unlike his "canny" counterparts, Derrida, for example, argues that the concepts and techniques of modern linguistics can be used to show that the discipline of linguistics itself will never succeed in formulating a systematic account of linguistic phenomena. Since all literary texts are made out of language, the same factors that prevent formulation of a systematic account of

8 J. Hillis Miller, "Stevens' Rock and Criticism as Cure, II," The Georgia Review, 30 (1976), 335-36. As Miller's remarks indicate, Jonathan Culler's book Structuralist Poetics (London: Routledge & Kegan Paul, 1975) constitutes an example of "canny" criticism, and, in fact, the tenth chapter of that book is devoted to a critique of "uncanny" views. Nevertheless, by the time he wrote On Deconstruction in 1982, Culler had become an advocate of such "uncanny" views. The fact that Culler has held incompatible views at different times should, therefore, be born in mind when the present study makes references to Culler's works.

linguistic phenomena will also prevent critics from formulating a systematic account of literary phenomena. Thus, although both the "canny" and "uncanny" critics make use of modern linguistic concepts and methodologies in their study of literary phenomena, they manifest radically different attitudes towards those concepts and methodologies and use them for quite different purposes.

The approach utilized by the critics Miller calls "uncanny" is often referred to as either "deconstructionism," or, more broadly, "post-structuralism"—but some writers, including Miller himself, classify both it and its "canny" counterpart under the rubric of "structuralism." Such a broad usage of the term, argue Miller and other like-minded critics, is legitimate because the deconstructionist

---


view represents the results of pushing the "canny" approach to its logical limits, where it suddenly turns into its own opposite; thus, deconstructionism is really just a logical extension of the "canny" approach, and the two should, therefore, be grouped together under the broad umbrella of "structuralism." In On Deconstruction, Jonathan Culler cites other reasons often advanced in favor of such a broad use of the term "structuralism": the names of many of the key figures in recent critical theory have been linked at one time or another with both approaches, and there are even cases where a writer appears to employ both approaches during the course of a single work.

The present study does not, however, adopt the broad use of the term "structuralism" promoted by critics like Miller—for several reasons. First of all, the present work is primarily concerned with the more "canny" of the two approaches described above, and it is inconvenient to

12 Miller, "Stevens' Rock," pp. 343-44; see also Lewis, "Post-Structuralist Condition," pp. 8-11. Incidentally, the above argument represents an example of a very common deconstructionist strategy, viz., the attempt to undermine and dislocate—to deconstruct—what appears to be an obvious opposition (in the present case the opposition between the approaches adopted by the "canny" and "uncanny" critics) by arguing that if pushed far enough one or both of the opposed items will collapse into the other. See, for example, Jacques Derrida, Positions, trans. Alan Bass (Chicago: University of Chicago Press, 1981), pp. 41-43--also, Culler, On Deconstruction, pp. 140, 143, 150.

13 Culler, Deconstruction, pp. 25-28.
have to differentiate that approach from deconstructionism by appending some qualificatory description to the term "structuralism" every time a reference to the approach is made. Second, despite the fact that it is not always easy to determine whether a given author, or even a given work, should be associated with one rather than the other of the two approaches, there are, nevertheless, paradigm cases of each. Thus, for example, even though Roland Barthes has written works of both sorts, Tzvetan Todorov's works belong in the "canny" group, and, without question, the writings of both Derrida and Miller belong to the set of "uncanny" works. Third, the argument that both approaches are versions of structuralism because one represents the logical opposite and, hence, the logical extension of the other, seems less than compelling; it is, after all, just such opposition that is usually considered the best reason for sharply distinguishing one theory or approach from another, and even if such oppositions were all shown to be "deconstructible," they would in all likelihood still be considered the best reason for distinguishing between theories or approaches simply because there do not appear to be any

---

14 For an example of Todorov's work, see his The Poetics of Prose, trans. Richard Howard (Ithaca: Cornell University Press, 1977); for an example of Derrida's work, see Of Grammatology; for an example of Miller's work, see "The Critic as Host," Critical Inquiry, 3 (1977), 439-47.
alternative criteria that would provide a stronger basis for drawing such distinctions. Finally, Saussure and Lévi-Strauss, the figures who are generally thought to be most responsible for development of the term "structuralism," both maintained an attitude toward the use of linguistic concepts and techniques in the study of cultural phenomena that is, like the optimistic views of the "canny" critics, diametrically opposed to the skeptical views of the "uncanny" critics.¹⁵

Except when it appears as part of a quotation, accordingly, the term "structuralism" is used throughout the remainder of the present work to designate an approach that is based on the assumption that modern linguistics has made progress towards the formulation of an accurate and systematic account of linguistic phenomena—an approach which, therefore, utilizes the concepts and techniques

¹⁵ As earlier remarks have already indicated, Saussure believed that the use of linguistic concepts and techniques in the analysis of other sorts of cultural phenomena would result in the formation of a new science—not simply a new academic field—but a science (see above, p. 1); and Lévi-Strauss made it clear that he felt the same way: "Structural linguistics will certainly play the same renovating role with respect to the social sciences that nuclear physics, for example, has played for the physical sciences" (Structural Anthropology, p. 33). To say that the "uncanny" critics do not share the preceding views would be, as remarks made during the course of the last several pages have shown, a definite understatement.
of modern linguistics in the study of other sorts of cultural phenomena, e.g., literary phenomena, in hopes that similar progress can be made in those areas. The term is, in other words, used to refer to an approach that both views and utilizes linguistic concepts and techniques in what might be called a "positive" manner. In addition, the term is also used to refer to an intellectual movement associated with the application and promotion of the preceding approach, a movement whose lineage includes both Saussure and Lévi-Strauss. Except when it occurs as part of a quotation, the term "Structuralism" is hereinafter routinely capitalized as a reminder that it is being used exclusively with reference to the preceding two characterizations.16

16 When the term occurs within the context of a quotation, it of course appears as it does in the original source. Whenever the term is associated with characterizations other than the ones advanced above outside the context of a quotation, it is not capitalized and is surrounded by quotation marks—in accordance with conventions adopted earlier—see above, p. 1, fn. # 1. Thus, for example, a remark might refer to Piaget's conception of "structuralism."

In order to avoid a further potential misunderstanding about the sort of "structuralism" that will be of primary concern in the present study, it may be prudent to discuss briefly one more of the various features that have at one time or another been posited as hallmarks of the approach. In the Introduction to his anthology on the subject, Michael Lane claims that "structuralism is effectively anti-causal. The language of structuralist analysis in its pure form makes no use of the notions of cause and effect: rather it rejects this conceptualization of the world in favour of 'laws of transformation'. . . . What the structuralists are
As Culler observes, the application of Structuralism in the area of literary studies results in the attempt to develop something like "grammars" for various types of literary works, an attempt which the "uncanny" critics believe cannot possibly succeed:

saying is essentially this: if we compare two patterns of social relations . . . separated by time or space . . . we observe differences in their respective structural configuration, the order and nature of the patterning of relations. Traditionally we should argue that a particular factor or factors caused the one to differ from the other. The structuralist would argue, instead, that we can only say that a certain structure is seen to be transformed into another structure, and that repeated observations permit us to say that a given structure is always transformed in a particular way, thus giving us not causal laws (since that concept has not been invoked) but laws of transformation" (Lane, Structuralism, p. 17). Some Structuralists have no doubt maintained views of the foregoing sort; Lévi-Strauss, for example, tends to talk about the relations between various different myths or various different social practices in transformational rather than causal terms (see, for instance, Claude Lévi-Strauss, The Raw and the Cooked, trans. John and Doreen Wightman [New York: Harper and Row, 1969], p. 2). Such views are, however, hardly maintained by all Structuralists. For example, while he thought that knowing the historical causes that led to the structural configuration of a language at a particular time would not help a linguist understand what enables a native speaker actively to speak and comprehend utterances of that language, Saussure, nevertheless, quite explicitly states that the structural configuration of a language at any given time does have definite causes and devotes several sections of the Course to an investigation of those causes (see pp. 18, 98, 147-73, 205-11). Moreover, the preceding remarks concern only one type of causation, viz., so-called "efficient" causation, and subsequent discussion will suggest that there may be a much more intimate connection between the Structuralist approach and other types of causation (see below, pp. 353-55, 483-94, 519-23). Even where the notion of efficient causation is concerned, however, the preceding remarks should furnish sufficient proof that Structuralism is perfectly compatible with adoption of a causal stance.
.... structuralists take linguistics as a model and attempt to develop "grammars"—systematic inventories of elements and their possibilities of combination—that would account for the form and meaning of literary works; post-structuralists investigate the way in which this project is subverted by the workings of the texts themselves. Structuralists are convinced that systematic knowledge is possible; post-structuralists claim to know only the impossibility of this knowledge.

The attempt to make advances towards a systematic characterization and explanation of literary phenomena through the use of linguistic concepts and techniques—especially the attempt to develop something like "grammars" of literature—may be referred to as the "Structuralist enterprise"; it is, in effect, a research program. The approach utilized by the "uncanny" critics will, outside the context of quotations, be referred to as "deconstructionism." 18

17 Culler, Deconstruction, p. 22.

18 "Deconstructionism" has been chosen rather than "post-structuralism" because most of the works commonly grouped under the former label manifest the sort of approach described during the course of the last few pages. In contrast, the term "post-structuralism" is commonly used as a label for not only deconstructionist, but all manner of critics that are not clearly Structuralist; see, for example, Josué V. Harari, "Critical Factions / Critical Fictions," in Textual Strategies: Perspectives in Post-Structuralist Criticism, ed. Josué V. Harari (New York: Cornell University Press, 1979), pp. 17-81.
Unfortunately, the enthusiasm with which literary critics first embraced the Structuralist approach during the 1960's and early 70's began to wane in the late 70's, when many of them began to wonder if the goals of the Structuralist enterprise could ever be realized. Without question, the activities of critics associated with newer, rival approaches were partly responsible for the aforementioned trend—especially the deconstructionists, who, as the last few pages have indicated, vigorously challenged the very foundations of the whole Structuralist enterprise. It would appear, however, that the trend was also due in part to certain practices on the part of critics in the Structuralist camp itself, practices that were not only problematic, but avoidable.

Many problems derived from the very way in which literary Structuralists treated the concepts and techniques they borrowed from linguistics. First of all, Structuralist critics tended simply to assume without question the adequacy of a given linguistic concept or methodology with regard to its original domain of application. The characterization of the Structuralist approach advanced above stated that it is based on the assumption that linguistics has made progress towards the formulation of an accurate and systematic account of linguistic phenomena, but that in no way implies that the Structuralist approach must assume that
all of the concepts and techniques that have been employed by linguists are fully adequate to the task of describing and explaining linguistic phenomena. For one thing, the underlying assumption is only that linguistics has made at least some progress towards attaining its goals—not that it has already arrived. Moreover, the historical process that led to what progress has already been made contained failures, false starts, the development of concepts and techniques that were wholly problematic, the development of some that were partly problematic, etc. Before a critic attempts to apply a given linguistic concept or technique within the domain of literary studies, it would, therefore, seem prudent for him to ascertain how useful it is as a tool for describing and characterizing the phenomena associated with the use of natural languages. If a critic overlooks the fact that a particular linguistic concept or technique is problematic with regard to its original domain of application, as subsequent remarks will suggest has been the case with some of Saussure's most influential concepts, the problematic features of that concept or technique are likely to be perpetuated when it is applied in the realm of literature.

---

19 See above, pp. 9-10.
20 See, for example, below, pp. 94-99, 163-67.
In addition to their tendency to apply linguistic concepts and techniques uncritically, Structuralist critics also tended to apply them in a superficial way. For example, literary applications of terms like "generative rule" and "transformation" have rarely borne any resemblance to the highly technical meanings such terms possess in their original domain. Such behavior naturally left Structuralists open to the charge that they were employing linguistic terms in name only in hopes that mere use of those terms would enable literary studies to acquire the reputation for rigor that linguistic studies have enjoyed.

Not all of the problematic aspects of previous Structuralist criticism derive, however, from the way in which linguistic concepts and techniques have been handled. Another problematic feature of such criticism lies in a tendency to ignore the work of critics from previous centuries, thereby eliminating the possibility that such work could be used in solving difficulties currently facing the Structuralist enterprise.

The present study can be broadly characterized as an attempt to formulate a Structuralist account of drama that avoids the aforementioned problematic aspects of previous Structuralist literary theory. The work as a whole can be
divided into two parts. The first is devoted primarily to an examination of how the former problems concerning the literary application of linguistic concepts and methodologies might be avoided. Chapter Two attempts to elucidate the crucially important concepts of "system" and "structure" and then examines the appropriateness of the analogy between literature and linguistics. The third chapter explores the literary application of concepts and techniques drawn from so-called "taxonomic" linguistics with a view to avoiding the sorts of problems described earlier. The fourth chapter, which brings to a close the first part of the study, does much the same thing vis a vis so-called "generative" linguistics.  

21 As the preceding remarks indicate, subsequent remarks often call into question the adequacy of a given linguistic concept in the realm of literary studies; but the present work by no means calls into question the adequacy of all linguistic concepts and procedures or the way in which all such concepts and procedures have previously been applied in the realm of literary studies. Naturally, the present work therefore does not call into question the foundational assumptions of the Structuralist approach either, viz., (1) the assumption that modern linguistics has made at least some progress in its attempt to formulate a systematic and accurate characterization and explanation of linguistic phenomena, and (2) the belief that using the concepts and techniques of modern linguistics could help scholars in other fields, like literary studies, make some progress towards the formulation of accurate, systematic accounts of those phenomena. Since deconstructionists have challenged not only the preceding assumptions, but have advanced arguments designed to undermine virtually every Structuralist concept and principle that is discussed.
In keeping with the overall goal of establishing a firm foundation for the Structuralist approach to literature, the discussions contained in the first four chapters not only consider the adequacy of various linguistic concepts and techniques vis-à-vis the domain of language more carefully during the remainder of the present work, it might be argued that this study must refute the deconstructionist arguments against a given concept or principle before that concept or principle can be accepted—for each and every concept and principle discussed, including the basic assumptions of the Structuralist approach itself. There are, however, several reasons why it does not appear to be necessary to provide such an exhaustive, point-by-point response to the arguments that have been advanced by deconstructionists.


Unfortunately, a long and complicated discussion is required in order to explain the second reason why it does not seem necessary to provide an exhaustive, point-by-point discussion of deconstructionist attacks on Structuralist concepts and techniques. Instead of attempting to include that discussion at this point, which would interrupt the flow of the present chapter, it is included as Appendix A at the end of this study; see below, pp. 546-73.
than have most previous writings, they also spell out the analogy between literature and language in more detail. Careful consideration is also given to various philosophical aspects of the material being discussed.

One of the most important points that emerges from the discussions contained in the first four chapters is that many—if not most—of the major concepts and techniques of modern linguistics are applicable to the domain of natural languages precisely because utterances can be segmented into a variety of hierarchically arranged types of units—from the phoneme to the sentence. The literary applicability of such concepts and techniques would, therefore, appear to depend upon whether literary works can be segmented in a similar fashion. Unfortunately, previous Structuralists have been unable to account for the overall structure of any type of literary work in terms of a linguistic-type hierarchy, and many have become pessimistic about the possibility of ever doing so, a state of affairs that would appear to constitute another source of the decline in Structuralism's popularity noted above: given that the success of the Structuralist enterprise is wholly dependent on

the degree to which linguistic concepts and techniques can be successfully applied in the domain of literature, and given that the successful application of such concepts and techniques is largely dependent on the isolation of a linguistic-type hierarchy in the field of literature, to give up on the possibility of ever finding such a hierarchy in the domain of literature is, in effect, to give up on the feasibility of the Structuralist program altogether.

The dissertation's second half attempts to show that pessimism with regard to the possibility of isolating a linguistic-type hierarchy in the field of literature could have been avoided if previous Structuralists had looked through the history of pre-twentieth-century literary theory for help. Specifically, the second part of the present study claims that the six so-called "qualitative" parts of tragedy identified by Aristotle in the Poetics exhibit the same sort of hierarchical organization that is exhibited by the utterances of natural languages. In order to defend the claim that Aristotle conceived of the qualitative parts as a series of hierarchically related elements, a fair amount of close textual exegesis, with the views of previous commentators in mind, is required. Over the last four centuries nearly every sentence of the Poetics has been subject to interpretive disputes, and no work that deals with the Poetics in any depth can completely ignore
the issues raised. Although such discussions tend to be rather technical, the fruits gained from them will hopefully justify their somewhat erudite character. Chapter Five provides a detailed characterization of each of the six qualitative parts, and the sixth chapter attempts to show that the parts do in fact bear hierarchical relations to each other.

If the arguments advanced in Chapters Five and Six are on the whole correct, they will have succeeded in showing that the decline in Structuralism was ill-founded. Further study of the qualitative parts in light of modern linguistic concepts and methodologies may yet lead to advances in the field of literary theory similar to those that have occurred in linguistics.
CHAPTER II
SYSTEMS, STRUCTURES, AND THE ANALOGY BETWEEN LITERATURE AND LANGUAGE

The characterization of Structuralism advanced during the previous chapter placed primary emphasis on the features that most distinguish it from other methodological approaches: it is an approach employed in the study of cultural phenomena based on a "positive" use of modern linguistic concepts and techniques. There is, however, an additional feature that should be included in any characterization of Structuralism, a feature that is also exhibited by both the broader, Piagetian brand of "structuralism" briefly identified at the outset of the preceding chapter and by work in what has come to be called "general systems theory"—namely, as Michael Lane puts it, "the emphasis [Structuralism] gives to wholes, to totalities."¹

Not only are Structuralists, Piagetian "structuralists," and those working in the area of general systems

¹ Michael Lane, Introd., Structuralism, p. 14. For the earlier remarks concerning Piagetian "structuralism," see above, p. 3. A rough characterization of general systems theory will be provided during the course of the next few pages.
theory all primarily interested in the study of holistic entities, they also hold similar views concerning the proper explanation of such entities. For example, (1) they all believe that the properties of a holistic entity can be adequately explained only in terms of the relations that exist between the entity's parts, and (2) they all believe that recognition of the preceding fact separates their approach from earlier ones. For instance, according to Lane:

Traditionally in Anglo-American social science, structure has been used as an analytical concept to break down sets into their constituent elements, an essentially atomistic excercise. As structuralists understand and emply the term, a new importance

2 A good deal of discussion will be necessary before a precise characterization of the distinction between holistic and non-holistic entities can be stated. However, in very rough terms, the distinction can be put as follows: whereas the parts of a whole contract strong functional relations with one another, the parts of a non-holistic entity do not. A more precise characterization is advanced below, pp. 42-43, 45-46.

3 As remarks contained in the preceding chapter indicated, Piagetian "structuralists" claim that although many of the applications of their views are new, the views themselves are not (see above, p. 3). They would not, therefore, claim the sort of novelty for their approach that the following remarks will show Structuralists and general systems theorists claim for their approaches. Like Structuralists and general systems theorists, they do, however draw a sharp distinction between their own views and those associated with certain earlier approaches they typically refer to as "atomism," "reductionism," or "mechanism" (see, for example, Piaget, pp. 4, 45). A discussion of some of the similarities between the views of general systems theorists and those associated with certain versions of "structuralism" can be found in Ervin Laszlo, Systems, Science, and World Order (New York: Pergamon Press, 1983), pp. 97-111.
has been given to the logical priority of the whole over its parts. They insist that the whole and the parts can be properly explained only in terms of the relations that exist between the parts. The essential quality of the structuralist method, and its fundamental tenet, lies in its attempt to study not the elements of a whole, but the complex network of relationships that link and unite those elements.\(^4\)

Ludwig von Bertalanffy, the man who has probably been most responsible for the development of general systems theory, draws a similar distinction when speaking of the two alternative approaches to biological phenomena that led him to posit the basic concepts of that theory—a contrast between, on the one hand, "biology, in the mechanistic conception, [which] saw its goal in the resolution of life phenomena into atomic entities and partial processes," and the "organismic" approach, whose goals are, instead, based on the assumption that

It is necessary to study not only parts and processes in isolation, but also to solve the decisive problems found in the organization and order unifying them, resulting from dynamic interaction of parts, and making the behavior of parts different when studied in isolation or within the whole.\(^5\)


In fact, von Bertalanffy's belief that the preceding distinction can be found virtually throughout academia constitutes one of the basic assumptions of his general theory of systems:

Classical science in its diverse disciplines, be it chemistry, biology, psychology, or the social sciences, tried to isolate the elements of the observed universe—chemical compounds and enzymes, cells, elementary sensations, freely competing individuals, as the case may be—expecting that by putting them together again conceptually or experimentally, the whole or system—cell, mind, society—would result and be intelligible. Now we have learned that for an understanding not only the elements but their interrelations as well are required: say, the interplay of enzymes in a cell or of many mental processes conscious and unconscious, the structure and dynamics of social systems, and so forth. This requires exploration of the many systems in our observed universe in their own right and specificities. Moreover, it turns out that there are correspondences or isomorphisms in certain general aspects of "systems." This is the domain of general systems theory.  

Thus, "structuralists" of various varieties and general systems theorists all draw a sharp contrast between their own tendency to highlight the importance of constituent relations in the explanation of holistic entities and what they claim was an earlier tendency to try to explain a whole

---

exclusively in terms of the properties that its parts exhibit in isolation, a tendency which they variously refer to as either "atomism," "mechanism," or "reductionism."  

7 The accuracy of the above contrast between the views of "structuralists" and general systems theorists on the one hand and views associated with previous approaches on the other might, however, be questioned on a variety of different grounds. First, the tendency to portray the Structuralist and "systems" interest in holistic phenomena, their emphasis on the importance of constituent relations in the explanation of such phenomena, etc. as new and innovative aspects of recent thought seems somewhat exaggerated. Perhaps the widespread application of such views is relatively new, but the basic ideas themselves—e.g., the notion that the properties of a whole cannot be explained exclusively in terms of the properties exhibited by its parts in isolation—are hardly novel; see, for example, the section entitled "From Aristotle to Brentano," which opens Barry Smith and Kevin Mulligan's discussion of developments in holistic thought in their paper "Pieces of a Theory," in Parts and Moments: Studies in Logic and Formal Ontology, ed. Barry Smith (Munich: Philosophia Verlag, 1982), pp. 15-25.

Associated with the preceding issue are questions concerning the accuracy of "structuralist" and "systems" characterizations of previous approaches. For example, their tendency to associate the terms "atomism," "mechanism," and "reductionism" with the claim that the parts of a whole will exhibit the same properties in isolation that they exhibit when combined and that the properties of a whole can, therefore, be explained exclusively in terms of the properties that its parts exhibit in isolation raises questions about which—if any—of the various specific scientific and philosophical doctrines that have been associated with each of those terms actually entail the claim in question. In support of his association of the claim with the term "mechanism," von Bertalanffy alludes to a passage in Bertrand Russell's Human Knowledge: Its Scope and Limits (von Bertalanffy, pp. 67-68). The relevant portions of the passage run as follows: "[The doctrine of organicism] holds that the body of an animal or plant is a unity, in the sense that the laws governing the behavior of the parts can only be stated by considering the
place of the parts in the whole. . . . The opposite view, which I should regard as correct, would say that to understand what [for example] an eye does, you need to know in addition to its own structure, only the inflow and outflow of energy. . . . The mechanistic view holds that if an eye is separated from its body, but preserves its structure and chemical constitution, and is provided with artificial nerves to drain away the impulses received from incident light, it will behave as it would if it were still in its proper place. . . . Speaking generally, scientific progress has been made by analysis and artificial isolation. It may be that, as quantum theory suggests, there are limits to the legitimacy of this process, but if it were not usually or approximately valid scientific knowledge would be impossible. It is therefore in any case prudent to adopt the mechanistic view as a working hypothesis, to be abandoned only where there is clear evidence against it" (Bertrand Russell, Human Knowledge: Its Scope and Limits [New York: Simon and Schuster, 1948], pp. 35-6). The fact that Russell's remarks seem to suggest that the doctrine of mechanism entails the claim in question does not, however, prove that it entails that claim; Russell might well be wrong. It would, no doubt, be necessary to embark upon a long and complicated discussion of the doctrine of mechanism itself in order definitely to determine whether it actually does entail the claim in question, and similar discussions would be necessary in the case of the other two terms. Instead of trying to settle such matters, the present study will simply adopt the following stipulative convention: the terms "atomism," "mechanism," and "reductionism" are to be exclusively associated, throughout the remainder of the present study, with the view that the properties of an entity—no matter what type of entity it might be—can be explained solely in terms of the properties exhibited by its parts in isolation; use of the terms is to carry absolutely no suggestion that any of the specific doctrines that have been associated with the terms entail the view in question. The present study will also refrain from either endorsing or denying Structuralist and "systems" claims that their interest in holistic entities, their emphasis on the importance of constituent relations in the explanation of such entities, etc. are new and innovative. The historical origins of such views are of little importance for the purposes of present discussion anyway; it is the views themselves and the fact that they have at least been maintained by Structuralists and general systems theorists—regardless of whether others have also maintained them—that will be of importance.
While the aforementioned writers' discussions of holistic entities have much in common, they also exhibit substantial differences—especially when it comes to the use and characterization of two crucially important terms—"system" and "structure." Whereas general systems theorists tend to employ only the term "system" when referring to holistic entities, Structuralists and "structuralists" tend to employ both terms. To further complicate matters, some of the writers who employ both terms seem to use them as synonyms, while others sharply distinguish them from one another. Worse still, those who differentiate the terms from each other seldom do so in the same way. The first part of the present chapter will, therefore, be devoted to the tasks of (1) formulating characterizations of the preceding terms that can be used throughout the remainder of the present study and (2) providing some idea of the role that the concepts of system and structure play within the Structuralist approach. ⁸ Once that has been accomplished,  

⁸ As might be expected, the inconsistencies in usage of the terms "system" and "structure" by previous writers rule out the possibility of trying to formulate some sort of consensus characterizations of them, and the following remarks are by no means to be seen as such an attempt. Nevertheless, the following discussion will draw heavily upon what has been said about the subject by previous writers—not only by Structuralists, but also by "structuralists," and general systems theorists. As a result, the characterizations of the terms advanced later
the concepts of system and structure will be used to help explain the appropriateness of the Structuralist analogy between literature and language.

TOWARDS A THEORY OF SYSTEMS AND STRUCTURES

Since the following discussion is to be primarily concerned with the characterization of holistic entities, the fact that the term "structure" has several different senses, only one of which is exclusively associated with such entities, warrants some comment at the outset. While one sense of the term is used, as earlier remarks have indicated, exclusively with reference to holistic entities, the term can also be used as a means of simply referring to a collection of items or parts--without necessarily implying in the present chapter will reflect major trends in previous thought about the subject even if they cannot be said to reflect an established consensus.

One additional point about the use of previous work during the following discussion bears special emphasis. Many previous authors--especially general systems theorists and "structuralists"--have claimed that their characterizations succeed in accurately capturing the essential attributes of holistic entities found in every domain of academic study--from physics to mathematics, biology, sociology, the arts, etc. While the following remarks will attempt to formulate characterizations that have fairly broad application, the formulation of characterizations that apply within the domains of linguistics and literary studies constitutes their primary goal, and subsequent remarks should by no means be construed as claiming that those characterizations are universally applicable.
that the collection in question exhibits what subsequent remarks will argue are the defining properties of a whole.\textsuperscript{9} The second sense of the term is, in other words, roughly synonymous with the expression "composite entity." Previous writers have rarely bothered explicitly to distinguish the foregoing senses of the term from each other, and, as a result, it is not always clear which of the two senses they have in mind. At one point in his book on Structuralism, for example, Robert Scholes asks the reader to "Consider for a moment the following verbal structures."\textsuperscript{10} Since Scholes has not, however, explicitly distinguished between the aforementioned senses of the term "structure," the exact implications of his statement are unclear: is he implying that the seven sentences that he then goes on to list are wholes in the strict sense, or is he simply referring the reader's attention to a list of verbal organizations without implying anything about the nature of their organization?

\textsuperscript{9} As the above remark suggests, the exact nature of the distinction between the preceding two senses of the term "structure" will become clearer when the distinction between holistic and non-holistic entities is fully explicated; see below, pp. 42-43, 45-46.

In order to avoid the preceding sorts of ambiguities, whenever the term is used to refer to holistic entities during the course of subsequent remarks (outside the context of quotation), it will be capitalized; whenever it is used to refer to a collection of items or parts without regard to the way in which those items or parts have been organized, it will not be capitalized.¹¹

There is, in addition to the preceding two senses, a third sense of the term "structure" that is, for example, employed by Laszlo in his statement that "Social systems can be analyzed from two distinct points of view. First, morphologically: here the structure of the system is under consideration."¹² It would appear that the term "structure" is being used in the preceding quotation as a means of designating the way in which the items or parts

¹¹ When the term occurs within the context of a quotation, it will, of course, appear as it does in the original. Citation of the word as it appears in the original will, however, be followed by a capitalized repetition of it enclosed in brackets when it seems clear that it is being used to designate holistic entities; when it is not clear exactly how the term is being used, no such attempt to disambiguate the meaning of the term will be made.

of a composite entity are arranged or organized rather than the set of items or parts arranged. The term is, in other words, being used to refer to an attribute of an entity—the way in which it is organized or arranged—rather than to the entity which exhibits that attribute.

There are thus at least three separate senses of the term "structure": (1) structure qua holistically organized collection of items or parts, (2) structure qua collection of items or parts—without regard to the way in which they have been organized, and (3) structure qua mode or manner of organization. Since it will rarely, if ever, be necessary to refer to the "structure" (in the third sense) of "structures" (in the second sense), and since it should be fairly obvious that the first occurrence of the term "structure" in phrases like "the structure of a Structure" refers to the third rather than the second of the three senses just listed, there appears to be no real need to graphically differentiate the second and third senses from each other. The term will, therefore, be capitalized only when it is being used to refer to holistic entities; all other occurrences of the term will be uncapitalized.

Whereas Structuralist and "structuralist" attempts to characterize the terms "system" and "Structure" have, as previous remarks indicated, produced widely divergent
results, general systems theorists seem to agree about the fundamental attributes of systems. Moreover, their discussions of the property of wholeness are among the most, if not the most, detailed treatments of the subject to date. The following discussion will, therefore, begin with a consideration of their work.

THE NATURE OF SYSTEMS

General systems theorists typically characterize the concept of "system" in terms of three fundamental attributes: (1) wholeness, (2) the capacity for dynamic activity, and (3) hierarchical organization. Discussion will begin with a consideration of what general systems theorists have said about the property of wholeness.

General systems theorists' discussions of the distinction between holistic and non-holistic entities, which they typically refer to as "aggregates," tend to focus on the well-worn maxim that "a whole is more than the sum of its parts." The meaning of the preceding maxim is, however, highly ambiguous, and general systems theorists have advanced two quite different explanations of its meaning. Sometimes they explain the point of the maxim in ontological terms, i.e., in terms of the properties that wholes and aggregates respectively possess. At one point, Laszlo says,
for example:

A whole possesses characteristics which are not possessed by its parts singly. Insofar as this is the case, therefore, a whole is other than the simple sum of its parts.  

Laszlo's characterization of aggregates one page later is, however, put in epistemological rather than ontological terms, i.e., in terms of whether it is possible to know by way of inference the properties of an aggregate solely on the basis of knowledge concerning the properties of its parts:

[In the case of non-holistic entities] it is sufficient to sum the properties of the parts to obtain the properties of the whole. Such wholes are better known as "heaps" or "aggregates" . . .

Von Bertalanffy advances a characterization of wholes that is similarly epistemological in tone:

The meaning of the somewhat mystical expression, "the whole is more than the sum of its parts" is simply that [a whole possesses] characteristics [that] are not explainable from the characteristics of isolated parts. The characteristics of the complex, therefore, compared to those of the elements, appear as "new" or "emergent." If,

---

13 Laszlo, Systems Philosophy, p. 36.
14 Laszlo, Systems Philosophy, p. 37.
However, we know the total of parts contained in a [whole] and the relations between them, the behavior of the [whole] may be derived from the behavior of the parts.\textsuperscript{15}

Unfortunately, general systems theorists do not have much to say about the relation between the two alternative ways of distinguishing wholes from aggregates, and it is not obvious that either characterization automatically entails the other. A person's inability to infer the properties possessed by a whole solely on the basis of information about the properties possessed by its parts might, for example, be due to the limitations of human reason rather than to some fundamental characteristics possessed only by those entities typically categorized as wholes. For reasons that will become apparent during the course of the following remarks, the present discussion will initially attempt to draw the distinction in epistemological terms.

The latter two quotations cited above seem to suggest that the distinction between wholes and aggregates is based on the following claim: it is possible to infer all of the properties possessed by an aggregate— but not a whole— solely on the basis of information about the properties that its parts possess in isolation— without, in other words, knowing

\textsuperscript{15} Von Bertalanffy, p. 55.
anything whatsoever about the way its parts are interrelated when combined. The preceding claim appears to be mistaken, however, where aggregates are concerned. In order to see why, it will be useful to consider an example of an aggregate—say, a heap of bricks, an example cited by both Laszlo and von Bertalanffy. 16 Among the properties possessed by any given heap will be those which concern the spatial relations between its constituent bricks or parts, e.g., the property associated with the fact that brick #1 is immediately to the left of brick #2. Such properties could not, however, be inferred solely on the basis of knowledge about the properties that each brick possesses in isolation and the knowledge that the heap includes bricks #1-#n. The foregoing will obviously hold true, moreover, in the case of any composite entity; the entity itself will possess certain properties associated with the interrelation between its parts that could not be inferred on the basis of knowledge about the properties that each part possesses in isolation and knowledge that the entity includes just the parts it does.

Perhaps the point of the preceding paragraph is that the distinction between wholes and aggregates concerns only

---

certain types of properties and interrelations, and, in fact, one of Laszlo's subsequent remarks seems to suggest as much:

A set of objects related merely by spatial or mechanical adjacency is a heap, or conglomeration, of externally related, mutually independent units. It is a summative complex, for it is properly described by the summation of individual parts. Taking away one part, or adding some, makes no difference to the character of the whole... Each part can be understood by itself; reference to the rest is not helpful or indicated.17

The foregoing paragraph has shown that the full range of properties possessed by an aggregate cannot be inferred, as Laszlo seems to be suggesting in the preceding quotation, by simply "summing" the properties possessed by its parts in isolation--at least if that means ignoring all of the interrelations that the parts contract when combined. The importance of Laszlo's remarks for present purposes lies, however, in the possible suggestion that the distinction between wholes and aggregates is to be drawn in terms of properties other than those which arise from merely spatial or mechanical relations--that the distinction must, in other words, be drawn on the basis of whether, given knowledge of only those properties that an entity's parts possess when

17 Laszlo, Systems Philosophy, p. 100.
isolated from each other, it is possible to infer—not all of the properties possessed by that entity—but rather properties of a particular type. If the properties in question are not, however, spatial or mechanical, what sorts of properties might they be?

Although the discussions of general systems theorists do not appear to contain an explicit answer to the foregoing question, they do seem to contain some suggestive hints. For example, the latter portion of Laszlo's characterization of aggregates, part of which was quoted several pages ago, states that non-holistic entities are better known as "heaps" or "aggregates," since the fact that the parts are joined in them makes no difference to their function—i.e. the interrelations of the parts do not qualify their joint behavior.\(^{18}\)

Elsewhere, Laszlo remarks that "systems form ordered wholes in which the law-bound regularities exhibited by interdependent elements determine the functional behavior of the totality."\(^{19}\) Perhaps, then, the distinction between wholes and aggregates is to be drawn in terms of functional properties and relations. It might, in other words, be


\(^{19}\) Laszlo, *World Order*, p. 41.
suggested that the distinction is to be drawn as follows: it is possible to infer all the functional properties of an aggregate—but not those of a whole—solely on the basis of information about the properties exhibited by the parts in isolation—without, that is, knowing anything about the interrelations contracted among the parts when combined.

Unfortunately, most of the examples of aggregates cited by previous writers could easily lead to the conclusion that the foregoing suggestion boils down to the claim that the distinction between wholes and aggregates is just another way of talking about the difference between functional and non-functional entities. In his discussion of aggregates, for instance, von Bertalanffy cites as examples not only a heap of bricks, but also "odds and ends." Now to describe a collection of things as "odds and ends" is to imply that the collection is associated with no clear purpose or function; description of a collection of bricks as a "heap" carries much the same implication. In fact, virtually all of the phenomena typically cited as examples of aggregates seem to carry much the same implication. If, however, all aggregates turn out to be entities associated with no clear

20 Von Bertalanffy, p. 67.
function or purpose, the distinction between wholes and aggregates will collapse into the distinction between functional versus non-functional collections of items—and most writers seem to think that there is more to the distinction than that.

What would be needed in order to show that the distinction does not simply collapse would be an example of a collective entity that is associated with the performance of one or more functions, but whose functional properties can be inferred and explained solely on the basis of knowledge about the functional properties exhibited by its parts in isolation. Can such examples be found? The answer would appear to be that they can. Imagine, for instance, a collection of people, each of whom has been trained to assemble a particular object on his own—while sequestered away in a room by himself perhaps. The group of people will possess the functional property of producing an average of roughly a certain number of objects per specified amount of time. That property can, however, be predicted and thoroughly explained solely on the basis of information about the average number of objects each individual person produces in the specified amount of time; the number of objects produced by the collective can be inferred by simply "summing" or adding together the number of objects produced by each individual. Such a collective would,
moreover, possess the other characteristics that, as an earlier remark has indicated, Laszlo associates with aggregates.21 The functional property of each member of the group—the average number of objects he is capable of producing—can be understood by itself; reference to the functional properties of the other members of the group—the average number of objects they produce—is not helpful or indicated.

In order to see why the same would not hold true in the case of a whole, consider a group of people, each of whom has been trained to assemble—not the whole object—but a particular portion of an object being produced, and suppose that the people have been arranged in such a way that the second person cannot assemble his portion of the object until the first person has assembled his portion, the third person cannot assemble his portion until the second person has assembled his, etc. Now it might be assumed that the average number of completed objects produced by the whole group will be equal to the average number of object-portions that the slowest member of the group produces in isolation, but that will not necessarily be so. When surrounded by faster workers, the slowest member of the group

21 See above, p. 36.
may end up producing more object-portions during the allotted amount of time than he normally would in isolation; if, on the other hand, he were surrounded by workers even slower than himself, the same worker might end up producing even fewer object-portions than he normally would in isolation. The same will hold true in the case of the other workers as well; the amount of object-portions each produces when part of the assembly line may vary depending on whether the workers on each side of him are slower or faster than he is. As a result, it will be impossible to infer or explain the average number of completed objects produced by the whole group during a specified amount of time solely on the basis of information about how many object-portions each individual can, on average, produce in isolation during that amount of time. It will not, in other words, be possible to infer or explain the functional properties of the whole solely on the basis of information about the functional properties that the parts exhibit when isolated from each other—additional information about the relations between the parts and the effects of those interrelations would be required in order to do that.

For purposes of exegesis, the remarks contained in the last two paragraphs have ignored the possibility that the groups referred to in the hypothetical examples might be associated with functions other than the production of
objects. That possibility should, however, be considered. Imagine, for example, that all the members of the group referred to in the first example belonged to a baseball team. It would then be appropriate to say that the group constitutes an aggregate vis a vis the production of objects, but that they constitute a whole vis a vis the playing of baseball. Since one and the same collection of items or parts may constitute an aggregate vis a vis some functions and a whole vis a vis others, it would appear that the aggregate/whole distinction should be drawn—not on the basis of whether it is possible to infer all of a collective's functional properties given information only about the properties possessed by its parts in isolation—but on the basis of whether it is possible to infer a specified function or set of functions given such information.

The distinction between wholes and aggregates may, therefore, be characterized as follows. A collection of items or parts is an aggregate vis a vis a specified functional property or set of such properties if and only if it is possible to infer that the collection will possess those properties on the basis of information about the properties that its parts exhibit when isolated from each other. A collection of items or parts is a whole vis a vis a specified functional property or set of such properties if and only if it is not possible to infer that the collection will
possess those properties on the basis of information about the properties that its parts exhibit when isolated from each other.

In order to prevent potential misunderstandings, two additional points should be made with reference to use of the term "function" in the preceding characterization. First of all, use of the term is in no way meant to suggest "purposive" behavior. Although the functional properties involved in the preceding examples are the sort that are typically guided by an intended purpose, neither the parts of a composite nor the composite itself need be guided by any sort of purpose in order to exhibit functional properties; they need only end up performing some function. Second, the distinction between functional relations on the one hand and merely spatial and mechanical relations on the other is not meant to imply that the two cannot be coextensive. On the contrary, functional relations are typically—if not always—instantiated by relations of a spatial, mechanical, or temporal sort. The converse is, however, not always true; spatial, mechanical, and temporal relations do not always instantiate functional relations. The word "merely" in the phrase "merely spatial or mechanical relations" has been included for the precise purpose of indicating that the relations in question are those which do not instantiate functional relations—as opposed to those which do.
Thus far, the "factory-worker" examples have been discussed from a purely epistemological point of view, but it is worth asking whether those differences between wholes and aggregates are linked to differences of an ontological sort. Does the fact that it is possible to infer the "object-producing" function of the entire collective solely on the basis of information about the "object-producing" functions of the parts in the former—but not the latter—of the two examples derive, for instance, from the limitations of human reason; or is it because the latter collective, or the parts of that collective, possess properties that the former collective, or its parts, do not possess? What sorts of properties could the latter collective or its parts possess that the former collective or its parts do not possess? The answer cannot, of course, be functional properties since both the collective itself and its parts possess functional properties in both cases. It would appear, however, that the parts of the latter collective do possess a species of functional property that the parts of the former collective do not possess, viz., relational functional properties. The reason that it is not possible to infer the functional properties of the whole collective in the latter case is that the parts of the collective in question contract functional relations with each other; the number of object-portions that any given worker will produce
when part of the assembly-line will depend upon the number being produced by the workers on either side of him. Alteration of the number of object-portions being produced by any given worker in the line might even end up exerting a ripple effect, thereby altering the number being produced by everyone in the collective. In sharp contrast, the reason that it is possible to infer the number of objects produced by the entire collective in the former case is that the individual workers do not contract functional relations with each other. The number of objects being produced by a given worker is in no way dependent upon the number being produced by any other worker, and alteration of the number being produced by any one worker will have absolutely no effect on the number being produced by any other worker.

The foregoing paragraph suggests that it may be possible to formulate an ontological characterization of the distinction between wholes and aggregates. The distinction may thus be characterized in the way suggested several pages ago, or it may be characterized as follows. A collection of items or parts is an aggregate vis a vis a specified functional property or set of such properties if and only if possession of those properties does not result from

---

22 See above, p. *.
functional interdependencies contracted between the items or parts in question. A collection of items or parts is a whole vis a vis a specified functional property or set of such properties if and only if possession of those properties does result from functional interdependencies contracted between the items or parts in question.

The preceding discussion has succeeded in explaining one characteristic possessed by all systems: all systems possess at least some functional properties that result from functional interdependencies contracted between their parts, functional properties that cannot, therefore, be inferred or explained solely on the basis of information about the properties that their parts exhibit in isolation. That is what is meant by saying that systems are wholes rather than aggregates.

The second property that general systems theorists attribute to all systems is a capacity for dynamic activity. In his discussion of the subject, for example, Laszlo attributes two such properties to systems: (1) what he calls the property of "adaptive self-stabilization," and (2) what he calls "adaptive self-organization."\(^23\) The first of the

\(^{23}\) Laszlo, *Systems Philosophy*, pp. 40-41. For the sake of exegetical simplicity, the above discussion will ignore certain complications that arise when the properties of closed as well as open systems are discussed--see von Bertalanffy, pp. 39-41, 124-26.
The aforementioned properties refers to a system's capacity to reestablish a previously realized organizational pattern in response to the effects of external forces in its environment. The second of the aforementioned properties refers to a system's capacity to realize a new organizational pattern in response to the effects of such forces.

The third property that general systems theorists discuss in connection with the concept of "system" is that of hierarchical organization. Laszlo defines a hierarchy as a "multi-holon" type structure in which the systems functioning as wholes on one level function as parts on the higher levels, and where the parts of a system on any level (with the exception of the lowest or "basic" level) are themselves wholes on lower levels. Systems belonging to a level below that of any chosen level are called "subsystems" in relation to the system of the chosen level, and the system belonging to the next higher level is a "suprasystem" in relation to it [i.e., the chosen level].

According to Laszlo, a given system and all its sub-systems constitute an "intra-systemic hierarchy"; a given system and all the other systems with which it combines to form a

---

24 Laszlo, Systems Philosophy, pp. 39-41.
25 Laszlo, Systems Philosophy, pp. 41-47.
26 Laszlo, Systems Philosophy, p. 51.
supra-system constitute an "inter-systemic hierarchy."\textsuperscript{27}

Unfortunately, the foregoing quotations do not appear to be entirely consistent with Laszlo's other views. If there were a hierarchy whose lowest level consisted of "basic," indivisible units, a possibility left open by the preceding quotation, and if "systems belonging to a level below that of any chosen level are called 'subsystems,'" which presumably means that they also count as systems; then the basic units of the aforementioned hierarchy would automatically qualify as systems--but since such units would not be divisible into parts, they would not constitute wholes--and Laszlo himself considers wholeness to be a necessary condition of all systems.

In order to avoid the foregoing problematic feature of Laszlo's account, the present study will consider not only wholeness and a capacity for dynamic activity--but also hierarchical organization--to be necessary conditions of all systems; and hierarchical organization will be characterized as follows: a hierarchy is a whole at least some of whose parts are themselves wholes. Thus, an entity will not be classified as a hierarchy, and will, therefore, not be classified as a system either, unless it exhibits \textit{at least} three

\textsuperscript{27} Laszlo, \textit{Systems Philosophy}, p. 52.
distinct levels. If there were a hierarchy whose bottom level units were indivisible, neither those bottom level units, nor the units on the level immediately above them would, therefore, count as systems; only units on the third and higher levels would count as systems. The present view thus leaves open the possibility that there may be wholes that do not count as systems; if there are wholes that consist of only two levels, they will not count as systems.

Since a whole possesses functional properties that none of its parts possess in isolation, and since a hierarchy constitutes a series of vertically arranged, increasingly inclusive wholes, the possibilities for diversity of function will tend to increase as each new level of a hierarchy is reached. Thus, as Laszlo observes:

The higher we raise our sights on [a] hierarchy, the more diversity of functions and properties we are likely to find, manifested by a smaller number of actualized systems. . . . many systems on one level constitute one system on a higher level, consequently higher level systems are less abundant and have a wider repertory of functional properties than systems on lower levels.  

Thus, an inverse relationship between the number of systems on a given level and the number of properties possessed by

28 Laszlo, Systems Philosophy, p. 49.
each of the systems on that level may be exhibited as successive levels of a hierarchy are considered; as the number of systems decreases, the number of functional properties possessed by each system increases—as the number of systems increases, the number of functional properties possesses by each system decreases. 29

Not only does Laszlo believe that new functional properties are exhibited as each new level of a hierarchy is reached, he also believes that natural systems are able to develop a wide range of complex functional properties more quickly if they join together to form a hierarchy than they could on their own:

The reason for the decreased time-span required for a hierarchical form of complexification is that any failure in the organization will not destroy the system as a whole, but only decompose it to the next stable subsystem assembly. Then, instead of starting all over again, the process of complexification starts from the stable hierarchy-level and reconstitutes the loss in a relatively short time. Hence the observed predominance of hierarchies among natural systems are explained . . . by the fact that hierarchical systems are the ones that have time to evolve.

If, in favorable environments of enduring order, natural systems not only

29 Although Laszlo believes that all hierarchies will exhibit the aforementioned properties, the present study will leave open the possibility that there may be exceptions.
maintain their already attained level of organization, but intermittently or continuously evolve toward more highly organized states, then, given the relative rapidity of hierarchical over non-hierarchical organization, systems sharing a common environment will tend to form hierarchical suprasystems, rather than set forth their self-organizing processes on their individual level.  

Thus, according to Laszlo, hierarchical organization is displayed by virtually all natural phenomena because such organization enables entities to develop a broad range of adaptive strategies more quickly than they could on their own; and the more an entity is able to adapt to its environment, the more likely it is to survive.

The foregoing discussion of general systems theorists' work should now allow the concept of "system" to be characterized in terms of a set of conditions that are disjointly necessary and conjointly sufficient: (1) a system exhibits at least some functional properties that result from functional interdependencies between its parts, and which cannot, therefore, be inferred or explained solely on the basis of information about the properties that its parts exhibit when isolated from each other, i.e., all systems are wholes, (2) a system exhibits a capacity for dynamic activity, and (3) at least some of a system's parts constitute

---

30 Laszlo, Systems Philosophy, p. 48.
wholes in relation to the parts out of which they are made, i.e., all systems are hierarchically organized. Attention may now be turned to the business of characterizing the term "Structure."

THE NATURE OF STRUCTURES

While previous writers have not disagreed about whether the term "Structure" should be associated with the properties of wholeness and hierarchical organization, they have disagreed about whether it should also be associated with the capacity for dynamic activity. For example, Piaget asserts that a "basic property of structures [Structures] is . . . that they are self-regulating, self-regulation entailing self-maintenance," while Anthony Wilden asserts that

a structure [Structure] is either a static framework which is predominant at low levels of organization (whereas at all higher levels, system is predominant) or it refers simply to the given code of a system and not to its evolutionary capacities. 31

The question to be decided is not, of course, whether Piaget

or Wilden is right—since they are both using the word as a technical term within the scope of their own work; the question to be decided is whether or not it will be useful—for purposes of the present work—to associate both the term "system" and the term "Structure" with entities that possess a capacity for dynamic activity. If all hierarchically organized wholes turned out to possess a capacity for dynamic activity, there would naturally be little reason for distinguishing the two terms from each other—but that does not appear to be the case. While paintings, for example, arguably possess both the properties of wholeness and hierarchical arrangement, they obviously do not possess a capacity for dynamic activity; once a painting is damaged by external forces, it cannot—by itself—reestablish its previous organizational pattern or actively reorganize itself. Since it will be useful to have a term for such non-dynamic wholes, and since there is already a perfectly good term for dynamic wholes, viz., "system," the term "Structure" will hereafter be used to refer exclusively to non-dynamic wholes. Thus, this study will distinguish between two distinct types of hierarchically organized, holistic entities—systems, which possess a capacity for dynamic activity, and Structures, which do not possess a capacity for dynamic activity.
Since Structures do not possess a capacity for dynamic activity, they obviously cannot come to possess functional properties on their own, so to speak, but come to possess such properties, instead, because dynamic entities, like human beings, use them to perform various functions. Moreover, since evolution is limited to entities that possess an internal source of dynamism, and since Structures do not, by definition, possess an internal source of dynamism, it will not be possible to explain the fact that Structures possess such characteristics in the same way that Laszlo explains the predominance of such features among natural systems. Explanation of the fact that Structures exhibit the properties of wholeness and hierarchical arrangement can only be provided by a theory of construction--one that assumes that certain complex natural systems, like human beings, are capable of building such properties into the objects they make. Perhaps human beings are also capable of building a capacity for dynamic activity into some of their constructions, e.g., some people might wish to argue that sophisticated computers constitute systems in the precise sense defined above. If there are such things as artificial systems, it will also be necessary to invoke a theory of construction to explain the fact that they possess the properties they do.
THE ANALOGY BETWEEN LITERATURE AND LANGUAGE

As earlier remarks have indicated, the Structuralist approach advocates the use of linguistic concepts and techniques in the study of not only literary works, but in the study of all types of cultural phenomena—and it may be useful to begin the following discussion by examining the reasons that have been adduced in support of using such concepts and techniques in the study of cultural phenomena in general. Following Saussure, Culler suggests that the case for using linguistic concepts and techniques in the study of such phenomena is based largely on the realization that just as the utterances of a language are not simply strings of sounds or marks on paper but sounds or marks with a meaning, "social and cultural phenomena are not simply material objects or events but objects or events with a meaning"—and that when

human actions or productions have a meaning there must be an underlying system of distinctions and conventions which makes this meaning possible. Confronted with a marriage ceremony or a game of football, for example, an observer from a culture where these did not exist could present an objective description of the actions which took place, but he would be unable to grasp their meaning and so would not be treating

32 See above, pp. 2, 9-10.
them as social or cultural phenomena. The actions are meaningful only with respect to a set of institutional conventions.\textsuperscript{33}

In a similar passage, Scholes remarks that

In language, of course, no utterance is intelligible to a speaker who lacks the language-system that governs its meaning. The implications of this for literature are striking. No literary utterance, no "work" of literature, can be meaningful if we lack a sense of the literary system into which it fits.\textsuperscript{34}

Thus, just as a person must learn a whole set of conventions in order to understand the utterances of an unfamiliar language, Structuralists maintain that a person must also learn a variety of conventions in order to understand an unfamiliar type of literary work—and there is a fair amount of evidence that appears to support such a claim. When critics first encountered plays written in the absurdist style, for example, most of them were thoroughly baffled. They

\textsuperscript{33} Culler, \textit{Structuralist Poetics}, pp. 4-5. The Saussurean passage in question states that "every means of expression used in society is based, in principle, on collective behavior or—what amounts to the same thing—in convention" (\textit{Course}, p. 68); see also pp. 16-17 of the \textit{Course}. As subsequent remarks will show, Saussure did not believe that the members of a given community have to be aware of the rules that underly regularities in behavior in order for them to count as conventions; see below, fn. # 39, p. 97. The present study will follow Saussure's usage of the term "convention."

\textsuperscript{34} Scholes, pp. 14-15.
understood the individual sentences of the play, but they did not know how to make sense of the plays as a whole. After a number of years had passed, such plays no longer seemed so hard to understand. Why had the initial difficulties disappeared? The most plausible answer would seem to be that critics had, over the course of the intervening years, acquired a specific set of conventions that are needed in order to understand such plays.

If all cultural phenomena derive their meaning from the existence of communally shared conventions, however, why should any one type of cultural phenomena—namely linguistic phenomena—be singled out as a heuristic exemplar to be used in the study of all the others? Saussure posited the following answer:

Signs that are wholly arbitrary realize better than the others the ideal of the semiological process; that is why language, the most complex and universal of all systems of expression, is also the most characteristic; in this sense linguistics can become the master-pattern for all branches of semiology although language is only one particular semiological system.

The preceding passage actually identifies two reasons why


36 Saussure, p. 68.
linguistic phenomena might be considered an appropriate heuristic exemplar. First, since the conventional character of semiotic phenomena is more apparent in language than in other types of cultural phenomena, use of it as an exemplar may, as Culler puts it, help scholars avoid the familiar mistake of assuming that signs which appear natural to those who use them have an intrinsic meaning and require no explanation. Linguistics, designed to study the system of rules underlying speech, will by its very nature compel the analyst to attend to the conventional basis of the phenomena he is studying.\textsuperscript{37}

Second, since language constitutes the most complex form of semiotic phenomena, it is the one most likely to contain analogues for all the features found in other sign-systems. There is, given the present state of research in the various disciplines that study cultural phenomena, one additional reason for utilizing linguistic concepts and techniques: linguistics has made far more progress in its study of semiotic phenomena than any other discipline.

To say that both literary works and linguistic utterances possess meaning and that the meaning of both types of objects is determined by communally shared conventions is, of course, to identify similarities of an exceedingly broad nature, and it is somewhat questionable whether a heavily

\textsuperscript{37} Culler, \textit{Structuralist Poetics}, pp. 5-6.
linguistic approach to literary studies could be successful­ly defended solely on the basis of such generalities. What would be needed to strengthen the case for the Structuralist approach to literary studies would be evidence that the meaning of literary works and linguistic utterances is not merely determined by conventions, but that they are determined by similar sorts of conventions—or that there are other similarities between literary and linguistic phenomena.

The following discussion will attempt to strengthen the case for literary Structuralism in both of the preceding ways. In particular, subsequent discussion will argue (1) that literary works and linguistic utterances are structured in significantly similar ways—specifically, that they both fit the characterization of Structures advanced earlier in this chapter, (2) that the preceding structural similarities suggest that the meanings of both types of objects are determined by similar sorts of conventions, and (3) that the foregoing similarities suggest that both types of objects are comprehended in roughly similar ways. In drawing an explicit connection between the case for the literature/

38 Since the primary purpose of the following remarks is to indicate some additional reasons why the analogy between literature and language might be considered appropriate, the ensuing discussion of linguistic phenomena will not attempt to do justice to all of the complexities and fine distinctions involved. Subsequent chapters will attempt to refine this somewhat rough-and-ready picture of language.
language analogy and the fact that both phenomena constitute structures, the following discussion will suggest that the analogy can be supported in ways that previous writers have not explicitly noted.

Modern linguistics has shown that the sentences of natural languages can be analyzed into units that exhibit the following compositional organization: sentences are composed of clauses; clauses are, in turn made out of phrases; phrases are made out of words; words are made out of units called "morphemes"; and morphemes are, in turn, made out of units called "phonemes." The sentences of natural languages can thus be analyzed into parts, which can be analyzed into further parts, etc.

Insofar as the possession of identifiable meanings is concerned, the aforementioned types of linguistic units can be analyzed into parts, which can be analyzed into further parts, etc.


Units that belong to each of the aforementioned types are not, however, always composed of more than one unit on the next lowest level. In the sentence *John hit a ball*, for example, the noun phrase *John* is composed of only one word, and the word *a* is composed of only one morpheme, which is, in turn, composed of only one phoneme. It might be said that in such cases, one and the same sequence of graphic marks or sounds simultaneously constitutes a token of two or more types that are located on different hierarchical levels. Kenneth L. Pike refers to such cases as "portmanteau" phenomena—Kenneth L. Pike, *Language in Relation to a Unified Theory of the Structure of Human Behavior* (The Hague: Mouton & Co., 1976), p. 440.
be divided into three groups. Some units possess identifiable meanings even though their constituent parts do not, e.g., although the word *barber* has an identifiable meaning, neither the two morphemes that it contains nor the six phonemes which they contain possess identifiable meanings. Such units will be called "basic expressions." A second group of units do not even possess identifiable meanings themselves, e.g., the phonemes and morphemes in the word *barber*; they will be called "elementary units." The third group of units not only possess identifiable meanings themselves, but at least some of their parts also possess

40 References to the "meaning of subsentential linguistic units" during the following remarks are not intended to suggest that the meanings of such units do not derive from the role they play within the context of complete sentences; they are only meant to suggest that it is possible to talk about the meaning of a subsentential unit without referring to any particular sentence in which it occurs. While the meaning of the word *dog* may, for example, derive from the role that the word plays in all the English sentences that are used to make assertions about, ask questions about, or issue commands concerning a particular sort of animal—it is nevertheless possible to talk about the meaning of the word without referring to any specific one of those sentences. The following references to the "identifiable meanings of subsentential units" need not, therefore, run afoul of Gottlob Frege's so-called "context principle"—at least as it is construed by Michael Dummett; see Michael Dummett, *Frege: Philosophy of Language*, second edition (1973; rpt. Cambridge: Harvard University Press, 1981), pp. 3-7, 192-196.
identifiable meanings, e.g., not only does the phrase the tall barber possess an identifiable meaning, its constituent words also possess identifiable meanings. Such units will be called "derived expressions." All phonemes and many—if not most—morphemes will count as "elementary units"; at least some morphemes and many—if not most—words will count as "basic expressions"; and at least some words, all phrases, all clauses, and all sentences will count as "derived expressions."

A basic tenet in most recent attempts to investigate the nature of meaning has been the so-called "principle of compositionality," derived from the work of the philosopher and mathematician Gottlob Frege, which states that "The meaning of the whole is a function of the meaning of the parts and their mode of combination."41 One implication of the preceding principle is, of course, that a person could not infer the meaning of a derived expression solely on the basis of information about the meanings of its basic expressions; in order to do that, he would also have to know how the the basic expressions have been ordered or related to each other. A person could not, for example, infer the

meaning of a phrase solely on the basis of information that it contains the words the, red, brick, and house and information about the meanings of each of those words, for the meaning of the resulting phrase will depend on how the words are arranged; the phrases the red brick house and the brick red house obviously do not possess the same meaning. Much the same will hold true in cases that concern the arrangement of phrases in a clause or the arrangement of clauses in a compound or complex sentence.

The preceding paragraph has suggested that phrases, clauses, and sentences fit the characterization of wholes advanced earlier in this chapter; the parts of each of the aforementioned units contract functional relations with each other, and, as a result, it is not possible to infer the functional properties—the meaning—of a given unit solely on the basis of information about the meanings possessed by its parts. Although the principle of compositionality does not apply to basic expressions, since they do not contain parts that possess identifiable meanings, such expressions nevertheless also fit the characterization of wholes. While the individual morphemes contained in a word may not, for example, themselves possess identifiable meanings, the word

---

42 See above, pp. 42-43, 45-46.
does—and that meaning may depend on how the morphemes have been arranged. A person could not, for instance, infer the meaning of a word solely on the basis of knowledge that it contains the two morphemes /lo/ and /say/, for the order of the two morphemes can be rearranged to form two different words—loci and silo.43 Much the same will hold true in the case of morphemes that possess an identifiable meaning. A person could not, for example, infer the meaning of a morpheme solely on the basis of knowledge that it contains the two phonemes /I/ and /l/, for the order of the phonemes can be rearranged to form two different morphemes—the negative prefix /Il/, as in ill-advised, and the adjective-forming suffix /Il/, as in manly (as the word is pronounced in certain British dialects). Moreover, since the presence and ordering of even morphemes that do not possess identifiable meanings affect the identity of the words in which they are contained, even they can be said to possess some function, and since the presence and ordering of phonemes will affect the identity of such morphemes, such morphemes can also be said to constitute wholes; since it is not possible to infer the functional properties of such morphemes solely on the

43 Whenever there is a need to symbolize particular phonemes, the present study will utilize the symbols employed by Bolinger, see pp. 44, 79.
basis of information about the phonemes that they contain, they too can be said to constitute wholes.

Since sentences constitute wholes at least some of whose parts constitute wholes, and since they do not possess a capacity for dynamic activity, they fit the description of Structures advanced earlier. It is, therefore, possible to conceive of a sentence as consisting of a series of hierarchical levels each of which corresponds to one of the six types of linguistic units mentioned above. At the "bottom" of the hierarchy will be the phonemic level, the morphemic level will come next, then the "logemic" level, the phrasal level, the clausal level, and finally the sentential level.

Thus far, the principle of compositionality has been treated as an observation about the structure of derived expressions, but many writers believe that the principle possesses further implications. Michael Dummett, for example, maintains that the principle suggests that we understand [a] sentence—grasp its sense—by knowing the senses of the constituents, and, as it were, compounding them in a way that is determined by the manner in which the words themselves are put together to form the sentence. We thus derive our knowledge of the sense of any given sentence from our

44 See above, p. 53.
previous knowledge of the senses of the words that compose it, together with our observation of the way in which they are combined in that sentence.

Thus, the principle not only constitutes an observation about the structure of derived expressions—it may also be construed as an observation about the way in which people come to understand such expressions.

While a person's understanding of a derived expression appears to be based on knowledge concerning the meanings of its basic expressions and the way in which those expressions are arranged, that is obviously not the case where basic expressions are concerned—the parts of such expressions do not even possess identifiable meanings. The preceding difference would appear to suggest that there are at least two different types of conventions involved in the comprehension of linguistic utterances. One type of convention associates basic expressions with identifiable meanings, e.g., there is a convention in English that associates the word barber with people who are involved in a particular profession. There would, in addition, also appear to be conventions that determine how the meanings of derived expressions are to be computed on the basis of the meanings of basic expressions.

---

45 Dummett, p. 4.
and their arrangements. Some authors refer to the set of all the former conventions as the "code" of given language, and they refer to the set of all the latter conventions as the "system" of that language. Subsequent remarks will associate the former conventions with the term "code," but the term "system" will, instead, be used to refer to all of the conventions that are associated with the use of a language, both those conventions that belong to the code and those that belong to what will be called the "combinatorial" component of a language system. The two types of conventions will be referred to as "code-conventions" and "combinatorial conventions."

Before the literary applications of the concept of "Structure" are discussed, it will be useful to distinguish between two quite different ways in which the concepts and techniques of linguistics can be used in the analysis of literary works—what Phillip Pettit calls the "homeomorphic" and the "paramorphic" uses of linguistic work. Since literary works are themselves composed of language, the concepts and techniques of linguistics can, of course, be used to


analyze such works into the very same sorts of units identified by linguists, viz., phonemes, morphemes, words, phrases, clauses, and sentences; that is what Pettit refers to as the "homeomorphic" use of linguistic work.\(^{48}\) As Pettit points out, the homeomorphic use of linguistic concepts and techniques is hardly peculiar to the Structuralist approach; homeomorphic procedures have been widely used since at least the advent of New Criticism in the 1920's and 30's and are a mainstay of so-called "stylistic" criticism.\(^{49}\)

Structuralists claim that linguistic concepts and techniques can, however, also be used to help identify other sorts of literary units, units that are, of course, made out of the aforementioned, strictly linguistic units but are not simply identical with them. In his "Introduction to the Structural Analysis of Narratives," for example, Roland Barthes speaks of "narrative" units that are, as he puts it, "vehicled" by linguistic units, but are, nevertheless substantially independent of linguistic units; they may indeed coincide with the latter, but occasionally, not systematically. [Such units] will be represented sometimes by units higher than the sentence (groups of sentences of varying lengths, up to the work

\(^{48}\) Pettit, p. 40.

\(^{49}\) Pettit, p. 40.
in its entirety) and sometimes by lower ones (syntagm, word and even, within the word, certain literary elements only).50

Although such narrative units are to be distinguished from the strictly linguistic units out of which they are made, Barthes nevertheless posits a homological relation between sentence and discourse insofar as it is likely that a similar formal organization orders [both of them] . . . . A discourse is a long 'sentence' (the units of which are not necessarily sentences) . . . . 51

Thus, Barthes suggests that it is possible to analyze a narrative into various types of units that are related to each other and to the work that contains them in much the same way that phonemes, morphemes, words, phrases, and clauses are related to each other and to the sentences that contain them. The foregoing application of linguistic work involves the "paramorphic" use of linguistic concepts and techniques.52 Thus, whereas homeomorphic procedures are employed in order to identify and study the relations


52 Pettit, pp. 42-43.
between the strictly linguistic units contained in a literary work, "paramorphic" procedures are employed in order to identify and study the relations between literary units that are analogous to, but not the same as, the strictly linguistic units out of which they are made. It is the paramorphic use of linguistic concepts and procedures that will be of major concern during not only the following remarks, but throughout the remainder of the present study.

Associated with the distinction between the homeomorphic and paramorphic uses of linguistic work is the fact that the concept of "Structure" is applicable to the domain of literary works on a variety of different levels. Whereas the homeomorphic use of linguistic work only implies that literary works contain Structures (by virtue of the fact that they contain sentences), the paramorphic use of such work suggests that literary works are themselves Structures. Just as a sentence constitutes a whole whose parts, in turn, constitute wholes in relation to their parts, etc., a literary work is, on the paramorphic hypothesis, a whole whose parts constitute wholes in relation to their parts, etc. A literary work is thus to be conceived of as a superstructure only some of whose substructures will be identical with strictly linguistic units.

Structuralists have, of course, hardly been the first critics to claim that literary works constitute
hierarchically organized wholes. Ever since Aristotle attributed the property of wholeness to tragedy and compared the organization of a tragedy to that of a living animal, the so-called "doctrine of organic unity" has been associated with the claim that just as the functional capacities of an animal derive from the interrelations between its constituent organs, whose functional capacities derive, in turn, from the interrelations between their constituent parts—so the various powers and capacities of a literary work derive from the interrelations between parts that, in turn, derive their powers and capacities from the interrelations between the parts they contain.\footnote{Aristotle, Poetics, trans. Kenneth A. Telford (South Bend, Indiana: Gateway Editions, Ltd., 1961), 7: 1450b21-8: 1451a36. Although the comparison between literary works and living organisms was first made by Plato in the Phaedrus, his contention that a whole cannot be other than the sum of its parts in the Theaetetus would seem to indicate he did not intend the comparison to suggest that literary works possess the characteristics attributed to wholes earlier in the present chapter; see Plato, Phaedrus, in The Collected Dialogues of Plato, ed. Edith Hamilton and Huntington Cairns, trans. R. Hackforth (1963; rpt. Princeton: Princeton University Press, 1982), 264D, and Theaetetus, in The Collected Dialogues, trans. F. M. Cornford, 204A-205E. However, in what appears to be a direct response to his teacher, Aristotle, in Book VII, Chapter 17 of the Metaphysics, flatly asserts that a whole is more than the sum of its parts or elements—a whole consists of its parts plus the form or organization given to those parts (Aristotle, Metaphysics, in The Basic Works of Aristotle, ed. Richard McKeon, trans. W. D. Ross [New York: Random House, 1941], VII: 17: 1041b1-33). In view of that passage and of what he says about the structure of organisms in his}
of Structuralist criticism lies in its use of the analogy between the structure of literary works and the structure of linguistic utterances instead of the traditional organismic analogy—but why, it might be asked, should the linguistic analogy be considered more appropriate than the organismic one?

One reason why the linguistic analogy seems more appropriate than the organismic one is that literary works do not possess a capacity for dynamic activity. Literary works are, therefore, more like linguistic utterances than like living organisms; whereas organisms constitute systems, both linguistic utterances and literary works constitute Structures. In addition, whereas both literary works and biological treatises, it does not seem unlikely that he intended his remarks in the *Poetics* to be construed as implying something very like the claim just associated with the doctrine of organic unity in the main text. For a survey of the origins of the doctrine of organic unity, see G. N. Giordano Orsini, *Organic Unity in Ancient and Later Poetics* (Carbondale, Illinois: Southern Illinois University Press, 1975). For a discussion of the organismic analogy that explicitly relates it to the aforementioned claim, see Roman Ingarden, *The Cognition of the Literary Work of Art*, trans. Ruth Ann Crowly and Kenneth R. Olson (Evanston: Northwestern University Press, 1973), pp. 74-78.

It might at first glance seem that the foregoing does not hold true in the case of a theatrical production; since the performers can alter their performance in response to audience reaction, it might be thought that a theatrical production constitutes a system. Notice, however, that the same could be said about the process of uttering a sentence; the speaker can adjust what he is about
linguistic utterances constitute essentially semiotic phenomena, organisms do not. Notice that the preceding fact not only provides a reason for thinking that the linguistic analogy is more appropriate than the organismic analogy, it provides a reason for preferring the linguistic analogy over an analogy between literary works and any non-semiotic phenomena, whether they be systems or Structures.

There are, moreover, additional reasons for thinking that the linguistic analogy is particularly appropriate. One additional reason is suggested by Barthes' claim that

however many [narrative] levels are proposed and whatever definition they are given, there can be no doubt that narrative is a hierarchy of instances. To understand a narrative is not merely to follow the unfolding of the story, it is also to recognize its construction in 'storeys'. . . . to read (to listen to) a narrative is not merely to move from one word to the next, it is also to move from one level to the next. 55

Thus, just as a person must not only consider the meanings
to say depending on the listener's reactions. Nevertheless, the utterance qua product—the sentence uttered—is obviously not a system. Just as it is necessary to differentiate process from product in the case of language, so it is in the case of theatre; and once that distinction has been made, it becomes clear that theatrical productions also constitute Structures. More will be said about the relation between written drama and theatrical performance below, see pp. 115-18.

of the basic expressions in a linguistic utterance, but must also consider how their arrangements give rise to the meanings of higher-level, derived expressions; so a person must not only consider the meanings of the individual words or sentences in a narrative, he must also consider how their arrangements give rise to the meanings of larger narrative units. Moreover, just as the rearrangement of basic expressions may alter the meaning of a sentence, a clause, and even a phrase, the rearrangement of sentences in a literary work might well alter the meaning of higher-level literary units, including the meaning of the work itself. If that is, in fact, the case, it would appear to provide at least some reason for thinking that the meanings of higher-level literary units are related to the meanings of their parts in accordance with something similar to the principle of compositionality and that at least some of the conventions associated with the comprehension of literary works are combinatorial in character. Thus, the similarities between literary works and linguistic utterances may well extend beyond the mere fact that they are both essentially semiotic phenomena whose meanings are determined in accordance with communally shared conventions; they may also be structured in a similar way, their meanings may be determined in accordance with similar types of conventions, and
there may be similarities in the way they are comprehended.

The primary purpose of these remarks has not been to prove that any of the preceding claims are, in fact, true, but rather to provide a rather rough overview of some of the primary reasons that the linguistic analogy might be considered appropriate. As a result, the reasons have been presented in fairly programmatic and speculative terms. The next two chapters will not only expand upon these claims and provide additional support for them, but will also point out additional similarities between literary and linguistic phenomena. The usefulness of the analogy between literature and language does not, however, lie exclusively in the similarities between them; there is also a good deal to be learned about the nature of literary phenomena by noting the ways in which literature and language are dissimilar, and subsequent chapters will attend to those aspects of the analogy as well.

MORE ON HIERARCHIES

When an investigator attempts to study a particular set of phenomena by isolating them from their surrounding context, the resulting analysis can be said to "assume closure." Assumption of complete closure between the levels of a hierarchy would allow an investigator to study the
"horizontal" interrelations between the units of one level in isolation from the "vertical" interrelations between those units and the units of higher and lower levels. So doing has obvious disadvantages, for, as M. D. Mesarović and D. Macko observe:

Starting from any given [level], understanding of a system increases by crossing [levels]: moving down the hierarchy one obtains a more detailed explanation while moving up the hierarchy one obtains a deeper understanding of its significance.  

Assumption of total closure is, in fact, the very sort of procedure that, as earlier remarks indicated, Structuralists and general systems theorists denounce as "atomistic," "reductionistic," etc.  

If, however, a complete absence of closure is assumed, epistemological problems result. If it is assumed that a person must understand all of the interconnections between the units of a given level and all of the sub-units out of which they can be composed and the interconnections between those units and all of the supra-units into which they can


57 See above, pp. 22-26.
be incorporated before the horizontal interconnections be-
tween the units of that level can be understood, a person
will be in the position of having to know everything before
he can know anything and will never be able to get started.

Since assumption of complete closure results in signi-
cificant losses of insight and assumption of zero closure
results in an epistemological crisis, relative closure must
be assumed. Assumption of relative closure would allow an
investigator to study the horizontal interrelations between
the units of a given level in isolation from the vertical
interrelations between those units and the units of higher
and lower levels—provided that: (1) the investigator
realizes that the entity itself does not exhibit complete
closure—he is temporarily assuming that closure exists for
purely methodological reasons, and (2) the investigator puts
the level being investigated back into perspective—he must
realize that even after he has finished examining the hori-
zontal relations between the units, the vertical interrela-
tions between those units and the units of higher and lower
levels still remain to be studied.

Not only does the principle of relative closure allow
an investigator to isolate temporarily the units associated
with a single level from the units located on adjacent
levels, it also provides for the isolation of what might be
called "level-groups." Isolation of a particular level-group will allow an investigator temporarily to ignore the fact that the units located on the lowest level of that group are, in fact, made out of still smaller units; it will, in other words, allow him temporarily to treat those units as if they were indivisible. It will also allow him to treat the units located on the highest level of the chosen level-group as if they were ultimate. Thus, for example, having isolated the level-group that extends from the phonemic to the sentential level, a linguist may temporarily treat phonemes as if they were basic, indivisible units— even though acoustic phonetics might show that phonemes actually constitute wholes in relation to still smaller parts— and he may treat sentences as if they were ultimate units— even though discourse analysis may prove that individual sentences actually combine to form units on still higher levels.

The structural analysis of literary works will presumably also proceed by isolating a particular level-group. In his analysis of narrative structure, for example, Barthes assumes that "Just as linguistics stops at the sentence, so narrative analysis stops at [the boundaries of a] discourse"— even though narratives may, in turn, constitute parts of larger units of meaningful behavior:
Narration can only receive its meaning from the world which makes use of it: beyond the narrational level begins the world, other systems (social, economic, ideological) whose terms are no longer simply narratives but elements of a different substance (historical facts, determinations, behaviors, etc.).

Likewise, the structural analysis of literary works will no doubt end up temporarily treating as basic and indivisible units that are, from the standpoint of some other discipline, complex wholes composed of still smaller parts.

---

CHAPTER III
LINGUISTICS AND LITERATURE: THE TAXONOMIC APPROACH

As Pettit observes, "The linguistics which provides structuralism with its model is, in the broadest sense, structural linguistics. In this sense, structural linguistics includes most twentieth-century theoretical linguistics."¹ The era of "structural" or "Structuralist" linguistics began when Saussure revolted against what had by then become the accepted approach in linguistics. During the nineteenth-century, linguistic work was dominated by the so-called "comparative-historical" method; linguists were primarily interested in comparing historically distinct phases of one or more languages as regards a limited number of specific attributes or features of those languages.² The Comparativists and Neo-Grammarians of the nineteenth-century thus tended to focus on isolated portions of the languages

¹ Pettit, p. 1.
they studied. Saussure argued that they were, as a result, missing the most important features of language; he believed that at any given point in its history a language constitutes a systematic, integrated whole and that linguistic facts can be understood only when they are related to the overall organization of the language with which they are associated.\footnote{Saussure, pp. 3-15. Since Saussure did not actually write the Course himself, there will always be some doubt about how accurately it reflects his thought—see above, p. 1. For the sake of convenience, the present study will, however, simply assume that it is accurate.} The structural or Structuralist approach to language is largely based on acceptance of those claims: "To envisage a language ... as a system organized by a structure to be revealed and described," says Emile Benveniste, "is to adopt the 'structuralist' point of view."\footnote{Emile Benveniste, Problems in General Linguistics, trans. Mary Elizabeth Meek (Coral Gables, Florida: University of Miami Press, 1971), p. 82.}

While most of the linguistic theories that have been formulated during the twentieth-century are similar enough to be classified under the broad rubric of structural linguistics, they also exhibit substantial dissimilarities. As Manfred Bierwisch observes, from the time of Saussure up until the 1950's, description of a linguistic utterance's structure typically involved the segmentation of the utterance into units that were then classified on the basis of
the relations they contract with each other and with other units of the same language. The structure of a language-system was accordingly conceived of as a set of classes of units that exhibit complex interrelations with each other. Bierwisch refers to that approach "by the now customary term 'taxonomic linguistics.'" He goes on to point out that taxonomic linguistics gave way to a new approach during the 1950's: "When [Noam] Chomsky published his Syntactic Structures in 1957, structural linguistics entered a new phase"--the phase of so-called "generative linguistics." A generative linguist's description of a linguistic utterance's structure does more than just segment the utterance into units and classify them, it also attempts to identify a process capable of generating that utterance. The structure of a language-system is accordingly conceived of as more than just a set of classes of units that exhibit interrelations; the system is also to contain a set of rules that specify the ways in which the units of particular classes can be combined in order to generate well-formed sentences.

---


6 Bierwisch, p. 21.

7 Bierwisch, p. 45.

8 Bierwisch, pp. 45-6.
While virtually all Structuralist critics agree that the concepts and techniques of taxonomic linguistics are applicable to the domain of literature, the applicability of generative concepts and techniques has constituted a source of considerable controversy. Because discussion of those ideas about which literary Structuralists generally agree will tend to be less complicated than discussion of those about which they disagree—and because development of the former ideas chronologically preceded development of the latter (in fact, the latter ideas probably could not have been developed without the foundation that the former ideas provided)—examination of the concepts and techniques developed by taxonomic linguists will precede examination of those employed by generative linguists. As the contributions to both approaches have been far too numerous to allow any sort of comprehensive survey of them within the present study, discussion of them will focus on some of the principal ideas developed by the linguists who pioneered them—Saussure in the case of taxonomic linguistics and Chomsky in the case of generative linguistics.\(^9\) The discussion will

not, however, be entirely limited to their work. Other linguists have attempted to amplify and refine some of Saussure and Chomsky's main ideas, and sporadic references to such other authors will, therefore, be made. Reference to the work of other writers will, in addition, facilitate attempts to identify the problematic aspects of the major concepts developed by Saussure and Chomsky. As remarks contained in Chapter One of this study have already indicated, the following discussions will not only attempt to explicate the theories advanced by Saussure and Chomsky, but will also criticize and extend those theories.10

After each of the major concepts has been critically examined from the linguistic standpoint, its various literary applications will be discussed. Although such discussions will frequently refer to ways in which previous critics have applied the concepts, they do not constitute an attempt to survey the work of literary Structuralists to date; references will be sporadic and anything but exhaustive.11 Besides referring to previously proposed


10 See above, pp. 16-18.

11 For surveys of the work of Structuralist critics, see Scholes, Jameson, and Culler, Structuralist Poetics.
applications, such discussions will also speculate about heretofore unexamined possibilities of application, and should, in so doing, identify a number of issues that will be pertinent to later discussions of the Poetics.

TAXONOMIC LINGUISTICS

As has already been mentioned, Saussure's major criticism of the Comparativists and Neo-Grammarians was that they tended to focus on isolated aspects of the languages they studied, thereby missing the systematic nature of language. Lévi-Strauss refers to such pre-Saussurean work as "atomistic" linguistics. Saussure did not, however, believe that linguistic studies had always been atomistic. On the contrary, he believed that traditional grammarians, like those who constructed the Port Royal Grammar, exhibited a "viewpoint [that] was absolutely above reproach"—insofar as it focused on the systematic nature of a particular language at a specific point in time. While Saussure believed that traditional grammar was superior to nineteenth-century linguistic work in that respect, he claimed

12 See above, pp. 80-81.
13 Lévi-Strauss, Structural Anthropology, pp. 34-35.
14 Saussure, p. 82.
that it was inferior in another respect; whereas nineteenth-century linguists attempted to describe linguistic usage, he believed that traditional grammarians had attempted to prescribe what they thought to be "proper" linguistic behavior. Thus, while the work of the traditional grammarians was non-atomistic, it was dangerously prescriptive. On the other hand, the work of the Comparativists and Neo-Grammarians was descriptive, but it was dangerously atomistic. Saussure believed both approaches were unsatisfactory. He believed that the study of language requires an approach that is both descriptive and non-atomistic—a belief that has remained a hallmark of structural linguistics. In order to create such a descriptive, Structuralist approach, Saussure developed a number of important concepts—concepts that have played a crucial role in shaping the course of modern linguistics.

**LANGAGE, LANGUE, AND PAROLE**

Saussure believed that many of the problems inherent in the work of previous linguists derived from the fact that they had never properly defined the object of linguistic

15 Saussure, p. 82.
According to Saussure, linguists should not take the totality of linguistic phenomena, which he referred to as "langage," as their object of study because it is too diverse and multi-faceted:

Taken as a whole, [langage] is many-sided and heterogeneous; straddling several areas simultaneously—physical, physiological, and psychological—it belongs both to the individual and to society; we cannot put it into any category of human facts, for we cannot discover its unity.

Saussure believed that the proper object of linguistic study could be isolated only by drawing a distinction within the domain of langage between what he referred to as "langue" and "parole."
According to Saussure, "**langue**" refers to "the whole set of linguistic habits which allow an individual to understand and to be understood."\(^{19}\) **Parole** refers to "an individual act"--either the act of expressing an utterance or the act of comprehending an utterance; Saussure also frequently used the term to refer to individual utterances themselves.\(^{20}\) Thus, as Culler observes in his book on Saussure:

> The distinction between **langue** and **parole** . . . is essentially a distinction between institution and event, between the underlying system which makes possible various types of behavior and actual instances of that behavior.\(^{21}\)

example, says that "**langage** is linguistic potential, **langue** is a language-system, and **parole** is individual utterance" (p. 14). Saussure is, however, quite explicit about the relation between the three terms: "within the total phenomenon represented by [**langage**] we first singled out two parts: [**langue**] and [**parole**]. [**Langue**] is [**langage**] less [**parole**]" (p. 77). Thus, when combined, **langue** and **parole** exhaust the domain of **langage**.

\(^{19}\) Saussure, p. 77.


Thus, "langue" refers to the system of rules or conventions that makes linguistic behavior possible, and "parole" refers to individual acts that exemplify or instantiate those rules or conventions.

In his discussions, Saussure repeatedly emphasizes two features of langue. First, he stresses the psychological nature of langue, insisting that it "can be localized in the limited segment of the speaking-circuit where an auditory image becomes associated with a concept."22 Second, he insists that "[Langue] is the social side of [langage], outside of the individual who can never create or modify it by himself."23

Because its scope is limited to those aspects of linguistic phenomena that are both psychological and social, Saussure believed that langue constitutes a much more appropriate object of linguistic study than either langage or parole. Whereas langage is heterogeneous, encompassing aspects of linguistic behavior that range from the psychological to the purely physical and physiological, langue encompasses only those aspects that are psychological.

22 Saussure, p. 14. When Saussure says that langue is a psychological phenomenon, he appears to have something like what is today called "cognitive psychology" in mind. More will be said about the matter below, see pp. 582-84.

Whereas parole is determined by a variety of factors that vary from one individual to the next and from one situation to the next, factors that are, from the standpoint of linguistics, accidental—like the mental and physical condition of a speaker or listener at a given time, the atmospheric conditions at the time of a given act of parole, etc.—the restriction of langue to the social aspects of linguistic phenomena excludes such accidental contingencies from its scope. Thus, unlike langage and parole, langue constitutes a self-contained whole and a principle of classification. As soon as we give [langue] first place among the facts of [langage], we introduce a natural order into a mass that lends itself to no other classification.

Thus, Saussure believed that langue provided precisely the sort of object that was needed in order to unify and organize linguistic study and insisted that "from the very outset we must put both feet on the ground of [langue] and use [langue] as the norm of all other manifestations of [langage]."

---

25 Saussure, p. 9.
26 Saussure, p. 9.
Unfortunately, there appear to be two basic problems inherent in Saussure's conception of *langue*. First, he thought that a fair amount of syntactic phenomena belonged to *parole* rather than *langue*. His reasons for thinking that appear to derive from two sources: (1) he maintained that *langue* properly includes only those aspects of *langage* that are based on communal convention rather than "the will of the individual," and (2) he seems to have taken the variety and complexity of syntactic phenomena as an indication that such phenomena, or at least a substantial subset of those phenomena, must derive from individual invention rather than communally shared conventions.

---

27 During the early sections of the *Course*, Saussure simply relegates to the domain of *parole* "the combinations by which the speaker uses the language code for expressing his own thought" (p. 14), or "individual combinations that depend on the will of the speakers" (p. 19)—without indicating that at least some syntactic phenomena also belong to *langue*. During a later section of the work, however, he indicates that some combinations do belong to *langue*: "pat phrases in which any change is prohibited by usage" (p. 124)—and even "sentences and groups of words built upon regular patterns" (p. 125). He then adds that with respect to the combinatorial aspects of linguistic phenomena "there is no clear-cut boundary between [*langue*], which is a sign of collective usage, and the fact that belongs to [*parole*] and depends on individual freedom. In a great number of instances it is hard to class a combination of units because both forces have combined in producing it, and they have combined in indeterminable proportions" (p. 125). It is, as a result, fundamentally unclear just how much of syntactic phenomena Saussure was willing to include as part of *langue*.

28 As evidenced by the remarks contained in the previous footnote.
While the work of generative linguists has substantiated Saussure's belief that syntactic phenomena are marked by an extraordinary amount of variety and complexity, it has shown that such phenomena can, nevertheless, be explained in terms of communally shared rules or conventions.\textsuperscript{29} Since Saussure would very likely have classified all syntactic phenomena under the rubric of \textit{langue} had he realized that its variety and complexity could be accounted for in terms of such rules, subsequent remarks will treat syntax as a part of \textit{langue}. Given Saussure's claim that \textit{langue} encompasses the whole set of linguistic habits or conventions that allow an individual to understand and to be understood, it seems perfectly appropriate to include syntactic phenomena; the combinatorial rules of syntax are, after all, part of what enables individuals to construct and decipher linguistic utterances.\textsuperscript{30}

During the last several decades, work in the philosophy of language has suggested that the set of linguistic habits or conventions that allow individuals to construct and decipher utterances are even more inclusive than the foregoing paragraphs have indicated. According to the so-called "theory of speech acts," developed primarily by the

\textsuperscript{29} See below, pp. 242–53.

\textsuperscript{30} See above, p. 88.
philosophers J. L. Austin and John R. Searle, the "illocutionary force" of an utterance—the fact that it constitutes a question rather than a request, or an assertion rather than a warning, etc.—also derives from the existence of certain communally shared rules or conventions. It therefore seems appropriate also to consider those conventions to be part of \textit{langue}, and, in fact, Searle explicitly claims that they are:

\begin{quote}
It . . . might seem that my approach is simply, in Saussurian terms, a study of "parole" rather than "langue." I am arguing, however, that an adequate study of speech acts is a study of \textit{langue}. 
\end{quote}

While Saussure has nothing to say about such matters and would probably have been as sceptical about the possibility of formulating rules to account for the various types of illocutionary force as he was about the possibility of formulating rules capable of accounting for the variety and complexity exhibited by syntactic constructions, classification of illocutionary conventions under the rubric of \textit{langue} nevertheless seems true to the spirit—if not the


\footnote{32 Searle, \textit{Speech Acts}, p. 17.}
letter—of Saussure's thought. The whole set of linguistic conventions that enable individuals to create and decipher utterances includes not only code conventions and combinatorial conventions, but also conventions of the sort discussed by Austin and Searle; the study of langue includes the study of all such conventions.

The second basic problem inherent in Saussure's conception of langue concerns its precise ontological status. During his discussion of the interrelation between langue and parole, he claims that "the former is both the instrument and product of the latter."33 A paragraph later, Saussure claims that langue "exists in each individual."34 However, if langue constitutes the instrument of parole—if it is identical with the body of knowledge in each individual which enables him to create and decipher utterances—it cannot also be equated with the social product of all individual's acts of parole: no individual will possess knowledge of every single convention mobilized by members of his linguistic community.

The problem is vividly illustrated by an analogy Saussure uses in his explication of langue. Langue, he says:

33 Saussure, p. 19.
34 Saussure, p. 19.
exists in the form of a sum of impressions deposited in the brain of each member of a community, almost like a dictionary of which identical copies have been distributed to each individual.\footnote{Saussure, p. 19.}

Several pages earlier, he claims, however, that \textit{langue} "is not complete in any speaker; it exists perfectly only within a collectivity."\footnote{Saussure, p. 14.} Since no individual in a given linguistic community possesses the complete, perfect copy of the "dictionary," \textit{langue} cannot, however, be identical with both the dictionary itself and that which each individual knows about the dictionary. Thus, \textit{langue} must be equated with either that which is in each individual or with the social product of which each individual has only partial knowledge. In view of Saussure's repeated emphasis on \textit{langue} as social product, the latter position seems preferable.

Even Saussure's conception of \textit{langue} as social product--as the sum of the various "dictionaries" possessed by the members of a given collectivity--cannot, however, be accepted as is. If each individual possesses a necessarily incomplete and flawed knowledge of \textit{langue}--a necessarily incomplete and flawed "dictionary"--the sum of all such "dictionaries" will constitute no more than a collection of

\footnote{Saussure, p. 19.}
\footnote{Saussure, p. 14.}
flawed works; but Saussure conceived of *langue* as being complete and perfect. The perfect—the complete and unflawed—"dictionary" must, therefore, be abstracted from from the entire collection. As Scholes observes:

> A language-system . . . has no tangible existence. The English language is not in the world any more than the laws of motion are in the world. In order to become an object of study a language, or a model of it, must be constructed from the evidence of individual utterances.  

Most generative linguists would add that the model of a language is based on not only the evidence furnished by individual utterances, but also by the judgments that native speakers make about such utterances—whether, for example, a given utterance is grammatically well-formed. However, the basic point remains the same: *langue* does not exist in a single individual or in a collectivity; it exists only as a model that has been abstracted from a collectivity.

Given the preceding view, an individual must be said to be capable of participating in acts of *parole* because he possesses, not *langue* itself, but a fairly extensive—a

---


fairly accurate—knowledge of *langue*. Chomsky has proposed a term that identifies the extensive, though incomplete, knowledge that is actually *in* each individual. He distinguishes "between *competence* (the speaker-hearer's knowledge of his language) and *performance* (the actual use of language in concrete situations)."  

---

39 Noam Chomsky, *Aspects of the Theory of Syntax* (Cambridge: The M. I. T. Press, 1965), p. 4. According to Chomsky, native speakers may be completely unaware of much of the knowledge that constitutes competence and enables them to create and decipher utterances: "As I am using the term, knowledge may be unconscious and not accessible to consciousness. It may be 'implicit' or 'tacit.' No amount of introspection could tell us that we know, or cognize, or use certain rules or principles of grammar . . . ."*--Rules and Representations* (New York: Columbia University Press, 1980), p. 128; see also, *Aspects*, p. 8. Like Chomsky, Saussure also insisted that "speakers are largely unconscious of the laws of [langue]" (*Course*, p. 72).

One additional implication of the foregoing remarks should be mentioned. A number of contemporary philosophers have argued that a particular sort of behavior cannot be considered a convention unless the participants know that it is a convention. For example, in his paper "Languages and Language," David Lewis argues that the fact that most people in a particular community conform to a certain regularity in behavior, along with a variety of other facts, must be "matters of *common* (or *mutual*) knowledge: they are known to everyone, it is known to everyone that they are known to everyone, and so on"; that, he claims, is a necessary condition for being a convention—*Philosophical Papers*, vol. 1 (New York: Oxford University Press, 1983), p. 165. On such an account, most rules of syntax could not be considered conventions, and, in fact, later in the same paper, Lewis argues that "there are other so-called rules of language which are not conventions of language and are not in the least like the conventions of language: for instance, 'rules' of syntax and semantics" (p. 180). It appears that Saussure must disagree. Since he believes that people are unaware of the rules of *langue*, and since the rules of
equivalent to parole, but competence is equivalent, not to langue, but to an individual's knowledge of langue. The present study will accordingly adopt the view that langue only constitutes the product of parole (more precisely, it constitutes an abstract model based on that product);

**langue** are identical with the body of conventions operative in a community, there must, in his view, be such things as "implicit" or "tacit" conventions. During a discussion of Lewis's position, Simon Blackburn has recently espoused a view similar to Saussure's, claiming that "it is correct to call things like word-order conventional even if their status is something the population does not understand. To analyze language we probably need the concept of a non-overt convention . . ." (Simon Blackburn, *Spreading the Word: Groundings in the Philosophy of Language* [Oxford: Clarendon Press, 1984], p. 122). As earlier remarks have indicated, the present study will adopt the Saussurean use of the term convention; see above, fn # 33, p. 56.

Actually, the concepts of performance and parole are not exactly the same. Whereas both terms are used to refer to actual occurrences of linguistic behavior—the composition and comprehension of particular utterance-signals—only the latter term can also be used to refer to the utterance-signals themselves.

As the foregoing pages have indicated, the study of linguistic phenomena requires the postulation of both the body of linguistic knowledge possessed by each individual and a model abstracted from the knowledge possessed by each individual. Chomsky associates the latter notion with the competence of "an ideal speaker-listener, in a completely homogeneous speech-community, who knows his language perfectly" (*Aspects*, p. 3). Chomsky's ideal speaker-listener is, in other words, a person who possesses a complete and perfect knowledge of langue. Thus, it is possible to draw the aforementioned distinction either in terms of a difference between the competence of actual individuals and langue as an abstract model or in terms of a difference between the competence of actual individuals and the competence possessed by an ideal speaker-hearer.
competence constitutes the instrument.

Because an individual's utterances may exhibit not only elements that are contained in langue, but also eccentricities not included in it—eccentricities that are not common to a sufficient portion of a given community and have not, therefore, been incorporated into the abstract model that constitutes langue—it is necessary to distinguish between linguistic competence and what many modern linguists refer to as "idiolect."\(^1\) Idiolect includes the knowledge that underlies all of a person's linguistic behavior—including eccentricities; competence only refers to the individual's knowledge of langue.

During the course of the foregoing paragraphs, langue has been primarily associated with the system of conventions shared by the speakers of a natural language like English or French. Between the idiolect of each individual and the system of conventions that underlie a natural language like English there are, however, a number of different dialects, and, it appears that each of those dialects is also based on a particular system of communally shared conventions. Thus, it would appear that the Sassurean concept of langue is also applicable in the case of dialects; whenever the linguistic

\(^1\) See, for example, Bolinger, p. 332.
behavior of a group of individuals derives from their knowledge of a system of shared conventions or rules, the concept of langue is applicable. The relations between the system of conventions that underlie a particular natural language and the systems that underlie the various dialects encompassed by that language are no doubt extremely complicated.

Having distinguished parole from langue and langage, Saussure went on to distinguish between two types of linguistic parole—spoken versus written—and insisted that the true object of linguistic study "is not both the written and the spoken form of words; the spoken forms alone constitute the object." Although the reasoning that Saussure used to justify his position seems somewhat unsound, modern linguists have, for methodological reasons, generally affirmed the priority of spoken parole. Because it


43 For a discussion of the viewpoint adopted by modern linguists, see Lyons, Theoretical, pp. 38-42. Saussure gave priority to the spoken form of parole because he believed that the written form "exists for the sole purpose of representing" it and is, therefore, more "artificial" than the "natural" spoken form (Course, pp. 23, 32). However, Saussure apparently contradicts his claim that the spoken form is more "natural" elsewhere, asserting instead that "what is natural to mankind is not oral speech but the faculty of constructing langue, i.e., a system of distinct signs [signifiers] corresponding to distinct ideas" (Course, p. 10).
exhibits linguistic factors that are absent from written speech, e.g., variations in pitch and stress, etc., linguists can learn more about the functioning of language by examining the spoken form.\(^{44}\)

Despite the fact that he considered spoken parole to be more important, Saussure did not, however, consider written language to be devoid of linguistic importance. Far from it, he drew attention to the fact that, because it is more durable than spoken parole, linguists are forced to rely heavily on the evidence provided by written language.\(^{45}\) He also believed that graphic signs aid the analysis of phonic ones—particularly in the early stages of analysis.\(^{46}\)

**LITERARY APPLICATION**

The literary relevance of the foregoing discussion concerns the literary applicability of the concepts of parole, langue, competence, and idiolect. Since the concept of parole seems to have the most obvious applications, discussion may profitably begin with it.

\(^{44}\) Lyons, *Theoretical*, p. 40.

\(^{45}\) Saussure, p. 23.

\(^{46}\) Saussure, p. 32.
Structuralists are unanimous in equating parole with individual pieces of literature.\(^\text{47}\) However, the fact that, in linguistics, "parole" refers to not only utterance-signals, but also to the processes of producing and perceiving utterance-signals indicates that, in literature, the processes of creating and comprehending literary works may also be referred to as instances of parole. The notion that comprehending a literary work constitutes an act of parole is especially important, for, without it, the number of people that could be said to be capable of literary parole would be extremely limited. One of the principle dissimilarities between the domains of language and literature is that the number of people capable of creating linguistic utterances is far greater than the number of people capable of creating literary works.

Literary application of the concept of idiolect will not be quite as simple as was the case with parole. In its most obvious application the term could be used to designate the literary system that underlies an individual author's works--including all of the idiomatic eccentricities they contain. There are, however, complications. Many authors

\(^{47}\) See, for example, Scholes, p. 14 and Jameson, p. 101.
have worked in more than one literary form, producing novels as well as plays or lyric poetry as well as short stories, etc. Moreover, authors frequently write in a variety of genres, e.g., comedy as well as tragedy, and a variety of styles, e.g., realism as well as expressionism. Should such an author be assigned a single idiolect that includes the knowledge required to produce all of the different sorts of works he has written, or should he be assigned a variety of idiolects?

Analogous complications arise in the case of language. As William G. Lycan observes, "an individual's speech habits change over time, so that a person speaks one idiolect at one time and another at another." Moreover, a person's speech habits may vary depending on his environment, e.g., a person may speak quite differently at the office than he does when he is out with friends, so that even at a given point in time it may be necessary to assign an individual several distinct idiolects.

The foregoing considerations indicate that it is appropriate to accord a person more than one idiolect in the case of language as well as literature. The exact criteria to be used in determining the identity conditions for

literary idiolects are far from clear. It would at least seem that it is inappropriate to ascribe a single underlying idiolect to works that differ as regards form, genre, or style, but there may be other factors that are relevant as well.

There is perhaps an additional way in which the term "idiolect" might be applied within the domain of literature. Individuals who are not capable of creating literary works might, nevertheless, be said to possess idiolects—insofar as they possess eccentric idioms or conceptions that influence their comprehension of literary works. It is not entirely clear whether it would be appropriate to use the term in the foregoing way, and although that use will not be ruled out, attention will be focused on the application of the term described in the preceding paragraph whenever the subject arises during the remainder of this study.

The notion that the concepts of *langue* and competence are applicable within the domain of literature is based on the following line of reasoning. Acquisition of the ability to understand the utterances of one natural language does not automatically enable a person to understand the utterances of other languages. That indicates that each language is associated with a particular set of conventions; a person's inability to comprehend the utterances of an unfamiliar language derives from the fact that he does not yet
know the particular conventions of that language. In order to understand the utterances of any language it is thus necessary to acquire a body of knowledge about a particular system of conventions; that body of knowledge constitutes competence, and the system of conventions constitute langue.

Likewise, acquisition of the ability to understand a certain class of literary works does not automatically enable a person to understand works of different sorts, as was indicated by the interpretive difficulties critics experienced when they first encountered absurdist dramas. It would appear then that each kind of literary work—each form, genre, and style—is likewise associated with its own special set of conventions; a person's inability to comprehend works of a particular kind indicates that he has not yet acquired the underlying set of conventions upon which works of that kind are based. Thus, it is necessary to acquire knowledge of a particular system of conventions in the case of literary works no less than in the case of linguistic utterances. It therefore seems appropriate to use the terms "competence" and "langue" with respect to the domains of both language and literature.

One of the questions raised by the analogy between linguistic and literary competence is whether it makes sense to speak of literary competence simpliciter or without qualification. When a linguist speaks of a person's linguistic
competence, he means competence with respect to a particular language or dialect; since it would be virtually impossible for a single person to acquire competence with respect to all of the world's different natural languages and dialects, it would not make sense to speak of linguistic competence simpliciter. Could a single person acquire competence with respect to all of literature's different forms, genres, and styles? While the present study will not rule out the foregoing possibility—unlikely as it seems—subsequent references to "literary competence" are to be construed as a generic way of referring to various, more limited competencies—not as references to some sort of all-encompassing competence.

The analogy between linguistic and literary competence also raises the following question: what sorts of behavior are to count as evidence that a given person possesses literary competence? Scholes remarks that

49 Scholes, p. 130.
works—to enjoy listening to or reading them—be taken as sufficient evidence that an individual, whether a child or an adult, possesses competence with respect to that class of works? A comparison of such a rudimentary capacity with the sorts of capacities associated with linguistic competence will reveal that it is not sufficient.

Among the various capacities that linguists typically take into consideration when discussing competence are the ability to paraphrase utterances, the ability to segment utterances into their primary structural constituents, and the ability to provide judgments concerning the well-formedness of utterances. The purported existence of linguistic competence is thus based on the existence of linguistic

50 See, for example, Chomsky, Aspects, pp. 18-19, 24, Baker, p. 5, and Lyons, Semantics, pp. 27-28, 389, 422. It should be added that Chomsky does not believe that the ability to perform the foregoing sorts of tasks constitutes the only possible evidence of a person's competence; since he regards linguistic competence as a body of knowledge that is realized in terms of mental state and structure, he insists that while "behavior . . . provides evidence for possession of knowledge, [so] might facts of an entirely different order—electrical activity of the brain, for example" (Rules, p. 48). During the course of the following remarks, it will be important to remember that "competence" does not refer to a person's linguistic capacities—but to the body of knowledge that underlies those capacities. To ask whether a person possesses competence is, thus, to ask whether he possesses a particular body of knowledge—not whether he possesses certain capacities—even though, at the present time, the former question could only be answered by obtaining an answer to the latter one.
capacities far more sophisticated than the mere ability to respond in a positive way to utterances, and the hypothesis that a literary counterpart exists apparently ought to be based on the existence of similarly sophisticated capacities.

Scholes is not, of course, unaware that the postulation of linguistic competence is based on the existence of more sophisticated capacities and decides, as a result, to reject the very possibility that a literary counterpart exists:

In linguistics it may be perfectly reasonable to rely on native informants for "competence" with respect to the grammar of their language. This works because "competence" is a minimal function. Everybody has it, or nearly everybody. In poetics, we cannot do this, because [literature] starts where mere "competence" leaves off.\footnote{Scholes, p. 39.}

Thus, Scholes rejects the idea that there is such a thing as literary competence apparently on the grounds that there are no literary counterparts for the more sophisticated capacities associated with linguistic competence—or that if such sophisticated literary capacities do exist, they are not possessed by a large enough number of people. Surely it is possible, however, at least to conceive of literary counterparts for the more sophisticated linguistic capacities described in the last paragraph—namely, the ability to
provide a rough synopsis or paraphrase of a literary work's plot, the ability to segment a literary work into its main structural constituents, and the ability to formulate judgments concerning the well-formedness of literary works.

When a linguist says that a person possesses the capacities mentioned earlier, he does not, however, mean that the person is able to provide just any judgments concerning the paraphrase, structural organization, and well-formedness of a given utterance; he means that the person is able to provide judgments about the utterance that are reliable or acceptable in the sense that they approximate a norm based on what other native speakers of the language would say about the same utterance. The linguist is, in other words, talking about what might be called "normative" as opposed to "non-normative" capacities. Thus, the case for postulating the existence of linguistic competence proceeds in several steps. Since there is widespread intersubjective agreement about matters of paraphrase, structural organization, and well-formedness, the linguist can postulate the existence of normative capacities; competence constitutes the internalized body of knowledge that presumably underlies those normative capacities.

---

52 See, for example, Baker, p. 11.
It would appear that a similar case will have to be made in order to establish the plausibility of postulating some sort of literary competence: the plausibility of postulating literary competence will depend on the plausibility of postulating the existence of certain normative capacities, which will, in turn, depend on the plausibility of the claim that there is at least a fair amount of intersubjective agreement among critics about the plot-synopsis, structural organization, and well-formedness of literary works. The notion that there is such intersubjective agreement among critics and that it is legitimate to speak of reliable or acceptable judgments about literary works is, of course, extremely controversial. Discussion of such matters will constitute a major portion of the next chapter of this study.\(^{53}\) There are, however, two points that should be mentioned now. First, while the plausibility of postulating the existence of some sort of literary competence will depend on whether there is intersubjective agreement among at least some individuals, it does not depend on how large that group of individuals is. If it turns out that intersubjective agreement occurs only among those who have made a career out of studying literature— if the number of people who possess literary competence turns out to constitute no

\(^{53}\) See below, pp. 283-305, 309-17.
more than a small percentage of those who enjoy literary works, whereas nearly everyone who uses a natural language possesses linguistic competence—that would not, as Scholes seems to suggest, count as evidence that there is no such thing as literary competence. Second, there is an obvious explanation for such a discrepancy between the number of people who might be said to possess the respective competencies.

As linguists often point out, it takes a baby literally years of intense effort to attain mastery of his first natural language. It does not seem unreasonable to suppose that acquisition of literary competence would also require a good deal of time and effort. Whereas acquisition of linguistic competence is vital to an individual's ability to function as a normal social being, acquisition of the latter is, however, far from vital to most individuals. As a consequence, a relatively small number of people are willing to spend the time and energy required in order to attain the latter. Indeed, the fact that institutions of higher education around the world contain departments whose primary enterprise is the study of literature would seem to constitute at least prima facie evidence that there is such

---

54 See, for example, Bolinger, p. 2.
a thing as literary competence; such departments exist precisely because there is such a thing as literary competence, and acquisition of that competence requires a considerable amount of time and energy in the form of intensive study.

The present study will be primarily concerned with competence *vis a vis* one particular class of literary works, viz., dramas. In his book *The Idea of a Theatre*, Francis Fergusson speaks of what he calls "histrionic sensibility": "The trained ear perceives and discriminates sounds; the histrionic sensibility (which may also be trained) perceives and discriminates actions." Subsequent remarks will utilize the phrase "histrionic competence" to refer to the body of implicit knowledge that presumably underlies a suitably trained histrionic sensibility.

Earlier remarks have provided some indication of the sorts of capacities that would be associated with the postulation of histrionic competence, viz., the ability to make fairly reliable judgments concerning the plot-synopsis, structural organization, and well-formedness of dramatic works; but would the possession of histrionic competence necessarily involve the ability to make such judgments with

---

respect to dramatic works of all styles and genres? Must a person possess competence with respect to every dramatic genre and style before he can be said to possess histrionic competence? A parallel question in the case of language might be whether a person must possess competence with respect to all of a particular language's dialects before he can be said to possess competence with respect to that language. The answer in the latter case is clearly no; competence with respect to only one dialect might even be enough in the case of language. The answer to the former question is, however, considerably less clear. If a person were baffled by comedies, it would seem rather inappropriate to say that he possessed histrionic competence—even if he were quite adept at comprehending tragedies and melodramas. The same would appear to be true in the case of the major dramatic styles. The present study will, therefore, proceed on the tentative assumption that histrionic competence does require knowledge of the conventions that underlie all—or at any rate most—of the major "dialects" of drama.56

56 The notion that the major forms of literature like drama, lyric, and epic stand in roughly the same relation to the various literary genres and styles as natural languages do to dialects suggests that it may not be possible to devise classifications that apply throughout the domain of literature. There may be no more reason to expect that the "dialects" of drama correspond to those associated with the epic or lyric forms than there is to expect that
While the preceding discussion has indicated most of the important implications of Saussure's ideas, it has not identified all of them. Another implication concerns the use of competence in deciphering instances of parole. Since linguistic utterances will sometimes exhibit idiolectic eccentricities, a person's linguistic competence alone will not always enable him to understand a given utterance. When an utterance contains eccentricities that fall outside the scope of the hearer's competence and idiolect, the hearer will have to depend on additional resources, e.g., his ability to make inferences based on context, etc. The nature of literary works would seem to require that an individual go beyond his own competence even more frequently. One of the primary qualities attributed to works of art is their uniqueness. Great works are said to be especially unique; they are often said to be great because they have gone beyond--risen above--the norm. If such is the case, then an individual's literary competence may be expected to guide

his attempts to understand a playscript—but will not take him all the way. Something more, something beyond his competence will be necessary. 57

The earlier discussion of linguistics has yet another implication: just as language can be manifested in two quite different forms, written and spoken, drama can also be manifested in two quite different forms, the written script and performance. The foregoing comparison raises two questions: (1) do the written and performed versions of a given play constitute instances of parole associated with the same langue much as written and spoken tokens of a sentence of English constitute instances of parole associated with the same langue?, and (2) if so, should one form of dramatic parole be given greater weight in constructing a theory of langue in the same way that spoken parole has been given greater weight in constructing models of natural languages?

It would appear that a number of twentieth-century writers think that the first of the foregoing questions should be given a negative answer. As Richard Hornby observes, writers like Antonin Artaud and Richard Schechner have insisted that "the idea of theatrical performance has nothing whatsoever to do with literature, that playscripts

57 The capacity to transcend the limits of literary competence will be discussed in a bit more detail below, pp. 136-37.
have only a shadowy existence except in performance."58 However, such views seem highly implausible for a number of reasons. First, while there is admittedly a sense in which a play has not been completely realized until performance, there is, as Hornby goes on to point out, also "a sense in which a good playscript is a finished work"; it is always "possible to read a play to oneself for pleasure."59 Moreover, the experience a person gains through reading a play to himself and the experience provided by a performance are not nearly as dissimilar as the aforementioned view seems to suggest. Far from preventing readers from attending to its sensual aspects, a written script actually encourages them to do so; the very way in which a play is written requires readers imaginatively to see and hear the dramatis personae as they move about and speak their lines. Of course a performance will render such things in much greater detail, but the difference is one of degree—not of kind. Even the tendency to speak of reading and attending a performance of the same play indicates that the relationship between the playscript qua piece of literature and the playscript


qua performance piece is much closer than the aforementioned view suggests; if written and performed drama really had nothing whatsoever to do with each other, it would be nonsense to talk about reading and attending a performance of the same play.

In direct opposition to writers like Artaud and Schechner, the present study will maintain that the earlier question should be answered in the affirmative: despite the differences between written and performed drama, they ought to be considered instances of parole that derive their meaning from the same set of underlying conventions or langue. In keeping with that view, the present study will not attempt to focus exclusively on written rather than performed drama or vice versa, but will, instead, attempt to focus on what they have in common.

The question remains, however, whether in attempting to develop a theory of dramatic langue—a theory of the conventions that make both written and performed drama possible—a theorist should give the evidence provided by one type of parole greater weight than that provided by the other. On the one hand, performances provide more detailed evidence of the sensual side of drama; on the other hand, written scripts provide evidence that is much less transitory. For the latter reason, the present study will tend to rely on the evidence provided by written scripts—but with an eye to
the sensual properties inherent in those scripts. The fore­
going position has, however, been adopted for purely prag­
matic reasons and does not, therefore, constitute an answer

to the question. The answer is unclear, and the present

study will not attempt to answer it.

The earlier discussion of Saussure's work possesses one
additional implication that warrants discussion. At one
point during the Course, Saussure concedes that "One might

if really necessary apply the term linguistics to [the study

of both langue and parole] and speak of a linguistics of

parole."60 A number of Structuralists have suggested a

roughly parallel division of literary studies: the branch

concerned with the study of langue is variously referred to

as the science of literature, literary theory, or poetics;

the branch concerned with the study of parole is usually

referred to as literary criticism.61

While there are similarities between the linguistic and

literary studies of langue and parole, there are also

important differences, most of which derive from the fact


60 Saussure, p. 19.

61 See, for example, Culler, Structuralist Poetics, p. 97, Scholes, pp. 143-47, and Todorov, Poetics of Prose, pp. 234-38.
that linguists and critics are not interested in *parole* for precisely the same reasons. Whereas linguists are interested in an individual utterance only because it furnishes evidence for the construction of a model of something else, e.g., a model of the English language, a model of the psycho-linguistic processes involved in the performance of linguistic acts, etc., dramatic critics are often interested in a drama for what it uniquely is. *Twelfth Night* is, for example, often studied, not as a means of finding out about the various conventions that underlie the phenomenon of drama in general or the psychological processes involved in writing or perceiving plays, etc., but in order to better understand *Twelfth Night* itself. Moreover, while the study of both linguistic and literary *langue* are assumed to be descriptive, the study of literary *parole*, unlike its linguistic counterpart, frequently concerns itself with prescriptive evaluation as part of its proper function. Finally, although study of literary *langue* may be considered primary in the sense that individual works cannot be studied without reference to some theoretical foundation, few writers would be willing to grant it the primacy that Saussure gave to its linguistic counterpart.
SYNCHRONIC AND DIACHRONIC APPROACHES

Having identified langue as the primary object of linguistic study, Saussure went on to observe that

The first thing that strikes us when we study the facts of [langue] is that their succession in time does not exist insofar as the speaker is concerned. He is confronted with a state. 62

Most native-speakers of a language do not, for example, know much, if anything, about the etymological history of the words they use—about the code-conventions of the past whose succession eventually gave rise to the code-conventions that underlie their own linguistic behavior.

Since most native-speakers of a natural language do not possess even an implicit knowledge of past linguistic conventions, those conventions cannot be counted as part of their langue—as part of the set of habits or conventions that enable them to communicate with each other. Thus, in order to individuate one langue from the next, linguists must not only distinguish English as it is currently spoken from French as it is currently spoken, etc., they must also distinguish English as it is currently spoken from English

62 Saussure, p. 81.
as it was spoken in the past, e.g., Old English, Middle English, etc. In order to deal with such temporal considerations, Saussure introduced the concept of a "language-state":

A language-state is . . . a certain span of time during which the sum of modifications [of linguistic conventions] that have superseded is minimal. The span may cover ten years, a generation, a century, or even more.63

Langue is thus to be equated with the relatively stable and unchanging set of conventions that obtain within a particular linguistic community during the span of a single language-state.

Saussure's realization of the role that temporal considerations play in the individuation of langues led him to insist that linguists draw a sharp distinction between what he referred to as the "synchronic" and "diachronic" approaches to linguistic study.64 Whereas the synchronic approach is exclusively concerned with the description of the more or less stable set of conventions that obtain within a linguistic community during a single language-state, the diachronic approach is concerned with the historical

63 Saussure, p. 101.
64 Saussure, p. 81.
relations that obtain between the conventions adopted during two or more successive language-states. 65

Although Saussure acknowledged the value of diachronic study, he argued that there are several reasons for thinking synchronic study to be of greater importance. One reason derives from the restriction of langue to the set of conventions that remain more or less constant throughout the span of a particular language-state:

it is evident that the synchronic viewpoint predominates, for it is the true and only reality to the community of speakers . . . . The same is true of the linguist: if he takes the diachronic perspective, he no longer observes [langue] but rather a series of events that modify it. 66

Since the facts of langue, which constitute the primary object of linguistic study, can be identified only by utilizing the synchronic approach, it must naturally be granted priority. Moreover, it would appear that all diachronic study is necessarily dependent on the results of synchronic study. Before a linguist can identify the historical relations that obtain between the conventions in force during successive language-states, he must first identify those

65 Saussure, pp. 81, 99-100.
66 Saussure, p. 90.
conventions—and that can be done only by using the synchronic approach.\textsuperscript{67} The dependence of diachronic on synchronic study thus constitutes an additional reason for granting the latter priority.

While Saussure insisted that the primary goal of linguistic study ought to be the description of \textit{langue}, which requires use of the synchronic approach, he realized that it is also important to understand the historical genesis of any given \textit{langue}—and for that task the diachronic approach is indispensable.\textsuperscript{68} According to Saussure, every diachronic change of linguistic conventions originates in the \textit{parole} of one or more individuals, but "not all innovations of [parole] have the same success, and so long as they remain individual," they do not constitute a part of \textit{langue}.\textsuperscript{69}

Thus, in the history of any diachronic change there are always two distinct moments: (1) when it sprang up in individual usage; and (2) when it became a fact of \textit{[langue]}, outwardly identical but adopted by the community.\textsuperscript{70}

The history of such transitions is the proper object of

\textsuperscript{67} Saussure, pp. 180-82.

\textsuperscript{68} Saussure, p. 90.

\textsuperscript{69} Saussure, p. 98.

\textsuperscript{70} Saussure, p. 98.
Saussure's observation that not all innovations meet with the same success raises the question of why some are eventually adopted while others are not. Although Saussure considered the foregoing question and discussed a number of possible explanations, he ultimately left the question open. Nevertheless, a number of subsequent linguists have suggested that the future success or failure of individual innovations is determined by a given language's current structure. For example, according to Edward Sapir:

As we look about us and observe current usage, it is not likely to occur to us that our language has a "slope," that the changes of the next few centuries are in a sense prefigured in certain obscure tendencies of the present and that these changes, when consumed, will be seen to be but continuations of changes that have been already effected. We feel rather that our language is practically a fixed system and that what slight changes are destined to take place in it are as likely to move in one direction as another. The feeling is fallacious.

Recently, linguists have begun to shed some light on the "obscure tendencies" to which Sapir referred. For example, Lyons has suggested that languages always contain a certain

71 Saussure, pp. 147-51.

number of grammatical anomalies, and Talmy Givón has point-
out that they also contain various conflicting communicative
principles that never reach a state of perfect balance or
equilibrium, e.g., increases in ease of articulation tend to
decrease ease of speech perception and vice versa. Thus,
each speech community may, perhaps, be unconsciously pre-
disposed towards adopting those innovations that seem to
solve the anomalies and the imbalances inherent in its
current set of linguistic conventions.

Perhaps such observations might provide a reason for
thinking that languages themselves possess a capacity for
dynamic activity and thus constitute systems in the sense
discussed in Chapter Two of this study. Sapir even won-
ders at one point whether "language [has] the power to
cchange of its own accord over and above the involuntary
tendency of individuals to vary the norm." However, such
an animistic view of language does not really seem very
plausible. Linguistic change, like language itself, ulti-
mately depends on the people who use it; it is they who

73 Lyons, Theoretical, p. 37 and Talmy Givón, On
Understanding Grammar (New York: Academic Press, 1979),
pp. 268-69.

74 See above, pp. 51-52.

75 Sapir, p. 154.
bring about changes—not language itself.  

Although Saussure tended to speak of diachronic change in evolutionary terms, a tendency that no doubt resulted from the preeminent position accorded to the evolutionary model during his lifetime, virtually all current linguists maintain that there is no basis for thinking that languages have increased in sophistication or complexity since ancient times. "All human societies of which we have knowledge," says Lyons:

> speak languages of roughly equal complexity; and the differences of grammatical structure that we do find among languages throughout the world are such that they cannot be correlated with the cultural development of the people speaking them and cannot be used as evidence for the construction of an evolutionary theory of human language.  

Thus, languages do not evolve; they just change.

After drawing the distinction between the synchronic and diachronic approaches to linguistic study, Saussure went 

76 There is, however, one other aspect of language that might be considered evidence of dynamic activity; it will be discussed during the course of the following chapter—see below, pp. 248-49. Because use of the expression "language-system" has become standard, subsequent remarks will continue to make use of it; its use is not, however, intended to suggest that languages constitute systems in the precise sense described in Chapter Two.  

77 Lyons, Noam Chomsky, p. 18.
on to ask whether there are any features of language "that are everywhere and forever verifiable," features whose universality would make it possible to study language from a "panchronic" viewpoint. His answer was affirmative:

Since phonetic changes have always occurred and are still occurring, this general phenomenon is a permanent characteristic of speech; it is therefore one of the laws of speech. In linguistics as in chess there are rules that outlive all events.

Another universal feature of languages is their hierarchical structure; all languages consist of a limited number of basic phonemes which are combined in various sequences to form units on a higher level, and those units are combined to form units on a still higher level, etc. Partly due to the heavy stress that Noam Chomsky has placed on the study of linguistic universals, the last several decades have been marked by a major surge of interest in what Saussure called the panchronic approach to linguistic phenomena.

---

78 Saussure, p. 95.
79 Saussure, p. 95.
80 See, for example, Bolinger, pp. 16-17.
81 See, for example, Chomsky, Aspects, pp. 27-30 and Bierwisch, pp. 60-62.
LITERARY APPLICATION

As earlier remarks have indicated, the plausibility of setting up a sharp distinction between synchronic and diachronic linguistics derives in large part from the plausibility of positing the existence of language-states—periods of time during which the set of linguistic conventions that obtain within a particular community remain relatively stable. There are, however, several reasons why the notion of "literary-states," and hence the literary applicability of the synchronic-diachronic distinction, might seem problematic.

One reason derives from the common suggestion that every new literary work brings about changes in the existing set of literary conventions. In the Theory of Literature, for example, René Welleck and Austin Warren assert that "We must conceive . . . of literature as a whole system of works which is, with the accretion of new ones, constantly changing its relationships, growing as a changing whole." The notion that the whole system of literary works—and the conventions with which they are associated—are constantly

82 See above, pp. 120-22.

changing might, however, appear to rule out the postulation of literary states, i.e., periods of time during which the set of literary conventions remains relatively stable. In fact, Todorov has explicitly rejected comparisons between literature and language where the temporal relations between parole and langue or, as he puts it, individual and species is concerned because, according to him, "every work modifies the sum of possible works, each new exemplar alters the species." 84

While there may be a good deal of truth to the claim that acts of literary parole tend to alter the existing set of literary conventions more often than acts of linguistic parole tend to alter the conventions of a natural language, the claim that the former conventions are affected by every single new work seems implausible. Literature abounds with examples of minor works that changed nothing, and how many works have been written only to be read by their authors and weary publishers—never to be heard from again? 85 Surely at least some of those works contained

84 Todorov, Fantastic, p. 6.

85 Todorov is not unaware of the fact that many works of literature do not alter literary systems, but he attempts to get around the problem by making "literature" an evaluative term: only great works, only works that produce a "change in our previous notion of [literary] activity" can be called "literature"; "Texts that do not fulfill this
innovations—innovations that never gained acceptance because they never gained public exposure. Thus, contrary to the views espoused by Todorov, the process by which literary conventions are changed can be divided into the same two phases that Saussure attributed to linguistic change: the point at which an innovative act of *parole* occurs, and the point at which the community adopts those innovations, and they become part of *langue*. Moreover, the notion that massive innovations are constantly being assimilated by the community appears to be a misconception, the origins of which are (a) the fact that critics and historians tend to focus their attention on those works that have resulted in convention-shifts, and (b) a tendency to mistakenly assume that literary conventions have always changed

condition automatically pass into another category: that of so-called 'popular' or 'mass' literature" (*Fantastic*, p. 6). Todorov's characterization of literature thus renders the above claim true—but vacuously so; literary systems must be changed by every new work of literature because the only works that qualify as "literature" are precisely those that change literary systems. Moreover, his remarks still leave the impression that all innovative works become influential, which does not appear to be true; all influential works may be innovative, but not all innovative works become influential. Welleck and Warren rightly reject evaluative characterizations of literature, pointing out that limitation of the term to great works makes "the development of literary genres," Todorov's very purpose, "incomprehensible" (*Theory of Literature*, p. 21).
as rapidly as they have over the course of the last century. A careful look at literary history will, however, show that large-scale literary innovations occur much less frequently than is often suggested. When they do occur, they are commonly seen as the initiation of new literary periods or movements. Thus, as Claudio Guillén suggests, "the concept of literary period (Renaissance, Baroque) [and] the notion of movement (Romanticism, Symbolism)" may allow identification of the "synchronous 'states of literature.'"^86

The foregoing remarks are not, of course, intended to suggest that the history of literature is marked by periods of complete stasis; but neither is the history of language. At one point in the Course, Saussure observes that "There is really no such thing as [linguistic] immobility. . . . To each period there corresponds some appreciable evolution."^87 As a result, "studying a language-state means in practice disregarding changes of little importance."^88 Thus, the existence of literary-states is perfectly compatible with the claim that literary conventions are always changing to

---


^87 Saussure, p. 140.

^88 Saussure, p. 102.
some degree; it is only incompatible with the claim that major changes are constantly occurring.

The preceding pages have been concerned with the possibility that the frequency of literary innovation might undermine attempts to apply the synchronic-diachronic distinction within the domain of literary studies, but there is a second reason why literary application of the distinction might be thought inappropriate—on based not on the frequency with which new conventions are adopted, but on the infrequency with which previous conventions are eliminated. Whereas the languages of previous linguistic states are rarely spoken today, e.g., outside of literature and historical linguistics no one speaks Elizabethan English today, previous literary "languages" are "spoken" constantly, e.g., Shakespeare is "spoken" on pages and stages throughout the world. Naturally, the foregoing fact also affects the way in which the respective competencies are acquired. Because the languages of previous linguistic states are no longer spoken, a child learns his first natural language by drawing exclusively upon contemporary—not historical—examples. An individual is, however, certain to draw upon historical examples as he acquires literary competence. Thus, whereas a normal native-speaker's linguistic competence does not include knowledge of past linguistic conventions, literary competence does seem to include knowledge of past
conventions.

As earlier remarks have suggested, Saussure stressed the importance of drawing the synchronic-diachronic distinction precisely because linguistic competence does not include knowledge of past conventions: since the linguist who wishes to study langue must concern himself with only those conventions implicitly known by the members of a given community, and since those individuals do not possess knowledge of past conventions, the linguist must adopt a strictly synchronic approach. Since literary competence does include knowledge of past conventions, it might seem that the synchronic-diachronic distinction is inapplicable within the domain of literary studies. The initial plausibility of the preceding claim is, however, diminished by the fact that not all literary conventions become incorporated in the langue of subsequent literary-states. For example, during the seventeenth and eighteenth centuries, many of the conventions of medieval drama fell into virtual obscurity. Thus, the synchronic-diachronic distinction seems to be applicable within the domain of literary studies after all. The synchronic approach will be concerned with identification of those conventions that obtain within a given

---

89 See above, pp. 120-22, 128.
community during the span of a particular literary-state—whether or not they have also been employed during the course of previous states. Thus, the synchronic approach will still be exclusively concerned with the conventions that obtain during the span of a given literary-state; it is just that some of those conventions will have been in force during the course of previous states as well. The diachronic approach will be concerned with the history of literary conventions—not only those that continued to constitute part of subsequent *langues*, but also those that fell into obscurity.

Earlier in the present discussion it was suggested that not all literary innovations become incorporated into *langue*. Thus, it would appear that certain sorts of innovations must be favored over others—but which ones? The Russian Formalists, whose ideas have heavily influenced the development of literary Structuralism, identified one sort of innovation that seems to be favored. Victor Shklovsky believed that one of the principal characteristics of an art

---

90 See above, pp. 129-31.

object is its capacity to renew perception by means of defamiliarization—by making things that are habitually taken for granted seem strange. Shklovsky went on to point out that techniques that at first enable defamiliarization eventually become hackneyed conventions; they begin to prevent the very effect that they originally made possible. Diachronic change, therefore, became, for Shklovsky, a necessarily continuous process whereby hackneyed techniques are replaced by new ones. Thus, just as a linguistic community adopts innovations that will help prevent malfunctions in their language's communicative efficacy, a literary community adopts those innovations that will help prevent malfunctions in the aesthetic efficacy of their artworks. The great, the influential, works are thus often the ones that replace the increasingly hackneyed portions of an existing literary-system with innovative conventions that promote defamiliarization.

Innovations may, however, also take another form—a form that stands in sharp constrast to the previous one. A literary work may become influential, not because it introduces new conventions, but because it actualizes certain

---

92 The Russian term used by Shklovsky is "ostranenie" (Jameson, pp. 50-51).

potentials of the existing conventions more fully than any previous work. Such a work may be said to be innovative because it reveals, for the first time, latent possibilities that had hardly been recognized, and, although such a work does not introduce new elements into langue, it may well alter the relative emphasis placed on various existing conventions.

While the preceding sort of innovation stands in sharp contrast to the sort discussed by the Russian Formalists, the two often work together. Shakespeare's plays are, for example, important and seem to have been innovative not only because they replaced portions of the dramatic system of conventions that he inherited from writers like Sackville, Norton, Kyd, and Marlowe, but also because they brought to fruition numerous latent possibilities inherent in the set of conventions employed by those writers.

The fact that outstanding works usually change existing literary systems reinforces a previously discussed idea: the idea that literary competence will take an individual most of the way—but not all the way—towards an understanding of great works of literature. Todorov says that "in art we are dealing with a language of which every utterance is

agrammatical at the moment of its performance."\(^9\)

Because of Todorov's evaluative definition of literature and, presumably, art in general, his statement requires some modification. Not every literary work is agrammatical, only the innovative ones—and only some of them are outstanding. Furthermore, influential works are almost never thoroughly agrammatical, and a reader's familiarity with the existing conventions that are employed may aid him in his attempts to come to terms with the innovative features of the work. Moreover, since the innovative features of a work often constitute reactions against existing conventions, even familiarity with existing conventions not employed may aid a reader's efforts to understand the innovative aspects of a work. Thus, an innovative work of art cannot be fully comprehended unless an individual both utilizes and goes beyond the system of conventions that enabled its composition.

Another issue raised by Saussure's discussion of the synchronic and diachronic approaches concerns the literary applicability of the non-evolutionary model of linguistic systems. Does the non-evolutionary model accurately characterize the relation between the systems of conventions that obtain during different states or periods of literary

---

history? As Scholes observes, belief that it does, which is fairly common among literary Structuralists, runs counter to the Marxist belief in the purposefulness of history and has, as a result, led many Marxist critics to vigorously attack Structuralism.\(^{96}\) The non-evolutionary view of literary history has not, however, been adopted by all Structuralists. While Shklovsky's account of diachronic change, which has been accepted by a number of Structuralists, is non-evolutionary, Frye, whose work is often characterized as Structuralist, implies, in his theory of modes, that literary change is evolutionary.\(^{97}\) In view of the fact that both views have been advanced, it would seem that Structuralism does not really dictate a dogmatic position on the issue. The matter would, in fact, appear to be entirely empirical. The empirical evidence does not seem to support an evolutionary model of natural languages; the

\(^{96}\) Scholes, p. 18. Most of Jameson's objections to Structuralism seem, for example, to be motivated by his Marxist inclinations. He argues that Saussure's distinction between the synchronic and diachronic approaches determined everything else that he said and continually argues that the distinction prevents Structuralism from formulating an adequate image of history. See, for instance, his remarks on pp. 5-7, 18-21, 38-39.

\(^{97}\) Discussions of Frye's work as part of the Structuralist corpus occur in Scholes, pp. 65, 118-28, Culler, pp. 119-22, and Todorov, pp. 8-19. Frye's theory of modes is contained in his *Anatomy of Criticism*, pp. 33-49. The evolutionary implications of his theory are discussed by Scholes, p. 123.
evidence pertinent to the study of other phenomena may or may not support an evolutionary model. In the case of literature, the appropriate interpretation of the empirical evidence seems, at present, far from clear.

The earlier discussion of Saussure's work also raises the question of whether there might be literary universals of some sort, the existence of which would make it possible to adopt a panchronic approach to the study of literary phenomena. Perhaps in the field of literary studies it would be most useful to separate questions about the existence of universal features into two sorts: (1) are there features that are universal to a given form or to a given genre of literature—are there, for example, features that are common to all dramatic works or to all tragedies, regardless of the time and place of their origin?, and (2) are there features that are universal to all forms of literature—features that could be properly called literary universals, as opposed to dramatic universals, narrative universals, lyric universals, etc.? Due to the much more general nature of the latter question, it would, of course, be much more difficult to answer with any confidence; and subsequent remarks will have little more to say about it. The question of whether there are dramatic universals will, however, arise at several different points during the remainder of this study.
THE UNITS AND PRINCIPLES OF LANGUE

Having argued that the primary concern of linguistics should be the study of langue from a synchronic perspective, Saussure turned his attention to the business of analyzing the nature of the principal elements or units of langue—linguistic signs. He began by observing that

Some people regard [langue], when reduced to its elements, as a naming-process only—a list of words, each corresponding to the thing that it names.98

While Saussure believed that proponents of such a view were right to maintain that "the linguistic unit is a double entity, one formed by the associating of two terms," he argued that they were wrong to suggest that those terms are non-psychological in character:

We have seen in considering the speaking-circuit . . . that both terms involved in the linguistic sign are psychological and are united in the brain by an associative bond. This point must be emphasized. The linguistic sign unites, not a thing and a name, but a concept and a sound-image.99

98 Saussure, p. 65.
In order to prevent ambiguities that might arise from the tendency to use the term "sign" as a means of referring to not only the combination of a sound-image and a concept, but also to the sound-image alone, Saussure decided to replace the terms "concept" and "sound-image" respectively by "signified [signifié]" and "signifier [signifiant]."\(^{100}\)

Saussure thus drew a sharp distinction between the internal constituents of a sign, its signifier and signified, and what will subsequently be called the "referent" of a sign—the entity or entities which the sign stands for or signifies.\(^{101}\) Of course, not all signs have a referent. For example, words like hello and goodbye and interjections like oh and drat do not function referentially at all.

---

\(^{100}\) Saussure, p. 67. Unlike the terms "langue," "langage," and "parole," commentators have generally adopted the practice of using the English terms "signified" and "signifier" in place of Saussure's "signifié" and "signifiant." The present study will follow that practice. Unfortunately, the Course is not itself free of the sort of ambiguity Saussure wished to avoid. For example, he sometimes uses the French equivalent of the term "word" when he is talking about the signifier; on other occasions he uses "word" when he is talking about signs. In order to avoid confusion, when the term "word" appears in a quotation from the Course, it will be followed by the term "signifier" enclosed in brackets if that is Saussure's intended meaning.

\(^{101}\) As subsequent remarks will indicate, some writers have a more technical notion in mind when they use terms like "reference" and "referent"—see below, p. 185. The terms will, however, be used in a looser, everyday sense in the present study.
According to Saussure, such terms are, however, composed of both a signifier and a signified.\(^{102}\)

Insofar as Saussure's analysis draws a sharp distinction between the concept associated with a particular signifier and the entity or entities to which that signifier refers, it resembles a common view of language whose history can, as Lyons observes, be traced back at least to the Middle Ages.\(^ {103}\) The extent to which Saussure's analysis can be said to resemble the traditional view in other respects is, however, a topic about which there appears to be some disagreement; a discussion of the matter is included in Appendix B at the end of this study.\(^ {104}\)

Having identified the internal composition of the linguistic sign, Saussure went on to state what he considered

\(^{102}\) Saussure, p. 69.

\(^{103}\) Lyons, \textit{Semantics}, pp. 96-97.

\(^{104}\) See below, pp. 584-89. There are several respects in which Saussure's views concerning the units and principles of \textit{langue} either raise controversial issues or seem in need of further elaboration. In view of the complexity of the issues involved, it seems best, however, to devote separate sections to the discussion of such matters so as not overly to complicate the initial exposition of Saussure's views. Issues that bear directly on the literary applicability of Saussure's views are discussed immediately following the present section; other issues are dealt with in Appendix B. Of particular concern is Saussure's vagueness with respect to the identity conditions of signifieds. While the following section of this chapter will attempt to eliminate much of that vagueness, the present section will not try to elaborate beyond what Saussure had to say about the matter.
to be the most important principle of langue: "the linguistic sign is arbitrary." One obvious sense in which the linguistic sign is arbitrary is that there is no necessary connection between its two components, the signifier and signified. That, says Saussure, is what distinguishes signs from symbols, for the relation between the two components of a symbol is never wholly arbitrary; it is not empty, for there is the rudiment of a natural bond between the signifier and the signified. The symbol of justice, a pair of scales, could not be replaced by just any other symbol, such as a chariot.

There is, however, no such natural bond between the components of a sign; "the signified 'ox' has as its signifier b-o-f on one side of the border and o-k-s (Ochs) on the other"--and neither signifier seems to be more naturally suited to the concept in question.

While Saussure claimed that the connection between the two components of all linguistic signs is arbitrary, he did not, however, believe that the connection is equally

105 Saussure, p. 67.
106 Saussure, p. 68.
107 Saussure, p. 68. The present study will follow Saussure's practice of distinguishing the names of particular signifieds from the names of signifiers by placing quotation marks around the former names and underlining the latter ones.
arbitrary in every case:

Some signs are absolutely arbitrary; in others we note, not its complete absence, but the presence of degrees of arbitrariness; the sign may be relatively motivated. 108

The distinction Saussure is trying to make may be illustrated by the distinction between the signs house and boat on the one hand and the sign houseboat on the other. Because the latter sign can be segmented into two separate parts each of which constitutes an independently meaningful sign in its own right, and because the meaning of the latter sign is a function of the meanings of those two parts and their mode of combination, the connection between the latter sign's signifier and signified may be said to be partially motivated by the prior connection between the signifiers and signifieds of its two parts. The sign is still arbitrary in the sense that the concept in question could have been correlated with some other signifier—say blurp or grimmel—rather than houseboat; but it is only relatively arbitrary. The foregoing cannot, however, be said about the signs house and boat; the connection between their signifiers and signifieds is absolutely arbitrary.

---

108 Saussure, p. 131.
One important feature of Saussure's distinction between absolutely and relatively arbitrary signs is that it allows him to distinguish between what the previous chapter of the present study referred to as "basic" as opposed to "derived" expressions. Whenever the principle of compositionality can be used to characterize the relation between a sign and its parts, that sign will count as relatively arbitrary; whenever the principle cannot be used to characterize that relation, the sign will count as absolutely arbitrary. Thus, all absolutely arbitrary signs constitute basic expressions, and all relatively arbitrary signs constitute derived expressions.

Culler argues that Saussure's remarks concerning the connection between the signifier and signified reveal only part of what he intended to suggest by saying that the linguistic sign is arbitrary; there is, argues Culler, evidence that Saussure thought the linguistic sign also contains a second, less obvious source of arbitrariness. In his remarks concerning the naming-theory of linguistic units, Saussure not only points out that the theory fails to recognize the psychological character of both components of

109 See above, pp. 61-62.

110 Culler, *Saussure*, pp. 20-23.
the sign, he also says that the theory is open to criticism because "It assumes that ready-made ideas exist before words [i.e., signifiers]." However:

If words [signifiers] stood for pre-existing concepts, they would all have exact equivalents in meaning from one langue to the next; but this is not true.

Consider, for example, differences in the way various cultures conceptualize colors. As Lyons observes, some languages have only two basic color terms, some have three or four, and some, like English, have as many as eleven.

According to Saussure, such conceptual variation across languages shows that the concepts employed by a particular linguistic community do not antedate the linguistic-system in which they are found—they emanate from that system. In fact, Saussure went on to claim that

\[\text{Psychologically our thought—apart from its expression in words—is only a shapeless and indistinct mass. . . . Without [langue], thought is a vague, uncharted nebula.}\]

\begin{enumerate}
\item Saussure, p. 65.
\item Saussure, p. 116.
\item Lyons, Semantics, p. 246.
\item Saussure, p. 117.
\item Saussure, pp. 111-12.
\end{enumerate}
Thus, according to Saussure, the evidence of conceptual variation across languages shows that it is the linguistic system itself that divides the continuum of thought into distinct units—and that linguistic communities can divide that continuum in a variety of different ways. In making the foregoing point, argues Culler, Saussure was implicitly suggesting that there is an additional source of arbitrariness in the linguistic sign: each language "has a distinctive and thus 'arbitrary' way of organizing the world into concepts or categories."\textsuperscript{116}

Much the same point can be made, moreover, with respect to the division of the sound-continuum into signifiers. As Lyons observes,

\begin{quote}
What counts . . . in establishing the inventory of phonemes in any language is whether the bits of the phonetic complex correlated with them stand in a relationship of functional contrast or not. Every language draws a more or less different, and in principle unique, set of distinctions in the continuum of sound and makes them functional by utilizing them to keep distinct the tokens of different word-types and utterance types.\textsuperscript{117}
\end{quote}

For example, the fact that in English the two acoustic sequences \textit{bet} and \textit{pet} count as different words rather than

\begin{flushright}
\textsuperscript{116} Culler, \textit{Saussure}, p. 23. \\
\textsuperscript{117} Lyons, \textit{Semantics}, p. 233.
\end{flushright}
as alternative pronunciations of the same word indicates that a \( b \)-sound and a \( p \)-sound stand in functional or, as subsequent remarks will put it, "institutional" contrast. The acoustic difference between the two sounds may therefore be said to constitute an "institutionalized" difference. There are, however, other languages in which the difference between the two sounds has not been institutionalized as a means of distinguishing word-types from each other, languages in which the two sounds may replace each other in the pronunciation of one and the same word.

Saussure presents similar examples of phonetic variation across languages, and concludes that, since linguistic communities can divide the continuum of sound in a variety of different ways by institutionalizing different acoustic contrasts, sound, as well as thought, must be amorphous apart from its embodiment in language. According to Culler, such remarks indicate that Saussure believed that it is not only the connection between the signifier and signified that is arbitrary, there is a sense in which signifiers and

---

118 The term "institutional" will be used because the term "functional" will be used to mark a different contrast later in the present chapter.
120 Saussure, pp. 112, 119.
signifieds are themselves arbitrary.  

Whether or not Culler is correct in thinking that Saussure saw a connection between the principle of arbitrariness and the phenomena of conceptual and phonic variation across languages, the latter certainly influenced Saussure's view of langue quite heavily. As the foregoing pages have shown, such phenomena convinced Saussure that language cannot be plausibly viewed as means of correlating two sets of already predelimited units. Instead, argued Saussure, language must be viewed as a mechanism by means of which both the realm of thought and the realm of sound are simultaneously differentiated into discrete units.

In order to bring out his point more forcefully, Saussure introduced the following analogy:

[Langue] can . . . be compared with a sheet of paper: thought is the front and the sound the back; one cannot cut the front without cutting the back at the same time.

Just as the front of a sheet of paper cannot be cut into a set of distinct shapes without at the same time cutting

121 Culler, Saussure, p. 23.
122 Saussure, p. 112.
123 Saussure, p. 113.
the back into corresponding shapes, so thought cannot be
differentiated into distinct units without at the same time
differentiating sound into distinct units—and vice versa.
Thus, at the moment when language comes into being, three
distinct—but inseparable—phenomena occur: the realm of
thought is divided into discrete units, the realm of sound
is divided into discrete units, and correspondences between
the units of thought and sound are established.124

As the foregoing analogy might suggest, Saussure be­
lieved that the identity of each sign is based, at least in
part, on the correspondence established between its two in­
ternal constituents.125 The identity of the English word
bed, for example, derives in part from the fact that it con­
sists of the sound-image bed and the concept "bed." That
claim, however, naturally raises questions concerning the
identity of signifiers and signifieds. To say that each
signifier derives its identity wholly from its association
with a particular signified—and vice versa—would, of
course, lead to a circular account of the identity of the
sign and its two components. Moreover, a theorist could
assume that each of a sign's two components derive their

124 Saussure, pp. 112-13. Of course, signifiers and
signifieds are not all linked in a relation of unique cor­
respondence, as the phenomena of synonymy and homonymy
clearly indicate.

125 Saussure, pp. 112-13.
identity entirely from their association with each other only if there were a necessary connection between them, and Saussure has already argued that such an assumption is thoroughly implausible. Thus, the identity of a sign's two components must derive, at least in part, from some other source—but what source could that be?

Saussure's answer was that the identity of a signifier derives largely from "the phonic differences that make it possible to distinguish [it] from all others." In other words, Saussure believed that its identity derives largely from a variety of institutionalized acoustic differences. For example, the identity of the signifier bed derives largely from the network of institutionalized acoustic differences that separate it from led, bad, bet, etc.

Likewise, the identity of a signified derives largely from a network of institutionalized conceptual differences that distinguish it from all other signifieds within the same linguistic system:

concepts are purely differential and are defined not by their positive content but negatively by their relations with the other terms of the system. Their most precise characteristic is in being what the others are not.

126 Saussure, p. 118.
127 Saussure, p. 117.
As Culler observes, color concepts provide a particularly clear example of what Saussure had in mind.  

Suppose that a person wished to teach the concept "brown" to a foreigner whose language only contains two basic color-terms, one for all the light colors and one for all the dark colors. It would not be enough simply to show the foreigner a number of brown objects, for he might well take that as an indication that the signifier brown is associated with the signified which in his language covers the entire range of dark colors. In order to acquaint the foreigner with the concept "brown," it would be necessary to distinguish it from all of the other color concepts of English. "The reason for this," says Culler:  

is that brown is not an independent concept defined by some essential properties but one term in a system of colour terms, defined by its relations with the other terms which delimit it.  

The identity of the concept "brown" is thus fixed by a network of relations that are essentially negative or differential: "Brown is what is not red, black, grey, yellow, etc.,

128 Culler, Saussure, p. 24.  
129 Culler, Saussure, p. 25.
and the same holds for each of the other signifieds."

The concept of differential relations thus plays an extremely important role in Saussure's account of language. An indication of just how much importance Saussure accorded the notion can be gathered from the following passage:

Everything that has been said up to this point boils down to this: in [langue] there are only differences. Even more important: a difference generally implies positive terms between which the difference is set up; but in [langue] there are only differences without positive terms. Whether we take the signified or the signifier, [langue] has neither ideas nor sounds that existed before the linguistic system, but only conceptual and phonic differences that issued from the system.

While previous commentators have rightly emphasized the importance of the preceding passage, they have generally failed, however, to point out that no more than four lines later Saussure goes on to add a crucial qualification to what he has said:

Although both the signified and the signifier are purely differential and negative when considered separately, their combination is a positive fact; it is even the sole type of facts that [langue] has, for maintaining the

---

131 Saussure, p. 120.
parallelism between the two classes of differences is the distinctive function of the linguistic institution.\textsuperscript{132}

While Saussure placed a great deal of emphasis on the negative character of the differences that account for the identity of signifiers and signifieds when they are considered in isolation from each other, he thus placed an equal amount of emphasis on the positive character of the correlations that come into view when both parts of the sign are taken into consideration.

It is important to note, moreover, that Saussure's statement that the signifier and signified are purely negative only when considered separately appears to imply not only that the identity of a sign derives, in part, from the correspondence between its two components, but also that the identity of each of the two components derives—at least to some extent—from their relation with each other. The identity of the concept "bed" may derive primarily from the network of differences that distinguish it from other concepts, but its association with the sound-image \underline{bed} also makes at least some contribution. After all, as Saussure's analogy with the sheet of paper indicates, it would be possible to establish neither the network of conceptual differences nor

\textsuperscript{132} Saussure, pp. 120-21.
the network of phonic differences without simultaneously establishing the correlations between the realms of sound and thought, correlations like the one that obtains between bed and "bed." Those correlations must, therefore, play at least some role in accounting for the identity of each signifier and signified.

According to Saussure, the identity of not only each sign, but also each signifier and signified thus derives from the establishment of two quite different sorts of relations: 1. the sort of relation that holds between a signifier and its corresponding signified, and 2. the sort of relation that separates a signifier from all other signifiers and separates a signified from all other signifieds. Jameson distinguishes between the two types of relations by referring to the former as "frontal" or "depth" relations and the latter as "lateral" relations. As both Barthes and Lévi-Strauss have observed, Saussure's conception of the relations between signs can be expressed as a series of homologies each of which laterally relate the signifiers and signifieds of two different signs and, at the same time, frontally relate the signifier and signified within each of

133 Jameson, p. 32.
the two signs.\textsuperscript{134}

The importance of Saussure's conviction that the identity of linguistic units is entirely based on the complex network of depth and lateral relations that obtain within a particular language-system would be difficult to overemphasize. Indeed, Lyons maintains that it constitutes the central thesis of linguistic Structuralism:

What, then, is the central thesis of structuralism? To put it first in its most general form, it is this: that every language is a unique relational structure, or system, and that the units which we identify, or postulate as theoretical constructs, in analyzing the sentence of a particular language (sounds, words, meanings, etc.) derive both their essence and their existence from their relationships with other units in the same language-system. We cannot first identify the units and then, at a subsequent stage of the analysis, enquire what combinatorial or other relations hold between them: we simultaneously identify both the units and their interrelations. Linguistic units are but points in a system, or network, of relations; they are the terminals of these relations, and they have no prior or independent existence.\textsuperscript{133}

Thus, all Structuralists have followed Saussure in emphasizing the fundamentally relational character of


\textsuperscript{135} Lyons, \textit{Semantics}, pp. 231-32.
language-systems.

It is important to note, moreover, that the foregoing thesis identifies a manifestation of holistic organization within the realm of linguistic phenomena not discussed in Chapter Two of the present study. Not only do individual utterances possess the property of wholeness, as was claimed in Chapter Two, the language-systems that create the possibility of those utterances are also holistically organized; the units out of which a language-system is constituted exhibit the same sort of functional interdependence that was attributed to the constituents of individual utterances in Chapter Two.136 *Langue*, says Saussure, "is a system of interdependent terms in which the [identity] of each term results solely from the simultaneous presence of the others."137 That is why it is appropriate to refer to natural languages as systems—or at any rate Structures.138

Having distinguished lateral from depth relations, Saussure went on to observe that the former relations

---

136 See above, pp. 63-65.

137 Saussure, p. 114.

138 The reasons for the inclusion of the qualificatory clause "or at any rate Structures" in the preceding statement are discussed above, pp. 125-26. Subsequent remarks will, however, continue to follow the practice of referring to natural languages as "systems" for reasons discussed above in fn. # 76, p. 126.
between linguistic terms fall into two distinct groups:

In discourse, on the one hand, words acquire relations based on the linear nature of language because they are chained together... Combinations supported by linearity are syntagms... In the syntagm a term acquires its value only because it stands in opposition to everything that precedes or follows it, or to both.

Outside discourse, on the other hand, words acquire relations of a different kind. Those that have something in common are associated in the memory, resulting in groups marked by diverse relations... They are associative relations.

For example, in the phrase a pride of lions, the sign pride contracts syntagmatic relations with the three signs that precede and follow it--a, of, and lions. More precisely, that sign contracts syntagmatic relations of two sorts with the other three signs. On the one hand, its signifier is distinguished from the other three signifiers by the acoustic differences between them; and on the other hand, its signified is distinguished from the other three

139 Saussure, p. 123. Although Saussure used the French equivalent of the term "associative" to refer to the second type of relation, the term "paradigmatic" has replaced it in modern linguistics, and it will, therefore, be used throughout the present study. It will also appear, in brackets, following the term "associative" within quotations from the Course. Lyons has pointed out that syntagmatic relations are not always sequential, citing counter-examples from intonational languages (Lyons, Theoretical, p. 76). Syntagmatic relations are, however, usually sequential, and thinking of them in such terms seems to facilitate understanding and identification of them.
signifieds by the conceptual differences between them.

Similarly, the sign contracts associative or paradigmatic relations of two distinct sorts. On the one hand, its signifier contracts relations with other acoustically related signifiers like bride, prude, prize, etc.; and on the other hand, its signified contracts relations with other, conceptually related signifieds like "group," "bevy," "flock," etc.

Because linguistic signs can be paradigmatically related to each other in so many different ways, Saussure remarked that "A particular word is like the center of a constellation; it is the point of convergence of an indefinite number of co-ordinated terms." 140 Although the implications of the preceding comment may seem, at first glance, to be radically holistic, suggesting a total absence of closure, a closer look at the analogy will reveal the limitations of Saussurean holism. If a term is like the center of a constellation, then its identity is likely to be more heavily influenced by those terms that are closest to it—both acoustically and conceptually. While the terms that lie farthest away may also be said to exert some influence, the degree of that influence may be infinitesimal. For

140 Saussure, p. 126.
example, as Culler points out:

If the word mutton were dropped from English certain local modifications would ensue: the value of sheep would change radically; beef, pork, veal, etc. would become slightly more anomalous with the disappearance of one member of their paradigm class . . . but vast areas of the language would not be affected in discernable ways. 141

Thus, the paradigmatic relations of a term may be thought of as forming a series of concentric rings that exhibit relative closure; the most important differential relations occur between the term in question and the members of the innermost circle—the least important between the term and the members of the outermost circle.

The examples that have thus far been used to illustrate Saussure's main points might suggest that he intended the extension of the term "sign" to include only individual words—and, in fact, Saussure himself virtually always uses individual words to illustrate his points. During his discussion of syntagmatic units, however, Saussure indicates that the extension of the term is much broader:

As a rule we do not communicate through isolated signs but rather through groups of signs, through organized masses that are themselves signs. 142

141 Culler, Structuralist Poetics, p. 14.
142 Saussure, p. 128.
Thus, the set of derived signs includes not only compound words like houseboat, but also phrases, clauses, and sentences. Does the set of signs then include all of the linguistic units identified in the preceding chapter of the present study—from the level of the phoneme to that of the sentence? There is reason to suppose that it does not.

At one point Saussure defines the sign as "a slice of sound which to the exclusion of everything else that precedes or follows it in the spoken chain is the signifier of a certain concept."\textsuperscript{143} Phonemes do not, however, meet the second of the preceding conditions; no phoneme by itself is associated with a distinct concept. Moreover, the same holds true of a significant number of morphemes. For example, neither of the morphemes in the word closet is associated with a distinct concept. Thus, it would appear that what were referred to as "elementary units" during the course of the preceding chapter of this study do not constitute signs; all other units do.\textsuperscript{144}

While elementary units may not constitute signs, Saussure believed that their identity must be accounted for in much the same way that the identity of signs is—in terms

\textsuperscript{143} Saussure, p. 104.

\textsuperscript{144} See above, pp. 61-62.
of the network of syntagmatic and paradigmatic relations that obtain within a particular language-system. In fact, Saussure's realization that the syntagmatic and paradigmatic relations that obtain within a particular language-system are primarily responsible for the identity of all linguistic units led him to place the specification of those relations at the center of his outline of a program for future linguistic research:

It would be interesting from a practical viewpoint to begin with units, to determine what they are and to account for their diversity by classifying them [on the basis of their syntagmatic and paradigmatic relations]. It would be necessary to search for the reason for dividing [langue] into words— for in spite of the difficulty of defining it, the word is a unit that strikes the mind, something central in the mechanism of [langue]. . . . Next we would have to classify the subunits, then the larger units, etc. By determining in this way the elements that it manipulates, synchronic linguistics would completely fulfill its task, for it would relate all synchronic phenomena to their fundamental principle.

As Lyons observes, the theoretical basis of the Saussurean program lies in the claim that

---

145 Saussure, pp. 119, 130-31.

146 Saussure, p. 111.
the structure of [a] language-system depends at every level upon the complementary principles of selection and combination. . . . To describe a language-system is to specify both the membership of [all] paradigmatic sets and [all] the possibilities of combination of one set with another in well-formed syntagms.147

The history of taxonomic linguistics could be appropriately described as the sustained attempt to carry out such a Saussurean program.

IMPLICATIONS AND EXTENSIONS OF SAUSSURE'S ANALYSIS OF LANGUE

As earlier remarks have indicated, there are several respects in which Saussure's views concerning the units and principles of langue either raise controversial issues or seem to be in need of further elaboration.148 Several of those issues do not bear directly on the literary applicability of Saussure's views and are therefore dealt with in Appendix B at the end of this study.149 One of the issues does, however, bear directly on the literary applicability of his views.

147 Lyons, Semantics, p. 241.
148 See above, p. 142.
149 See below, pp. 574-89.
In his discussion of conceptually oriented theories of language, Lyons raises the following objection:

As the term 'concept' is used by many writers on semantics, it is simply not clear what is meant by it; and that is perhaps of itself a sufficient criticism of their use of the term. It is after all a term with a long and controversial history; and anyone who [relies heavily on its use] owes his readers some explanation of what kind of thing [a] concept might be.\textsuperscript{150}

Such theories of language do, of course, typically provide a very general indication of what kind of things concepts might be--they indicate that concepts are mental entities of some sort--but that is obviously not the sort of information Lyons has in mind. He is probably suggesting that most mentalistic theories fail to specify a sufficient amount of information about the identity conditions of concepts; without such information, talk about concepts is really too vague to be of much explanatory value.

Unfortunately, the Course is far from immune to the foregoing sort of criticism. Saussure does, of course, say that the identity of a concept is determined by both its

\textsuperscript{150} Lyons, \textit{Semantics}, p. 113.
association with a particular signifier and by the lateral relations it contracts with other concepts, but that is not enough.\textsuperscript{151} What is needed is a more detailed account of the latter sorts of relations, one that specifies the salient differences between concepts the way that phonemic analysis has succeeded in specifying the salient differences between signifiers.\textsuperscript{152} Without such an account, the explanatory value of saying that each signifier is associated with a specific concept will remain largely empty. There is, however, a problem inherent in Saussure's original conception of signifieds that must be resolved before such an account can be attempted.

As earlier remarks have indicated, Saussure tended to think of langue as "a dictionary of which identical copies have been distributed to each individual."\textsuperscript{153} As those remarks pointed out, one problem with such a view is that each member of a particular linguistic community actually possesses implicit knowledge of only a proper subset of all the "entries" that would have to be included in the complete

\textsuperscript{151} See above, pp. 149-55.

\textsuperscript{152} For a discussion of phonemic analysis, see Bolinger, pp. 77-83.

\textsuperscript{153} See above, p. 95.
and perfect "dictionary" of their language. There is, however, another important problem inherent in Saussure's view: it assumes that the information associated with each entry in the dictionary is the same for each individual, and that is obviously not the case. As a result, it appears that the signified must be identified with an abstract model that is based on--but is nevertheless distinct from--the implicit knowledge possessed by the individual members of a linguistic community. There is, moreover, reason to think that in a substantial number of cases at least part of the information included in the specification of a signified will be known by only a fraction of the members of the relevant community.

In "The Meaning of 'Meaning'" Hilary Putnam draws attention to the fact that there are a substantial number of cases where two or more extensionally distinct terms are--for many or even most members of the relevant linguistic community--associated with concepts that are indistinguishable from each other. Putnam speculates, for example, that, like himself, most English speakers probably associate exactly the same concept with the expressions elm tree and

---

154 See above, pp. 95.
beechn tree. Nevertheless, says Putnam, there are at least some speakers, e.g., botanists, who associate quite different concepts with the two terms. In such cases, says Putnam, a "division of linguistic labor" takes place; the knowledge possessed by a minority of a linguistic community becomes "part of the social meaning of [a] word while being unknown to almost all [other] speakers who acquire the word." In order to insure that all non-synonymous signifiers are associated with distinct signifieds, it appears that such information must be included in the specification of signifieds.

While Putnam argues that a person need not know all of the information associated with an expression of the aforementioned sort before he can be said to have mastered its use, Putnam thinks that knowledge of certain pieces of information is required for mastery. For example, a person must, according to Putnam, know of certain expressions that they are associated with a communally accepted stereotype. According to Putnam, it is, for instance, necessary that a person who has mastered the expression tiger know that stereotypical tigers are large, feline animals,

156 Putnam, pp. 226-27.
157 Putnam, p. 228.
158 Putnam, pp. 247-52.
that they are yellow with black stripes, and that they are fierce.\textsuperscript{159} Often the features included in a stereotypical characterization only hold true for paradigmatic members of the expression's extension (there are, for example, albino tigers); sometimes the features do not hold true for any members of the extension (at one time, one of the features associated with the expression \textit{whale} was that of being a type of fish).\textsuperscript{160}

Towards the end of "The Meaning of 'Meaning'," Putnam suggests that the meaning of a substantial number of expressions should be described in terms of a finite sequence, or 'vector,' whose components should certainly include the following (it might be desirable to have other components as well): (1) the syntactic markers that apply to the word, e.g. 'noun'; (2) the semantic markers that apply to the word, e.g. 'animal', 'period of time'; (3) a description of the additional features of the stereotype, if any; (4) a description of the extension.\textsuperscript{161}

Putnam goes on to say that "the components of the vector all represent a hypothesis about the individual speaker's competence [i.e., the information specified by each component

\begin{itemize}
\item[\textsuperscript{159}] Putnam, pp. 250-52.
\item[\textsuperscript{160}] Putnam, pp. 250-51.
\item[\textsuperscript{161}] Putnam, p. 269.
\end{itemize}
is assumed to be implicitly known by anyone who has mastered a particular expression], except the extension," adding that, although it is necessary to employ a description in order to specify the fourth component of the vector, "we think of the component in question as being the extension (the set), not the description of the extension."\textsuperscript{162}

The relevance of Putnam's suggestion to the present discussion lies not in his claim that a description based on the aforementioned schema will provide an adequate analysis of an expression's meaning, but in the possibility that criteria for specifying the identity conditions for signifieds might be extracted from his proposal. If such criteria are to be obtained, however, the proposed schema will have to be modified in several ways.

The major modifications that are required derive from the limited applicability of Putnam's schema; it is, as Putnam himself indicates, suitable for describing only the subset of referring expressions usually called "kind terms."\textsuperscript{163} However, Saussure thought that all linguistic expressions are associated with a signified. As earlier remarks have indicated, for example, he maintained that

\textsuperscript{162} Putnam, pp. 269-70.

\textsuperscript{163} Putnam, p. 242.
interjections consist of both a signifier and a signified. If Putnam's schema is to provide a set of criteria that are adequate for use in individuating signifieds, it must therefore be modified in a way that will render it useful in describing not only additional types of referring expressions, but also expressions like oh and drat, which have no referents or extension at all. Moreover, the schema will also have to be modified in a way that will render it useful in describing derived as well as basic expressions—but how are those tasks to be accomplished?

The basis for an answer to that question can, perhaps, be extracted from some of the remarks contained in Ludwig Wittgenstein's Philosophical Investigations. During the course of the Investigations, Wittgenstein repeatedly criticizes earlier approaches to language for failing, in general, to place sufficient emphasis on the facts of language use—and for failing, in particular, to recognize the fact that reference is not the only use to which linguistic expressions are put. "Think of the tools in a tool-box," says Wittgenstein:

164 See above, pp. 141-42.

there is a hammer, pliers, a saw, a screwdriver, a rule, a glue-pot, glue, nails and screws.—The functions of words are as diverse as the functions of these objects.\textsuperscript{166}

Just as such tools could not be described without specifying their various different functions or uses, so linguistic expressions cannot be described without specifying their various different functions or uses. Perhaps then the applicability of Putnam’s schema can be extended by including a component that specifies the typical use or uses of each signifier.

A revised version of Putnam’s schema that could be used to individuate signifieds might thus include the following components: 1. a specification of whether the signifier in question is basic or derived; 2. a specification of the syntactic features of the signifier, e.g., the signifier is a noun, a verb phrase, etc.; 3. a specification of the typical use or uses of the signifier, e.g., the signifier is typically used to refer, to express emotion, to make an assertion, to issue a command, etc.; 4. where the signifier in question is basic and where one or more of the typical uses specified in 2 are non-referential, a specification of additional information that is relevant to those non-referential uses, e.g., information that the signifier drat

\textsuperscript{166} Wittgenstein, section # 11.
is typically used to express frustration or anger rather than joy, affection, etc.; 5. where the signifier in question is basic and where one or more of the typical uses specified in 2 are referential, a specification of the features associated with the relevant stereotype or stereotypes, or a specification of the experts to whom a non-expert should defer in his application of the signifier, or both; 6. where the signifier in question is basic and where one or more of the uses specified in 2 are referential, a specification of the extension or extensions according to relevant experts; and 7. where the signifier in question is derived, a specification of the way in which its typical use or uses are determined by the typical uses of its constituents and their mode of combination.

Several features of the foregoing proposal require a bit of additional explication. First, it should be noted that it will be necessary to associate two or more separate descriptions with a homonymous signifier like bank. Second, a signifier is not to be described as having more than one use merely because it could be used in more than one way—the additional uses must be typical. For example, although the signifier chair could, in certain circumstances, be used to express emotion, it is not typically used to so; in contrast, certain obscenities are typically used both for referential and for emotive purposes.
Moreover, since the proposal is only intended to specify the typical use or uses of a signifier, it is not by itself intended to account for the apparently limitless variety of different uses that can be made of a signifier in figurative statements, e.g., metaphor, irony, etc. However, in light of Searle's argument that the figurative uses of an expression derive from its literal uses plus the application of a set of inferential principles, it seems likely that, when combined with a set of principles like the ones discussed by Searle, the present proposal could account for those uses as well.  

Third, Putnam's requirement that the schema include a component that specifies the semantic markers associated with a referring expression has been dropped on the grounds that such information will be specified by either component 5 or 6, if not by both. Next, the second disjunct in the specification of the fifth component has been included to allow for cases in which mastery of a referring expression does not seem to depend on knowledge of a stereotype, but does seem to depend on knowing that a person should defer to certain experts' use of the expression. For example, it seems implausible to think that a person must possess knowledge of some sort of stereotypes before he can be said

to have mastered use of the signifiers quark and gluon, but mastery of those terms does seem to require that a person at least know that they are used to refer to some sort of entities discussed by physicists and that he should therefore defer to physicists with regard to their proper use.

Next, the sixth component is not to be thought of as being the extension itself, as it is on Putnam’s account, but as the best description of the extension provided by current experts. As a result, a total specification of all of the components will not necessarily determine the extension of a signifier, but that is perfectly in keeping with the purpose for which the schema is being proposed; application of the schema to a signifier is supposed to provide a specification of the signified with which the signifier is correlated, and, as remarks contained in Appendix B indicate, signifieds do not by themselves determine extension.

Since, in many cases, only a small group of experts will possess knowledge of the information specified by the sixth component, that knowledge is assumed not to be necessary for mastery of a given expression. However, the present proposal is intended to suggest that implicit knowledge of the information specified by all of the other components

168 See below, p. 587.
is necessary for mastery of a given signifier. Application of the descriptive schema to a particular signifier thus results in a sort of dossier or file of information about the syntactic role and typical use or uses of that signifier; mastery of the signifier requires knowledge of most but not all of the information contained in that dossier.

Finally, although the present proposal assumes that all of the information specified in the dossier is somehow stored in the brains of individuals (even the information specified by the sixth component is stored in the brains of the relevant experts), it makes no assumptions about whether such information is stored in the form of sentences, propositions, images, templates, or exemplars, etc. Specifically, it leaves open the possibility that the information associated with different types of signifieds might be stored in different ways—and even that the various pieces of information included in the specification of a single signified might be stored in different ways. For example, it might be the case with respect to some referring signifiers that the information that they are used for referential purposes is stored in the form of sentences or propositions, while information about the stereotypes with which they are associated is stored in the form of images or exemplars. Of course, hypotheses that certain sorts of information are stored in one form or another already involve
abstract levels of description; ultimately, the information is presumably stored in the neuro-chemical structure of the brain.

When taken together, the components of the foregoing descriptive schema appear to provide a set of criteria that could be used adequately to specify the identity conditions of signifieds. The overall plausibility of Saussure's theory of language can thus be considerably strengthened by modifying and extending it so as to include something like the preceding account.

The foregoing discussion has suggested that Saussure's theory of language be altered and extended in several respects. The overall view that results from incorporation of those suggestions may be summarized as follows: in order to account for a particular community's regular use of linguistic signs, it is necessary to formulate an abstract model that specifies (1) the signifier and signified associated with each basic sign of the language in question, and (2) a set of combinatory rules that can be used to specify the signifier and signified associated with any derived sign on the basis of information about its constituents and their mode of combination. The community's regular use of linguistic signs is assumed to derive in part from the fact that every competent member of the community possesses for each of a substantial subset of the basic signs of his
language at least partial knowledge of its associated signifier and signified and also possesses at least partial knowledge of the combinatorial rules that allow specification of derived signs. It is the preceding, modified version of Saussure that will be of primary concern throughout the remainder of the present study.

LITERARY APPLICATION

To the extent that the modified version of Saussure's theory provides insights into the basic units of language—from phonemes to sentences—it will, of course, be of use in analyzing the strictly linguistic structure of literary works. The following discussions are, however, primarily concerned with the possibility that the modified version of Saussure's theory might be of additional use to literary theorists. As earlier remarks have indicated, Structuralist critics typically claim that literary works can be segmented into units over and above the strictly linguistic units they contain. The following remarks are primarily concerned with the possibility that the modified version of Saussure's

169 Remarks contained in Appendix B indicate that the community's regular use of signs also derives in part from the fact that its members are situated in the world in a certain way; see below, pp. 588-89.

170 See above, pp. 68-70.
theory might provide insights into the nature of and relations between such literary units.

More specifically, the following remarks will explore the possibility that, like the strictly linguistic units out of which they are composed, such distinctively literary units constitute signs. According to the modified version of the Saussurean account, that would mean that such literary units consist of two components, a signifier and a signified. It would also mean that the identity of each such unit is based on a system of lateral, purely differential relations that distinguish its signifier from all other signifiers, a system of lateral relations that distinguish its signified from all other signifieds, and a depth relation which secures a bond between its signifier and signified. Before the plausibility of such hypotheses can be adequately assessed, however, it will be necessary to say something about how the two components of such units might be characterized, about the reasons for thinking that such units consist of both a signifier and a signified, and even about the reasons for postulating the existence of such units in the first place.

The initial plausibility of postulating the existence of distinctively literary units derives from several sources. First, as Pettit points out, it appears that the meaning of a literary work
is more than what is grasped just by understanding the individual sentences in which the text consists: I may understand these and miss the meaning of the whole.\footnote{171}{Pettit, p. 43.}

The possibility that a person might understand each sentence of a literary work perfectly well and nevertheless fail to understand the work as a whole suggests that the whole cannot be fully understood without grouping the individual sentences into larger, distinctively literary units. Additional motivation for postulating the existence of such literary units derives from the fact that the meaning of an entire work appears to depend on the order in which whole blocks of sentences are arranged; just as the meaning of a sentence would be altered if portions of it were rearranged, so too the meaning of a literary work would be altered if whole blocks of sentences were rearranged. If readers did not perceive literary works in terms of larger, distinctively literary units, it is difficult to see why such transpositions would have the effects they do.\footnote{172}{As Barthes points out, not all distinctively literary units will consist of two or more sentences: distinctively literary units will sometimes constitute "units higher than the sentence (groups of sentences of varying
that transcend the level of individual sentences, they also provide reasons for thinking that such larger units possess meaning—that they consist of both a signifier and a signified. Since the arrangement of such units appears to effect the meaning of the whole of which they are parts, and since the constituents of such units—the individual sentences out of which they are made—possess meaning, it seems plausible to assume that they too possess meaning. How, then, are the two components, the signifier and signified, of such units to be characterized?

It seems plausible to suppose that all of the signifiers in lyric and narrative works could be characterized in terms of their phonemic make-up and various prosodic features such as metrical pattern; since an individual sentence qua signifier may be characterized in terms of its phonemic make-up and prosodic features, blocks of sentences, up to

lengths, up to the work in its entirety) and [will] sometimes [constitute] lower ones (syntagm, word and even, within the word, certain literary elements only)" ("Narratives," p. 91). For example, an instance of alliteration, or assonance, or dissonance might span only several words, or even several syllables within a word, and would thus constitute a distinctively literary unit below the level of sentence. For purposes of exegetical simplicity, however, the following remarks will, with a few exceptions, focus on the description of larger units. Unless otherwise qualified, the phrase "distinctively literary unit" is therefore to be taken as a reference to such larger units.
and including the entirety of a lyric or narrative work, could be characterized in similar terms. When considered apart from the domain of signifieds, the identity of each signifier contained in a lyric or narrative work might thus be specified in terms of the phonemic and prosodic differences that distinguish it qua acoustic sequence from (1) the other phonemic sequences that constitute its syntagmatic environment within the work in question and (2) all those phonemic sequences which bear an acoustic resemblance to it and which therefore constitute the paradigmatic group to which it belongs.

While it may be feasible to characterize the signifiers contained in lyric and narrative works in purely phonemic and prosodic terms, it is not at all clear that the signifiers contained in dramatic works could be exclusively characterized in such terms. Of course, the words on the printed page could be exhaustively characterized in such terms, but there is a sense in which at least some of the sights and non-verbal sounds that are actualized in the theatrical performance of a play are also implicitly contained in the written text. Moreover, it seems plausible to say that such sights and sounds, whether they be explicitly manifested in a theatrical performance or implicitly suggested by a written text, signify the perceptible features of a fictional or imaginary world—the world inhabited by
the _dramatis personae_.

In the case of dramatic works, it may, therefore be necessary to draw a distinction between primary signifiers—the acoustic strings which make up the dialogue and stage directions of the written text—and secondary or derivative signifiers—sights and non-verbal sounds that are inferred from but are not themselves composed out of the former sorts of signifiers. Moreover, the primary and secondary signifiers contained in a dramatic work will often occur at the same time—e.g., a section of dialogue, a sequence of sights, and a sequence of non-verbal sounds will occur simultaneously. As a result, many, if not most, of the signifiers contained in a dramatic work will have to be characterized as an ordered sequence of "mixed percept-images," each percept-image consisting of simultaneously occurring phonemes, visual images, and non-verbal, acoustic images.

When considered apart from the domain of signifieds, the identity of each mixed signifier contained in a dramatic work might thus be specified in terms of the acoustic and visual differences that distinguish it from (1) the percept-images that constitute its syntagmatic environment within the work in question and (2) all those percept-images which

---

173 The claim that such visual and acoustic sequences signify or refer to objects and events in fictional worlds is controversial and will be discussed at some length shortly.
bear an acoustic or visual resemblance to it and which therefore constitute the variety of paradigmatic groups to which it belongs. Formulation of such characterizations would, of course, constitute an enormously complex task, one which would require theorists to utilize the techniques employed not only in the field of linguistics, but also those employed in kinesiology, proxemics, and a number of other disciplines as well. Practical difficulties aside, such characterizations nevertheless constitute at least a theoretical possibility.

The existence of such "mixed" signifiers raises the possibility that dramatic works may contain distinctively literary signifiers that do not possess an identifiable meaning—visual and non-verbal acoustic images that, like phonemic units, are not individually associated with a

---

174 Much of the work in kinesics, i.e., the study of body motion qua meaningful phenomena, has been done by Ray L. Birdwhistell; see, for example, his *Kinesics and Context: Essays on Body Communication* (Philadelphia: University of Pennsylvania Press, 1970). Much of the work in proxemics, i.e., the study of spatial relations qua meaningful phenomena, has been done by Edward T. Hall; see, for example, his book *The Hidden Dimension* (New York: Doubleday, 1966).

175 Several writers have, in fact, already made a significant amount of progress with respect to the description of the visual and non-verbal acoustic aspects of mixed-percept images; see, for example, Elam, pp. 50-87 and John George Gutting, "Toward a Structuralist Poetics of the Drama," Diss. The Ohio State University 1980, pp. 87-100.
definite meaning even though their elimination or replacement would alter the meaning of the larger units in which they are contained. The following remarks will, however, focus primarily on higher-level units that are composed of both a signifier and a signified.

Towards the end of the previous section of the present chapter, it was suggested that the typical uses of a given signifier would, in the case of strictly linguistic units, play a crucial role in the characterization of its corresponding signified. Perhaps an examination of the typical uses of literary signifiers will also prove fruitful in the search for a means of characterizing the signified component of distinctively literary units.

As earlier remarks have indicated, one of the most obvious uses of strictly linguistic signifiers is that of reference. Recently, there has, however, been a good deal of disagreement about whether linguistic units that normally function referentially, e.g., proper names, also function referentially within the context of a literary work. Although those disagreements are primarily concerned with the referential status of the strictly linguistic units contained in a literary work, they also bear on the

176 See above, pp. 170-72.
177 See above, pp. 171-72.
question of whether distinctively literary units function referentially, the major topic of present concern. Before the nature of those disagreements is considered, however, more should be said about how the term "reference" will be used during the course of the following remarks.

Many previous writers, e.g., Searle, draw a sharp distinction between the referential and the predicative use of an expression.¹⁷⁸ According to such writers, the phrase a man is used referentially in the sentence A man came, but it is not used referentially in the sentence John is a man.¹⁷⁹ Other writers, e.g., Michael Devitt, employ "reference" as a generic term and maintain that "proper names, predicates, variables, and so on, may all [be said] to refer."¹⁸⁰

During the course of the following remarks, the term "reference" will be construed in an even broader fashion: whenever a linguistic expression can be said to "stand for" something else, it will be said to function referentially, and whatever the expression stands for will be called the "referent" of that expression. As a result, not only the proper parts of sentences, but also whole sentences themselves will be said to function referentially. In the case of

¹⁷⁹ Searle, Speech Acts, p. 27.
John is a man, for example, John may be said to refer to a particular individual, is a man may be said to refer to a particular property, and the whole sentence may be said to refer to a particular state of affairs—the state of affairs that consists in a particular individual, John, possessing the property of being a man. By way of contrast, interjections and sentences such as Hello, Goodbye, etc. do not stand for any particular individuals, properties, or states of affairs and hence do not function referentially.

Most previous writers who have written on the subject of fictional reference have concerned themselves with not only the referential status of fictional names as they occur within literary discourse itself, they have also concerned themselves with the referential status of such names as they occur in talk about literary works. As a result, such writers tend to claim that there are really two questions that arise in connection with the issue of fictional reference: 1. do fictional names like "Hamlet" function referentially when they occur within sentences contained in a fictional work?, and 2. do such fictional names function referentially when they occur within sentences that are used

181 See, for example, Searle, Expression, pp. 70-71 and Devitt, pp. 170-72.
to talk about a fictional work? As a matter of convenience, sentences that are contained within a fictional work will hereafter be called "f-sentences," and sentences that are used to make comments about a fictional work will be called "c-sentences."

Previous writers can be roughly divided into three main camps on the basis of their answers to the foregoing two questions. There are those who maintain that fictional names function referentially in both f-sentences and c-sentences, those who maintain that fictional names do not function referentially in either context, and those who claim that such names function referentially in c-sentences but do not function referentially in f-sentences. Writers who fall within each camp will be called "liberals," "hard-liners," and "moderates" respectively. In addition, the liberal and moderate camps may each be subdivided into two distinct subgroups. Underlying not only the initial division of writers into three main camps, but also the subdivisions within the liberal and moderate camps are three primary sources of disagreement.

The first major source of disagreement derives from conflicting views concerning the plausibility of what Searle has called the "axiom of existence"—the claim that "Whatever
is referred to must exist."\textsuperscript{182} The second source of disagreement derives from conflicting views about whether there are such entities as "fictional objects," and the third source derives from conflicting views about whether, assuming that there are such entities, they can be said to exist. Thus, underlying the various conflicting answers to the two questions concerning the referential status of fictional names contained in f-sentences and c-sentences are conflicting answers to three prior questions: 1. should the axiom of existence be accepted?, 2. are there such entities as "fictional objects"?, and 3. if there are such entities as fictional objects, can they be said to exist, or do they possess some form of being other than existence, eg. subsistence?

Writers who fall within the liberal camp have virtually all agreed that the second of the foregoing questions should be answered affirmatively, but they have advanced two different sets of answers to the first and third questions. It is that difference which accounts for the subdivision of the liberal camp into two distinct subgroups. The writers who belong to one of the two subgroups claim that although fictional names refer to fictional entities, such entities constitute examples of what the turn-of-the-century German

\textsuperscript{182} Searle, \textit{Speech Acts}, p. 77.
philosopher Alexius Meinong called "non-existent" objects.\textsuperscript{183} Since such writers claim that it is possible to refer to objects that do not exist, they must, of course, reject the axiom of existence. Thus, such writers advance negative answers to both the first and the third of the aforementioned questions.

Writers who belong to the other subgroup within the liberal camp not only claim that there are such entities as fictional objects, they also claim that fictional objects exist (although there is a good deal of disagreement among such writers concerning various other issues).\textsuperscript{184} Since

\begin{flushleft}

It is possible that the differences between the two subgroups within the liberal camp are largely terminological rather than substantive; the labels "existent" and "non-existent" may perhaps be Meinongian substitutes for the labels "actual" and "non-actual." The liberal camp is divided into two subgroups in the main text in order to allow for the possibility that such is not the case.

\textsuperscript{184} See, for example, Elam, pp. 99-117, 150-51; David Lewis, "Truth in Fiction," in his Philosophical Papers, pp. 261-80; and Nicholas Wolterstorff, Works and Worlds of Art (Oxford: Clarendon Press, 1980), pp. 106-97. Whereas Lewis treats fictional objects as inhabitants of possible worlds, towards which he adopts a fully realistic attitude
such writers are prepared to say that fictional objects exist, there is no need for them to reject the axiom of existence, and, in fact, many of them explicitly endorse the axiom. \(^{185}\) Thus, members of the second subgroup posit affirmative answers to both the first and third questions, whereas Meinongian writers posit negative answers to both those questions.

Like the writers who belong to the second subgroup within the liberal camp, those who belong to the hard-liner camp wholeheartedly endorse the axiom of existence. Unlike the former group of writers, however, the hard-liners firmly reject the notion that there are any such things as fictional objects, existent or otherwise. For example, Russell insisted that

\[
\text{To say that unicorns have an existence in heraldry, or in literature, or in imagination, is a most pitiful and paltry evasion.} \ldots \text{Similarly, to maintain that Hamlet, for example, exists in his own world, namely, in the world of Shakespeare's imagination} \ldots \text{is to say something deliberately}
\]

\(^{185}\) See, for example, Wolterstorff, p. 137.
confusing, or else confused to a degree which is scarcely credible. There is only one world, the "real world". . . .

According to Russell, liberals like Meinong base their belief in fictional objects on something like the following reasoning: people can assert true propositions that contain terms like "unicorns" and "Hamlet" as grammatical subjects—hence, such terms must refer to entities that "have some kind of logical being, since otherwise the propositions in which they occur would be meaningless." 187

Russell agreed with the liberal writers about two points: 1. propositions that contain terms like "unicorns" and "Hamlet" as grammatical subjects are not normally meaningless, and 2. the assumption that such terms do not refer to anything would entail the view that such propositions are meaningless—if the logical form of such propositions ever turned out to be that of a simple subject-predicate

186 Bertrand Russell, Introduction to Mathematical Philosophy (1919; rpt. New York: MacMillan Co., 1938), p. 169. Although Russell maintained a hard-liner view throughout most of his career, it is interesting to note that he advanced a Meinongian position in his early writings: "Numbers, the Homeric gods, relations, chimeras, and four-dimensional spaces all have being, for if they were not entities of a kind, we could make no propositions about them. Thus being is a general attribute of everything, and to mention anything is to show that it is" [Bertrand Russell, The Principles of Mathematics (New York: W. W. Norton & Co., 1903), p. 449].

187 Russell, Mathematical Philosophy, p. 169.
construction. However, argued Russell, the superficial grammar of such sentences is systematically misleading. On analysis, such terms turn out to be disguised descriptions instead of names, and the superficial grammatical structure of the sentences containing such terms gives way to a more complex logical structure.  

Consider, for example, the c-sentence "Titania fell in love with a donkey." According to Russell's account, the signifier Titania should be treated—not as an instance of a proper name—but as an abbreviated version of some description like "the queen of the fairies," and the whole c-sentence should not be construed as a simple subject-predicate construction, but should, instead, be assigned roughly the following form: "There is an individual x such that x is a queen of the fairies, x is the only queen of the fairies, and x fell in love with a donkey." Russell claimed that, unlike the original, the foregoing paraphrase does not carry the misleading appearance of referring to a specific individual named "Titania"; and since the paraphrase captures the "real" underlying form of the original,


there is no need to posit a realm of fictional entities in order to account for the fact that the sentence in question is perfectly meaningful.\textsuperscript{190} Thus, Russell maintained that he could do justice to the widespread intuition that sentences containing fictional names are meaningful without sacrificing either the axiom of existence or his own ontological scruples. The question of whether Russell actually succeeded will be discussed shortly.

Since writers who belong to the moderate camp claim that fictional names function referentially when they occur in \(c\)-sentences but not when they occur in \(f\)-sentences, they must provide an answer to a question that does not arise for those who belong to the other two camps: how is it possible for fictional names to function referentially within the former sentences when they do not function referentially within the latter ones? Two different answers have been proposed. According to Searle:

\begin{quote}
By pretending to refer to (and recount the adventures of) a person, [an author] creates a fictional character. Notice that [the author] does not really refer to a fictional character because there was no such antecedently existing character; rather, by pretending to refer to a person [the author] creates
\end{quote}

\textsuperscript{190} Russell, \textit{Mathematical Philosophy}, pp. 169-71.
Thus, Searle posits affirmative answers to the second and third of the questions raised earlier: there are such things as fictional objects, and they exist. However, fictional objects do not come into existence until after the relevant fictional work has been written. That is apparently the reason why Searle claims that fictional names contained in f-sentences do not function referentially. Once the fictional names contained in f-sentences have been used to bring the relevant objects into existence, however, the way is clear for using the names referentially in c-sentences.

Unlike Searle, the other writers within the moderate camp follow Russell in insisting that there are no such things as fictional entities, antecedent to an author's creative efforts or otherwise. Unlike Russell, however, such writers claim that the fictional names contained in c-sentences nevertheless function referentially. While Devitt, for example, denies that fictional names either designate or denote in any straightforward sense, he nevertheless claims that the occurrence of such a name in a sentence used to

191 Searle, Expression, pp. 71-72.
talk about a fictional work (which he unfortunately calls an "f-sentence" rather than a c-sentence) does stand "in some other referential relationship to the world: it 'F-designates' those parts of the [relevant] fiction" which were causally responsible for use of the term in the sentence in question and which bear on the truth conditions of that sentence. Thus, the signifier Titania in the sentence "Titania fell in love with a donkey" refers to various

192 Devitt, p. 181. In an earlier passage, Devitt claims that the occurrences of a fictional name within the original work itself are causally linked to the imaginative acts involved in producing the various sentences of the work in question, which raises the possibility that occurrences of the name in sentences used to talk about the work ultimately refer to the relevant imaginative acts rather than the parts of the original work which flowed from those acts (Devitt, p. 175). If that were the case, however, it seems that Devitt would have to say that occurrences of the name within the fictional work itself also "F-designate" those imaginative acts and, hence, function referentially, a view which would place Devitt in the liberal rather than the moderate camp. While it would not be entirely implausible to construe Devitt as holding such a view, there are several reasons for thinking that he is best construed as a moderate. First, Devitt appears to restrict the notion of "F-designation" to the occurrences of fictional names within sentences used to talk about fictional works, and second, the passages in which he introduces the notion of "F-designation" link such occurrences of fictional names to parts of the original work itself rather than the imaginative acts which gave rise to those parts of the original work. Thus, it seems that Devitt should be construed as holding the following view: although the causal network that underlies the occurrence of a fictional name in a sentence about a fictional work ultimately stretches back to the imaginative acts from which the sentences of the original work flowed,
portions of Shakespeare's play *A Midsummer-Night's Dream*.

As an earlier remark indicated, the primary relevance of the foregoing disputes about the referential status of fictional names lies in their bearing on the referential status of the distinctively literary signifiers contained in a fictional work, especially the larger units that encompass whole blocks of sentences. If such units did indeed function referentially, it would be reasonable to suppose that they referred to complex states of affairs. However, in order for there to be such a state of affairs there must be one or more entities that either exhibit specific properties or stand in specific relations to each other, and that is just what writers who belong to the hard-liner and moderate camps (with appropriate qualifications in the case of Searle) have denied is the case. If such writers were correct in denying that there are such things as fictional entities, it would thus follow that distinctively literary signifiers do not function referentially. Hence, the views of the hard-liners and moderates warrant further

only the causal link between the occurrence of the name in the sentence used to talk about the work and the occurrences of that name in sentences contained within the work qualify as referential links—not all causal relations count as referential relations. Fictional names can, therefore, be said to function referentially when they occur in sentences used to talk about a fictional work, but not when they occur within the fictional work itself.
Russell believed that his views enabled him to account for the widespread intuition that sentences containing fictional names are meaningful without sacrificing either the axiom of existence or his ontological scruples, but there are other widespread intuitions about such sentences that Russell's account fails to capture. Recall that on Russell's account the c-sentence "Titania fell in love with a donkey" should be assigned roughly the following logical form: "There is an x such that x is a queen of the fairies, x is the only queen of the fairies, and x fell in love with a donkey." Since Russell would claim that there is no individual who possesses the property of being a queen of the fairies, the first conjunct of the foregoing construction would, according to Russell, turn out to be false, and when one conjunct in such a construction is false, the whole construction turns out to be false. Thus, the original c-sentence turns out to be false. In fact, Russell claimed that all sentences that contain a fictional name in what he called a "primary"—as opposed to a "secondary"—occurrence turn out to be false. Thus, a myriad number of c-sentences, sentences like "Hamlet was the Prince of Denmark,"
"Oedipus killed his own father," "Sherlock Holmes lived on Baker Street," etc., all turn out to be false on Russell's account.

Most people would, however, regard all of the foregoing sentences as true. Such sentences are, as John Woods puts it, "bet-sensitive."\(^{194}\) Suppose, for example, that, Smith says that the sentence "Sherlock Holmes lived on Baker street" is true, and, in response, Jones bets that the sentence is false, that Holmes lived on Berczy Street instead. As Woods observes, Smith would obviously win the bet.\(^{195}\) As Woods goes on to observe, not every class of sentences containing so-called "empty singular terms" is, however, bet-sensitive: "As between 'The present king of France is bald' and 'The present king of France is vigorously hirsute' there is no rational wager to make."\(^{196}\) Russell's account suggests, however, that a sentence like "Sherlock Holmes lived on Baker Street" should be analyzed in precisely the same way as "The present king of France is bald."\(^{197}\) Thus, his account provides no way of distinguishing bet-sensitive

---


\(^{195}\) Woods, p. 13.

\(^{196}\) Woods, p. 13

sentences from those that are not bet-sensitive.

What has been said during the foregoing paragraphs about c-sentences would, moreover, be equally applicable if Russell's account were applied to f-sentences. For example, according to such an account, when, at the outset of the second act of A Midsummer-Night's Dream, one of the fairies says, "Oberon is passing fell and wrath," the line should be construed as an assertion that there is an x such that x is a king of the fairies, x is the only king of the fairies, and x is exceedingly angry. Given such an analysis, the line turns out to be false; but that seems highly unintuitive. Moreover, the proposed analysis does not even seem to capture the meaning of the original line, a point which becomes even more apparent in cases where an f-sentence consists of a name concatenated with an interjection. At the end of Julius Caesar, for example, Dardanius has a line that consists entirely of the words "O

198 Since Russell does not draw a distinction between c-sentences and f-sentences in either of the discussions of fictional names contained in "On Denoting" and Mathematical Philosophy, it is not certain that he intended his account to cover both cases—although the very fact that he does not bother to draw such a distinction might be taken as evidence that he believed the two sorts of sentences should be treated in the same way.

The notion that such a line should be construed as an assertion that there is an x that uniquely satisfies a description of some sort seems implausible to the point of absurdity.

During the course of the last several pages, a number of criticisms of Russell's account have been advanced: it fails to provide any means of distinguishing bet-sensitive sentences from sentences containing empty singular terms that are not bet-sensitive, it yields highly unintuitive verdicts about the truth-value of c-sentences, and when applied to f-sentences, it produces a variety of counterintuitive results. While an account like that advanced by Devitt does not fall prey to the first of the foregoing criticisms and yields more intuitive verdicts concerning the truth-value of c-sentences, it nevertheless yields

---


201 As Searle notes, Russell's theory produces counter-intuitive paraphrases when it is applied to any non-assertional type of utterance like a command or a question regardless of whether it contains any fictional names (Searle, Speech Acts, pp. 160-61). Of course, it might be objected that Russell intended his analysis of referring expressions to apply only in the case of assertional utterances. Nevertheless, non-assertional utterances do contain referring expressions, and some account must be provided for those occurrences. Russell's account either provides a counter-intuitive analysis of such occurrences, or it provides no analysis at all.
equally if not more unintuitive results with respect to f-sentences.

As earlier remarks indicated, Devitt links the occurrence of a fictional name in a sentence used to talk about a fictional work to relevant portions of the original work itself.\textsuperscript{202} He goes on to say that such a sentence should be counted as true or false depending on what is contained in the relevant portions of the work.\textsuperscript{203} For example, since the Sherlock Holmes stories contain sentences like "Sherlock Holmes lived on Baker Street," when Smith asserts the corresponding sentence in talking about those stories, what he has said is true. Sentences that contain what Devitt calls a "failed" name, an empty name that is not causally linked to a fictional work, are, however, normally untrue.\textsuperscript{204} Thus, Devitt's account provides a way of marking the distinction between sentences that are bet-sensitive and those which are not, and it yields intuitive verdicts concerning the truth-value of sentences used to talk about fictional works.

\textsuperscript{202} See above, p. 195-96.

\textsuperscript{203} Devitt, p. 175.

\textsuperscript{204} Devitt, p. 175.
Devitt maintains, however, that sentences contained within a fictional work are all truth-valueless, i.e., they are neither true nor false. The problem with such a view is that it rules out the possibility of saying, for example, that various portions of Iago's speeches are false, while Desdemona's lines are true; but surely that is counter-intuitive. Iago lies; hence what he says must be false. Desdemona does not lie; what she says must be true. According to Devitt's account, however, the lines are all on an equal footing as regards their truth-value.

In addition to the foregoing problem, Devitt's account either assigns Russellian-style paraphrases to sentences of a fictional work that contain fictional names, or it assigns them paraphrases that possess even less plausibility, depending on whether or not the casual links between those occurrences of the names and the imaginative acts which gave rise to them are assumed to count as referential relations. If Devitt does not think that such causal links qualify as cases of "F-designation" or some other suitable sort of referential relation, it seems that he will have to embrace a Russellian-style account in order to explain the

---

205 Devitt, p. 171.

206 As earlier remarks have indicated, it is perhaps possible to construe Devitt as holding either view; see above, fn # 192, pp. 195-96.
fact that sentences of the sort being considered are not meaningless.

If Devitt does, however, count such causal links as cases of reference, it will turn out that the occurrence of a fictional name in a sentence of fiction refers to a particular imaginative act of that sentence's author, a process in the author's mind. Hence, the sentence will have to be assigned a paraphrase that mentions—not a fictional entity of some sort, nor even a set of properties in the manner of a Russellian-style description—but, instead, a particular mental state. Since Devitt maintains that, with the exception of the fictional names employed, the author of a fictional work "takes over the full vocabulary [of ordinary discourse] and uses it in the same sense," a sentence like Lodovico's "Cassio rules in Cyprus" would have to be assigned a paraphrase like "Mental-state number 53 rules in Cyprus." The counter-intuitive character of Russellian-style paraphrases obviously pales in comparison with the absurdity of what appears to be Devitt's only alternative.

An account like that offered by Searle suffers from much the same defects as does Devitt's account. According

207 Devitt, p. 171.

to Searle:

What distinguishes fiction from lies is the existence of a separate set of conventions which enables the author to go through the motions of making statements which he knows to be not true even though he has no intention to deceive.209

Thus, Searle thinks that f-sentences are either false or truth-valueless. As remarks have already indicated, either view faces serious problems.

With respect to the problem about how f-sentences that contain fictional names can be meaningful even though the names do not refer, Searle has little to say. Whereas Devitt never explicitly denies that such occurrences of fictional names function referentially, which leaves open at least the possibility that he might choose to extend the notion of "F-designation" to cover the use of fictional names in such contexts, Searle explicitly denies that fictional names contained in f-sentences refer.210 In view of the fact that he vigorously attacks Russell's theory of descriptions in his book Speech Acts, it seems unlikely that he would wish to pursue that option either--but how else could the meaningfulness of such f-sentences be explained, given

209 Searle, Expression, p. 67.

210 See above, pp. 193–94.
the assumption that the fictional names they contain do not refer? Searle seems to think that the problem can be dealt with by saying that although such uses of fictional names do not really refer, they "pretendedly" refer; but what do "pretended" references amount to, and how are sentences that contain "pretended" references to be understood? Without answers to such questions, Searle's account remains obscure at best.

Curiously enough, it appears that Searle could have easily avoided at least the foregoing problem. As earlier remarks have indicated, Searle's claim that the fictional names contained in f-sentences do not refer appears to be based on his contention that the required referents do not exist until after the author has completed his work. However, when he discusses the axiom of existence in *Speech Acts*, Searle explicitly qualifies the claim that "Whatever is referred to must exist," by adding that "'Exist' has to be construed tenselessly. One can refer to what has existed or what will exist as well as to what now exists." If it

---


212 See above, p. 193.

is possible in ordinary discourse actually to refer to future existents, what prohibits the author of a fictional work from actually referring to the fictional objects that will have come into existence by the time he has finished his work? Why must he only pretend to refer to them?

Perhaps Searle thinks that since an author is only pretending to make assertions, ask questions, issue commands, etc. when he writes the f-sentences contained in a fictional work, he must also be merely pretending to refer. Surely it is possible, however, actually to refer to objects while making a pretended assertion. Suppose, for example, that Jones pretends to assert that Richard Nixon was never president of the United States. Even though Jones is only pretending to make an assertion, he nevertheless actually refers to Richard Nixon and the United States. Thus, there appears to be no real need for Searle to deny that fictional names function referentially in f-sentences.

So far, very little has been said about the plausibility of the various views advanced by writers within the liberal camp, and they are not without their problems.\(^\text{214}\) In

\(^{214}\) For example, once a theory admits fictional objects into its ontology, it may also have to admit impossible objects like the round square. Moreover, a theory that allows f-sentences to be true or false must develop some way of blocking certain sorts of inferences that proceed from claims about what is true in fiction to claims
light of what has been said about the plausibility of the various alternatives to those views, a lengthy discussion of such problems does not, however, seem necessary. Each of the three major alternatives to the liberal camp faces what would seem to be the most serious objection to any account of fictional discourse; each of the three positions is incompatible with a number of basic, widely shared intuitions about the character of such discourse. Of course, it may not be possible to accommodate all such pretheoretic intuitions on any account, but it seems that, at the very least, an adequate account ought to preserve most of them. The account advanced by the hard-liners and the accounts advanced by both groups within the moderate camp all fail to satisfy that criterion, and that would seem to constitute sufficient grounds for regarding them as untenable.

The remainder of the present study is, therefore, based on the assumption that any adequate account of fictional discourse must be built around the two claims that are common to all versions of the liberal view: 1. there are

about what is true outside of fiction, e.g., an inference that would proceed from the claim that time-travel is possible in certain works of science fiction to the claim that time-travel is actually possible. Woods discusses a broad range of such problems and considers the relative merits of attempting to solve them by employing various types of so-called "deviant" logics.
such things as fictional objects, and 2. regardless of whether they are contained in c-sentences or f-sentences, fictional names refer to such objects. Subsequent remarks do not, however, take a stand on any of the questions that separate the various versions of the liberal view, e.g., the question of whether fictional objects exist or subsist, the question of whether fictional objects constitute kinds or particulars, etc. \(^{215}\) Thus, subsequent remarks assume that the most plausible account of fictional discourse will turn out to be some version of the liberal view, without attempting to determine exactly which version that is.

Once the referential status of various strictly linguistic units like fictional names has been accepted, it becomes clear that the larger distinctively literary units contained in a fictional work also function referentially. Consider, for example, the opening passage of D. H. Lawrence's *Women in Love*:

> Ursula and Gudrun Brangwen sat one morning in the window-bay of their father's house in Beldover, working and talking. Ursula was stitching a piece of brightly-coloured embroidery, and Gudrun was drawing upon a board which she held on her knee. They were mostly

\(^{215}\) The foregoing questions are briefly discussed above, pp. 189-90.
silent, talking as their thoughts strayed through their minds.\textsuperscript{216}

Just as the two fictional names, "Ursula" and "Gudrun," may be said to refer to two fictional personages, the whole first sentence may be said to refer to a particular state of affairs, a state of affairs that consists in two fictional personages each possessing the properties of sitting in the window-bay of their father's house, working, and talking. Likewise, the second and third sentences go on to refer to additional states of affairs involving Ursula and Gudrun. When taken as a whole unit, the passage may thus be said to refer to a complex state of affairs that consists in the amalgam of the less complex states of affairs referred to by each sentence of the passage.

The referential features of the distinctively literary units contained in a dramatic work are even more complex. Consider, for example, the following lines from Shaw's \textit{Arms and the Man}:

\begin{quote}
Catherine: Sergius is the hero of the hour, the idol of the regiment.
Raina: Tell me, tell me. How was it! [Ecstatically] Oh, mother, mother, mother!
\end{quote}

\begin{footnotesize}
\begin{enumerate}
\end{enumerate}
\end{footnotesize}
Catherine's line refers to a particular state of affairs, a state of affairs that consists in a fictional personage named "Sergius" having the property of being the hero of the hour, but that does not exhaust the referential function of the line. Because the line is to be uttered by Catherine, it may also be said to refer to a state of affairs in which a fictional personage named "Catherine" has the property of uttering the words in question. Likewise, Raina's line refers to a state of affairs in which a fictional personage named "Raina" has the property of uttering various other words. Moreover, the passage implicitly contains various sights and non-verbal sounds, which also serve a referential function, e.g., the passage implicitly contains a sequence of sights that refer to a state of affairs in which Raina pulls Catherine down onto an ottoman. Thus, at least one of the primary uses of the strings of mixed percept-images which constitute the distinctively literary signifiers in a dramatic work is that of reference.

Reference does not, however, constitute the only function of distinctively literary signifiers. As Searle points out:

serious (that is, nonfictional) speech acts can be conveyed by fictional texts, even though the conveyed speech act is not represented in the text. Almost any important work of fiction conveys a "message" or "messages" which are conveyed by the text but are not in the text. 218

For example, Martin Esslin has contended that in the climactic sequence of Eugene Ionesco's The Lesson, which depicts the sexually sadistic stabbing of a female student by her male teacher:

What Ionesco is saying is that even behind so apparently harmless an exercise of authority as the teacher-pupil relationship, all the violence and domination, all the aggressiveness and possessiveness, the cruelty and lust are present that make up any manifestation of power. 219

Thus, the sequences of mixed percept-images that constitute the distinctively literary units contained in a dramatic work may be used to make a thematic statement, as well as to refer to a particular fictional state of affairs.

218 Searle, Expression, p. 242.
219 Esslin, p. 119.
The thematic function of distinctively literary units has, in fact, constituted the primary focus of much—if not most—of the previous work done by Structuralist critics. Such work has led to a number of significant insights and is, therefore, of undeniable value. However, it has also tended to encourage an almost obsessive concern with thematic phenomena, thereby obscuring the fact that literary units possess uses over and above their referential and thematic functions. That tendency appears especially problematic because at least one of those additional functions seems every bit as important as do the referential and thematic functions—if not more so—namely, the use of literary units to arouse emotional responses. While the climactic sequence in Ionesco's The Lesson, for example, may well possess the thematic function discussed by Esslin, it obviously also functions as a means of arousing a sense of horror and revulsion.

Despite the obvious importance of such emotive uses of literary signifiers, when a Structuralist like Greimas or Barthes asks himself why a particular passage is contained in a literary work, he is almost sure to talk about its thematic value rather than its value as a means of creating suspense, arousing fear, dissipating emotional tension,
etc. As earlier remarks have indicated, the later Wittgenstein criticized philosophy of language up through the early part of the twentieth-century for exhibiting a myopic concern with the referential uses of language. Unfortunately, much of the work that has been done by Structuralist critics could, likewise, be criticized for exhibiting a myopic concern with thematic matters.

Despite the general tendency towards an excessive concern with thematic matters, some Structuralists have, however, managed to formulate a more balanced approach. In his commentary on Baudelaire's *les Chats*, for example, Michael Riffaterre suggests that "The characteristic common to [poetic] devices must be that they are designed to draw

220 In the entirety of his 295 page book *Structural Semantics: an Attempt at Method*, Greimas mentions the affective features of a work only once (p. 295), and then only as a by-product of the spacing of narrative events or functions; A. J. Greimas, *Structural Semantics: an Attempt at Method*, trans. Daniele McDowell, Ronald Schliefer, and Alan Velie (Lincoln: University of Nebraska Press, 1983). Moreover, his 39 page analysis of the literary works of Georges Bernanos at the end of the book contains no mention of emotional effects at all (pp. 257-95). Even more surprising is the absence of comment on emotional effects in Barthes's study of Racine's plays; Roland Barthes, *On Racine*, trans. Richard Howard (New York: Octagon Books, 1977). Barthes mentions various emotions during the course of his study, but they are always treated as the focus of thematic concerns rather than effects that the plays might have on readers or audience members.

221 See above, p. 170.
responses from the reader," and his analysis of Baudelaire's poem makes frequent reference to the work's ability to create emotional effects like suspense and shock.\(^{222}\) Some of Todorov's essays also exhibit a concern for the emotive uses of literary units.\(^{223}\) If Structuralism is to fulfill its potential in the field of literary criticism, it must lean in the direction of writers like Riffaterre and Todorov, who exhibit a concern for the emotive as well as the thematic function of literature—especially in the case of dramatic works, where the importance of emotional effect cannot be overestimated.

Although distinctively literary units may exhibit still other functions, it seems likely that the referential, thematic, and emotive uses constitute the primary functions of such units. As such, those functions will presumably play a crucial role in the characterization and individuation of the signified component of distinctively literary units. Such a characterization could be expected to include a specification of the way or ways in which a given signifier found in a particular work is used. Where one of those uses


\(^{223}\) See, for example, Todorov, Fantastic.
is referential, the characterization should contain a description of the intended referent—of the fictional state of affairs to which the signifier in question refers. Where one of the uses is thematic, the characterization should include a specification of the thematic statement conveyed; and where one of the uses is emotive, it should indicate the sort of emotional response evoked.

In addition, the characterization of the signified component might be expected to include a specification of the syntactic category to which the unit in question belongs. For example, a signifier in a dramatic work might constitute the inciting incident, a complicating incident, the crisis, the climax, the denouement, etc. Naturally, the claim that literary units can be assigned to some sort of syntactic categories and the claim that the foregoing terms designate such categories both require a good deal of discussion. Since those claims will be discussed at some length during the course of the next chapter, little more will, however, be said about them at this point.\footnote{See below, pp. 267-75.}

As in the case of strictly linguistic units, the characterization of the signified component of a distinctively literary unit should also indicate whether the unit in
question is basic or derived. If the unit in question is derived, the characterization should also specify the way in which the functions of the unit's components and their mode of combination help to determine the functions possessed by the unit as a whole.

The identity of the signified component of a distinc-tively literary unit may thus be said to be based—in part—on the syntactic and functional differences that distinguish it from (1) the signified components of the units that constitute its syntagmatic environment within the work in question and (2) all the signifieds to which it is paradigmatically related due to partial similarities of either a syntactic or functional nature, signifieds which are associated with units contained in other literary texts. The syntactic and functional features mentioned in the characterization of the signified component of a literary unit thus serve to indicate the differences that separate it from all other signifieds, differences which partially account for its identity.

The preceding remarks concerning the paradigmatic relations between literary signifieds, and earlier remarks concerning the paradigmatic relations between literary signifiers, appear to shed light on one of the most important concepts discussed by previous Structuralist writers—the
concept of "intertextuality." As Culler observes in his book *In Pursuit of Signs*, virtually all previous writers have agreed "that literary works are to be considered not as autonomous entities . . . but as intertextual constructs: sequences which have meaning [in virtue of their] relation to other texts." 225 Although such writers rarely draw an explicit connection between their views concerning intertextuality and Saussure's remarks about paradigmatic relations, there appears to be a clear way of linking the two. Together, the claim that paradigmatic relations link the signifier and the signified associated with a literary unit contained in one work with signifiers and signifieds associated with units contained in other works—and the claim that the identity of every signifier and signified derives in part from the institutionalized differences that distinguish it from its paradigmatic alternatives—imply that the identity of literary signifiers and signifieds, and hence the identity of a literary work as a whole, is partially constituted by a complex network of intertextual relations.

---

225 Jonathan Culler, *In Pursuit of Signs: Semiotics, Literature, Deconstruction* (Ithaca: Cornell University Press, 1981), p. 38. As Culler notes, some previous writers have used the term "intertextuality" as a means of referring exclusively to the phenomena of literary allusion and literary borrowing, while others have used the term to cover relations between texts that are more general than the specific relations of allusion and borrowing (Culler, *Signs*, pp. 104, 116-18). Only the latter sense of the term will be of concern in what follows.
When Saussure's views concerning paradigmatic relations are applied within the domain of literature, the result is intertextuality.

Naturally, the notion that the meaning and identity of a literary work derives in part from the phenomenon of intertextuality has led previous writers to claim that intertextuality is also involved in the perception of literary works. In his discussion of the topic, Culler, for example, quotes Laurent Jenny's statement that "Outside intertextuality the literary work would be quite simply imperceptible, in the same way as an utterance in an as yet unknown language." 226

The notion of intertextuality thus encompasses a psychological as well as a metaphysical dimension, and Saussure's remarks concerning paradigmatic relations seem to be as relevant to an account of the former as they are to an account of the latter. Just as a foreigner will be unable to identify and grasp the value of the various linguistic units employed in a given language until he comes to know which acoustic and conceptual differences have been institutionalized as a means of distinguishing signifiers and signifieds, a person will be unable to identify and grasp the

226 Culler, *Signs*, p. 104.
value of the various literary units contained in works of a given kind or genre until he comes to know which perceptual and conceptual differences have been institutionalized as a means of distinguishing signifiers and signifieds within that genre.

Unfortunately, a number of previous writers, especially the deconstructionists, seem to have misunderstood the role that intertextual relations play in the perception and understanding of literary works. Deconstructionist discussions of differential relations often seem to suggest that since the identity of a given linguistic or literary unit is partly determined by the differences that distinguish it from all other units, a person would have to take account of its relation to every one of the other units in order to understand it.\(^{227}\) Such a view rests on an assumption that the paradigmatic relations between particular units possess a psychological importance that they do not have.

It is not knowledge of the differential relation between a given unit and any other particular unit—much less knowledge of the relations it bears to every other particular unit—that is essential to a person's perception and understanding of that unit. It is rather knowledge that certain

\(^{227}\) See, for example, Derrida, Positions, pp. 26, 33 and Miller, "Steven's Rock," pp. 333-34.
types of differences have been institutionalized within the relevant language or genre that is crucial. The psychological value of any particular paradigmatic relations is that acquaintance with them may lead a person to recognize which of the various types of differences between units have been institutionalized within the relevant system.

A person could not, for example, identify and grasp the value of the English word *bin* without knowing that the difference between a *b*-sound and a *p*-sound has been institutionalized in English. Comparison of various pairs of words like *bin* and *pin*, *bet* and *pet*, etc. may lead a person to the realization that the difference in question has been institutionalized. However, any set of pairs of the appropriate type would serve the purpose. The ability to identify and grasp the value of *bin* does not, therefore, depend on a person's having taken account of the particular relation between it and *pin*—much less the particular relation it bears to every other English word that begins with a *p*-sound.

Likewise, a person could not fully grasp the value of Desdemona's death at the end of *Othello* without knowing that in tragedy a killing is not always used to evoke a feeling of pity—without knowing that tragedy has, for example, institutionalized a contrast between killings that evoke pity and killings that evoke a sense of righteous satisfaction. Comparison of Desdemona's death with that of Claudius at the
end of *Hamlet* might well lead a person to a realization that such a contrast has been institutionalized. However, portions of a number of other plays would have served as well. The ability to understand the value of Desdemona's death does not, therefore, depend on a person's having taken account of the particular relation between it and Claudius' death—much less the relation it bears to the killings depicted in every other tragedy.

The primary psychological importance of intertextuality thus lies in the fact that it is comparison of various works within a genre that enables a person to infer the appropriate set of institutionalized difference-types, knowledge of which guides his perception and understanding of other works within that genre. Once the real role that intertextuality plays in the perception and understanding of literary works has been recognized, it becomes clear that intertextuality does not, however, entail the radically holistic and sceptical epistemological position espoused by deconstructionists.

Thus far, attention has been primarily focused on the role that paradigmatic relations play in determining the identity of literary units; such relations set up a network of institutionalized differences that insure the uniqueness of each of a literary unit's two components. However, syntagmatic relations also play a role. In fact, such
relations appear to play a double role in helping to determine the identity of literary units. Insofar as they separate the signifier and signified associated with one unit in a syntagmatic sequence from the signifiers and signifieds associated with all the other units of the sequence, they serve as an additional means of establishing the network of institutionalized differences that insure the uniqueness of signifiers and signifieds. However, it appears that syntagmatic relations also play a crucial role in establishing the depth relation between a given signifier and its corresponding signified, and, as earlier remarks have indicated, depth relations constitute the other major source of a literary unit's identity.²²⁸

In the case of language, for example, a person decides whether the acoustic sequence *rose* constitutes part of one sign rather than another on the basis of whether it occurs, for instance, in the syntagmatic frame "Mary picked a . . ." rather than "The sun . . . in the east." The paradigmatic contrasts between *rose, nose, raise*, etc. constitute the primary source of the uniqueness of the unit's signifier, and the paradigmatic contrasts between "tulip," "daisy," "violet," etc. on the one hand—and "fell," "remained," "set," etc. on the other—insure the uniqueness of the

²²⁸ See above, p. 178.
signifieds that could be involved; but it is the syntagmatic context that specifies which of the possible depth relations actually obtains.

Likewise, when Hamlet kills Claudius, a person comes to realize which of various possible literary units is involved on the basis of its syntagmatic frame. The paradigmatic contrasts between the sequence of percept-images associated with that unit and perceptually similar images contained in other plays insure the uniqueness of that unit's signifier, and the paradigmatic contrasts between killings that evoke righteous satisfaction, e.g., Barabas' death in Marlowe's Jew of Malta, Lorenzo's death in Kyd's The Spanish Tragedy, etc., on the one hand—and the contrasts between killings that evoke pity, e.g., Desdemona's death, Julius Caesar's death, etc., on the other—insure the uniqueness of the various signifieds that could be involved; but it is the syntagmatic frame that specifies which of the various possible correlations actually obtains. Claudius's killing evokes righteous satisfaction because of its syntagmatic relation to the reprehensible deeds that he has committed during earlier portions of the play and to the reports of reprehensible deeds he has committed prior to the beginning of the play.

Given the obvious resemblance between the mixed percept-images contained in dramatic works and the states of
affairs to which they refer, the question arises whether the combination of a dramatic signifier and a signified can be said to constitute a sign rather than a symbol. In the passage from the *Course* cited earlier, Saussure seems to suggest that meaningful phenomena can be sharply separated into two mutually exclusive groups, those that constitute signs and those that constitute symbols, depending on whether "there is the rudiment of a natural bond between the signifier and the signified." However, in the same section of the *Course*, Saussure seems to indicate that the distinction between signs and symbols does not constitute such a sharp and easily drawn dichotomy after all:

One remark in passing: when semiology becomes organized as a science, the question will arise whether or not it properly includes modes of expression based on completely natural signs, such as pantomime. Supposing that the new science welcomes them, its main concern will still be the whole group of systems grounded on the arbitrariness of the sign. In fact, every means of expression used in society is based, in principle, on collective behavior or—what amounts to the same thing—on convention. Polite formulas, for instance, though often imbued with a certain natural expressiveness (as in the case of a Chinese who greets his emperor by bowing down to the ground nine times), are nonetheless fixed by rule; it is

---

229 See above, p. 143.
Thus, it would appear that the distinction between signs and symbols does not sharply bifurcate the set of meaningful phenomena into two disjoint subsets after all; rather, the two terms designate the terminal nodes of a continuum.

Admittedly, the resemblance between mixed percept-images and the states of affairs to which they refer seems to place dramatic units towards the symbol-end of the continuum. However, as earlier remarks have indicated, the signifieds associated with such percept-images do not consist entirely of information about the intended referent. They also contain information about the thematic and emotive functions of those signifiers, and the links between those functions and the signifiers with which they are associated do not seem to constitute a "natural bond." As earlier discussion has suggested, a given signifier could be associated with quite different thematic and emotive functions depending on its syntagmatic context, the genre of the work in which it is contained, etc. As a result, there would appear to be sufficient justification for classifying the units contained in dramatic works as signs after all.

230 Saussure, p. 68.
One further literary implication of the Saussurian approach warrants discussion. In his brief sketch of a program for future linguistic work, which was cited earlier, Saussure remarked that it would be necessary to identify and classify not only the set of words employed within a given language, but also the subunits which they contain and the larger units which, in turn, contain them. Implicit in Saussure's remark is, of course, the notion of hierarchically arranged levels of units; words must be assigned to one hierarchical level of linguistic structure, the subunits to other levels, and the larger units to still other levels. Saussure's remark thus constitutes an early formulation of what has become a fundamental tenet of all subsequent work in linguistics: identification of the hierarchical levels of linguistic structure constitutes an absolutely essential step in the study of language.

From the outset, Structuralist critics realized that the same holds true in the case of literature. Indeed, Barthes claimed that the same will hold true in any case of structural analysis:

In order to conduct a structural analysis, it is . . . first of all necessary to distinguish several levels or instances of

---

231 See above, p. 162.
Thus, critics like Barthes realized that the success of the Structuralist enterprise would ultimately depend on whether or not theorists succeeded in isolating a hierarchy of levels in the realm of literature analogous to those isolated by linguists. Unfortunately, previous writers have been largely unsuccessful in their attempts to isolate such a literary hierarchy.

Part of the problem lies in the fact that most previous writers have followed the Russian formalist Vladimir Propp in focusing on only two putative levels of literary structure. In his *Morphology of the Folktale*, completed in 1928, Propp argues that it is possible to account for the structure of all Russian fairy tales in terms of thirty-one possible events or, as he calls them, "functions," and seven possible character-types or, as he calls them, "spheres of action." Each character-type is characterized in terms of his typical functional role, e.g., the "villain" pursues


the "hero" and they eventually become involved in some form of fight or struggle; and each event is characterized in terms of a particular sort of activity involving one or more of the character-types, e.g., the "struggle" is marked by the fact that "the hero and the villain join in direct combat."234

While the theories advanced by Propp and by Structuralist writers who have adopted much the same approach, e.g., Greimas, are of substantial value in a number of other respects, they provide little or no insight into the hierarchical structure of literary works.235 At most, they could be claimed to have isolated only two distinct hierarchical levels—the level on which the various "functions" are located and the level on which the character-types are located—and surely more than just two levels will need to be distinguished if all the complexities of literary structure are to be explained.

It does not, however, appear that they have even succeeded in isolating two levels, at least not levels that are at all analogous to the levels of linguistic structure. One

234 Propp, pp. 79 & 51 respectively.

235 See Greimas, Structural Semantics. For a survey of various other writers whom Scholes refers to as "the progeny of Propp," see Scholes, pp. 91-117.
of the obvious features of the integrative relations between linguistic units located on adjacent hierarchical levels is that each unit on one level is sequentially coextensive with, and therefore completely contains, the units on the lower-level out of which it is composed; a morpheme is temporally coextensive with the phonemic sequence out of which it is made, a word is temporally coextensive with the morphemic sequence out of which it is made, etc. However, since the "hero-type" or the "villain-type" typically appears in a number of distinct narrative events, events which may not themselves be temporally contiguous, the same sort of relation obviously does not obtain. Propp's functions and character-types are thus not related to each other in anything like the way that hierarchically adjacent linguistic units are related.

Even Barthes's analysis, which constitutes one of the better efforts to date, seems largely unsatisfactory. In his "Introduction to the Structural Analysis of Narratives" Barthes posits units located on three different hierarchical levels: functions, sequences, and actions. A function consists of "any segment of the story which can be seen as the term of a correlation," a sequence consists of "a small group of functions," and an action consists in a "major articulation of praxis ([e.g.,] desire, communication,
Barthes illustrates the relation between functions and sequences by means of the following example. The activities of ordering a drink, obtaining it, drinking it, and paying for it are each correlated with each other in the following way: ordering a drink implies one of two consequences—obtaining it or not obtaining it—obtaining it likewise implies drinking it or not drinking it, and drinking it implies paying for it or not paying for it. Since each activity is correlated with at least one other term in the series, each constitutes a function; and since the series of functions as a whole constitutes an obviously closed sequence, it being impossible to put anything before the order or after the payment without moving out of the homologous group "Having a drink," the series constitutes a sequence.\footnote{Barthes, "Narratives," pp. 89, 101, & 107 respectively.}\footnote{Barthes, "Narratives," p. 101.} If that sequence were to occur as part of a narrative, it might, in turn, be combined with other sequences as part of a major articulation of praxis like a struggle, which would constitute an action.

Unfortunately, Barthes goes on to insist that "the
units of a sequence, although forming a whole at the level of that very sequence, may be separated from one another by the insertion of units from other sequences." Thus, Barthes's analysis falls prey to the same sort of objection raised against writers like Propp and Greimas. Moreover, Barthes provides no criteria that might be used to tell when a string involving an attempt to satisfy a desire, successfully communicate, or win a struggle is to count as "a major articulation of praxis," and hence an action. Presumably a person who orders a drink has a desire of some sort, to quench his thirst or to get drunk, etc. Without some sort of criteria that will help to identify a major articulation of praxis, it is not clear why Barthes's example of a sequence, "having a drink," does not constitute an action. Finally, even if Barthes had succeeded in clearly identifying three distinct hierarchical levels, three levels still does not seem sufficient to account for all the complexities of narrative structure.

The general lack of success in isolating a series of hierarchical levels capable of accounting for the full range of structural phenomena associated with any form or genre of literary works obviously constitutes a major stumbling block

for the Structuralist enterprise. Some writers have even given up the search for such a hierarchy as hopeless. In The Semiotics of Theatre and Drama, for example, Kier Elam rejects as hopeless any attempt to locate a comprehensive hierarchy of levels in the areas of drama and theatrical performance—the very phenomena with which the present study is most concerned.239 The last two chapters of this study argue, however, that, contrary to the claims of writers like Elam, the contents of Aristotle's Poetics provides a firm foundation for the description of the hierarchical structure of drama and theatre. Isolation of such a hierarchy will help to remove one of the chief obstacles to the success of the Structuralist enterprise.

239 Elam, p. 48.
CHAPTER IV
LINGUISTICS AND LITERATURE: THE GENERATIVE APPROACH

As was mentioned at the outset of the foregoing chapter (p. 83), previous writers have disagreed about the literary relevance of the various concepts and methodological procedures that have been developed by generative linguists. One source of those disagreements may well be the unsettled state of generative linguistics itself. Since the inception of the generative approach, linguists have formulated a variety of different generative models of language, and they have disagreed about the merits of rival models so vociferously that one period has even been called the "period of the linguistic wars."\(^1\) Even when consideration is limited to the work of a single generative linguist, the formulation and revision of several different models may have to be examined.

In order to avoid the complexities that would be involved in a comprehensive survey of generative linguistics, the following discussion will be restricted in a number of

\(^1\) Newmeyer, p. 133.
ways. First, although there will be occasional references to the work of linguists other than Chomsky, his work will occupy the focus of attention. Second, the discussion will only deal with Chomsky's work on the phrase structure component of generative grammars. Moreover, because Chomsky's formulation of the rules included in a phrase structure grammar has undergone periodic changes, subsequent discussion will focus on the models of phrase structure grammar developed in only two of his major works, *Syntactic Structures* and *Aspects of the Theory of Syntax*. As a result of these limitations, little or nothing will be said about Chomsky's views concerning semantics and pragmatics, his work in generative phonology, his views on the overall organization of the various components needed for a complete description of language, his work on the development of the transformational component of generative grammars, his views on "universal grammar," innate knowledge, and a number of other important issues.²

² For a survey of the work being done in generative semantics, see Janet Fodor's book. For examples of the work being done in generative pragmatics, see the selections contained in Tuen A. van Dijk, ed., *Pragmatics of Natural Language and Literature* (New York: American Elsevier, 1976). For discussions of the other topics to which reference has been made, see Lyons, Chomsky, Newmeyer, and, of course, Chomsky's own works.
The fact that the preceding topics are being excluded from consideration in the present study should not, however, be taken as an indication that they are irrelevant to the concerns of literary Structuralism. Their exclusion simply reflects limitations of time and space. Chomsky's work on phrase structure grammar has been selected as the focus of concern for two main reasons. First, the general approach of generative linguistics can be most easily illustrated by a discussion of that work; consideration of the work that has been done in the areas of generative phonology, semantics, or pragmatics would require a much more technical discussion. Second, the literary applications of such work are more obvious than the applications of work that has been done on the other topics.

In order to fully appreciate the development and the importance of Chomsky's approach to language, it must be seen against the backdrop of its taxonomic predecessor, the American Structuralist or "Bloomfieldian" movement.

---

3 See, for example, van Dijk's discussion of the literary applications of generative pragmatics; Tuen A. van Dijk, "Pragmatics and Poetics," in van Dijk, pp. 23-57.

4 For discussions of Bloomfieldian linguistics, see, for example, Bierwisch, pp. 39-44, Ivlić, pp. 156-62, Lyons, Chomsky, pp. 28-34, and Newmeyer, pp. 1-17.
view of the fact that Chomsky studied under one of the major figures in that movement, Zellig Harris, similarities between the two approaches should come as no surprise.\(^5\) In fact, Chomsky has himself indicated that Bloomfieldian work provided the impetus behind his development of phrase structure grammar:

The concept of "phrase structure" grammar was explicitly designed to express the richest system that could reasonably be expected to result from the application of Harris-type procedures [those of segmentation and classification] to a corpus. Harris, and the other methodologists of the 1940's, were developing an approach to linguistic analysis that one can trace at least to Saussure.\(^6\)

In spite of the fact that the Saussurean, Bloomfieldian, and Chomskian approaches differ on a number of important issues, Chomsky's suggestion that the latter two can be seen as, in certain ways, an extension of the former is worth emphasizing; all three approaches represent attempts to specify the complex paradigmatic and syntagmatic relations contracted among the basic units of language. It is the continuing interest in accomplishing the foregoing task that links the various, diverse approaches manifested by twentieth-century linguists,

\(^5\) Newmeyer, p. 33.

establishing a sense of continuity among them. As the following discussion of Bloomfieldian syntactic work will show, the Bloomfieldians provided Chomsky with more than just an impetus, however; they provided a number of concepts that greatly facilitated his efforts to further specify the complex relations contracted among the basic units of language.

Unlike both Saussure and the Bloomfieldians, Chomsky believes, however, that a complete description of syntactic phenomena will have to go beyond the specification of syntagmatic and paradigmatic relations. Chomsky claims that there are two basic levels of syntactic structure, "surface structure" and "deep structure," and he believes the two levels are related to each other by way of "transformational" rules (Mind, p. 191). While he believes that paradigmatic and syntagmatic relations are capable of specifying the phenomena associated with surface structure, he does not believe that they are capable of specifying the phenomena associated with deep structure and grammatical transformations (Mind, p. 19). Although Chomsky's views concerning deep structure and grammatical transformations have been extremely influential, there are a number of contemporary linguists who advocate a non-transformational, generative approach to language; a brief survey of such work is contained in Newmeyer, pp. 236-48. In keeping with the limitations of the present chapter, little more will be said about deep structure and grammatical transformations.

As the results of Bloomfieldian syntactic work are about to be discussed, their attitude towards syntax should, perhaps, be mentioned. Heavily influenced by empiricist philosophy and behaviorist psychology, the Bloomfieldians believed that any truly "scientific" investigation should begin by describing those phenomena that are most accessible to direct observation. In the case of language, that naturally meant beginning with what is actually heard. As a result, Bloomfieldians believed that linguists should begin by examining phonemic structure and should proceed to an examination of successively higher levels of linguistic structure only after completing analysis of all lower
In 1933 Leonard Bloomfield introduced the notion of "immediate constituents":

Any English-speaking person who concerns himself with [the] matter, is sure to tell us that the immediate constituents of Poor John ran away are the two forms poor John and ran away; that each of these is, in turn, a complex form; that the immediate constituents of ran away are ran... and away; and that the constituents of poor John are... poor and John.

As Lyons has observed, the procedure of isolating immediate constituents is quite similar to "the traditional procedure of 'parsing' sentences into 'subject' and 'predicate,' and each of those, where appropriate, into words, phrases, and

levels. Naturally, such an attitude led them to focus on the relations contracted among phonemes and morphemes, thereby postponing analysis of the syntactic relations contracted between higher-level units. As Newmeyer observes, "the little syntactic work which WAS done [by the Bloomfieldians] was, in a sense, the result of 'cheating'--a complete morphemic analysis had never been worked out even for English" (Newmeyer, p. 9). Thus, while the Bloomfieldians did achieve some significant insights into the nature of syntactic phenomena, such insights were somewhat rare and were regarded as methodologically premature by the very movement that originated them. Discussions of the influence of empiricism and behaviorism on Bloomfieldian linguistics are contained in Newmeyer, pp. 3-17 and Lyons, Chomsky, pp. 25-34.

clauses of various types."\(^{10}\) Both procedures facilitate identification of the integrative or compositional relations that obtain between units located on adjacent hierarchical levels within the linguistic hierarchy, e.g., the relation between a phrase and the various words out of which it is composed.

Although Bloomfield did not develop the notion of immediate constituents much further, two of his students, Harris and Rulon Wells, went on to develop explicit procedures for identifying the immediate constituent structure of utterances.\(^{11}\) Among other things, those procedures enabled Harris and Wells to abstract "sentential formulae," sentence-length syntactic patterns, from utterances.\(^{12}\) For


\(^{11}\) Newmeyer, pp. 7-8.

\(^{12}\) Although the procedures developed by Wells and Harris enabled linguists to abstract roughly equivalent sentential formulae, the two methods exhibited some interesting differences. For example, Wells, like Bloomfield, worked from the top down; he started from the sentence as a whole and divided it into successively smaller constituents (Rulon Wells, "Immediate Constituents," *Language*, 23 [1947], 81-117). Harris, however, worked from the bottom-up; he classified individual morphemes into lower-level syntactic categories and then combined them to form units belonging to successively higher-level categories (Zellig Harris, "From Morpheme to Utterance," *Language*, 22 [1946], 161-83). In spite of the fact that the two procedures led to the formulation of roughly equivalent sentential formulae, the
example, the sentential formula \( \text{Adj (Adjective)} - \text{N(Noun)} - \text{V(Verb)} - \text{Adv(Adverb)} \) could be abstracted from the sentence *Poor John ran away.*

Like the Bloomfieldian approach, Chomsky's approach also enables linguists to specify the immediate constituent structure of utterances. Because of that similarity, as well as others, some linguists have, in fact, suggested that the former can be interpreted as a version of the latter,

Bloomfieldian's empiricist and behaviorist inclinations tended to focus attention on the differences in directional orientation. Harris's procedures were, of course, more in keeping with those inclinations.

Interestingly enough, various generative models exhibit much the same difference. Chomsky's phrase structure grammar generates utterances from the top-down, whereas categorial grammars, e.g., those developed by Yehoshua Bar-Hillel and Richard Montague, generate from the bottom-up; see Yehoshua Bar-Hillel, "A Quasi-Arithmetical Notation for Syntactic Description," *Language*, 29 (1953), 47-58, and Dowty, Wall, and Peters. At present, the implications of such differences in directional orientation seem unclear (see, for example, Janet Fodor's discussion of the matter, pp. 107-16). Little more will be said about the subject in what follows—see, however, the remarks contained in fn # 16, pp. 242-43.

13 Harris went on, during the 1950's, to work out additional procedures for stating the "formal relation among [sentential formulae], by virtue of which one sentence structure may be called the transform of another sentence structure (e.g. the active and passive, or in a different way question and answer)" (Zellig Harris, "Co-occurrence and Transformation in Linguistic Structure," *Language*, 33 [1957], 283). Harris's formulation of transformation relations and Chomsky's formulation of transformational rules are, of course, related. For a discussion of both the similarities and differences between the two, see Newmeyer, pp. 8, 23.
but Chomsky strongly disagrees:

the suggestion that the various theories of immediate constituent analysis might be interpreted as generative phrase structure grammars... certainly goes beyond what is explicitly stated by linguists who have developed these theories, and very likely beyond their intentions as well.14

One important difference between the two approaches lies in the fact that they utilize different formal devices to specify the immediate constituent structure of utterances.

Whereas the Bloomfieldians specified the immediate constituent structure of each utterance contained in a given corpus by compiling a list of sentential formulae, Chomsky utilizes rules. Thus, as R. H. Robbins puts it, "A contrast is drawn between the 'grammar of lists'... and the 'grammar of rules.'"15 That difference might not, at first glance, appear to be of much significance—it might even seem that the two means of specifying the immediate constituent structure of utterances are mere notational variants

14 Chomsky, Aspects, p. 205. Newmeyer cites several statements by Harris and Charles Hockett that seem to indicate some generative tendencies, but he goes on to point out that such statements do not accurately reflect their overall approach and were not further developed by them (Newmeyer, p. 37).

of one another—but, as the following rudimentary account of phrase structure rules will show, such is not the case.

A phrase structure grammar contains a set of rules that can be used to generate sentences from the top-down.16

16 As Janet Fodor has observed, there are at least three distinct, although related, uses of the term "generative" in generative linguistics (Janet Fodor, pp. 112-13). The first, and most basic, use has to do with whether or not a grammar is capable of specifying an infinite set of sentences, or formal descriptions of sentences, by finite formal means: "Technically, what makes generative grammar generative is that it recursively enumerates the infinite set of well-formed sentences of a language" (Janet Fodor, p. 112).

A grammar may be said to be "generative" in an additional sense if it specifies that infinite set by means of providing explicit procedures for constructing each member of the set (Janet Fodor, p. 113). The foregoing does not, however, constitute the only means of specifying such a set. "To use a test procedure (filter) to winnow out of the set of all strings those which are ill-formed" would, as Fodor observes, successfully specify the infinite set of well-formed sentences, and would, therefore count as "generative" in the first sense of the term but would, nevertheless, "not count, in this special [constructive] sense, as generating the well-formed sentences" (Janet Fodor, p. 113). Nevertheless, nearly all attempts to specify the infinite set of well-formed utterances of a language have so far employed the "constructive" method of doing so and have, therefore, been generative in both senses (Janet Fodor, p. 113).

A third sense of the term derives from its use as a means of distinguishing a grammar, or a component of a grammar, that provides a means of producing an infinite set of structures from a finite input from a so-called "interpretive" component—one which requires an infinite input in order to produce an infinite output (Janet Fodor, p. 113). While the phrase structure component of a Chomskian grammar is generative in this third sense, the transformational, phonological, and semantic components are all interpretive; "they simply define mappings from one infinite domain of
Each phrase structure rule specifies the syntagmatic structure of a syntactic unit's subcomponents in the following way: a symbol representing a particular syntactic paradigm is followed by an arrow, which is, in turn, followed by a sequence of symbols that represent the paradigmatic categories of the original unit's permissible subcomponents. Such representations to another" (Janet Fodor, p., 113). The fact that the phrase structure component is the only component of a Chomskian grammar that is generative in all three senses is, incidentally, another reason why it was picked to illustrate the generative approach.

Subsequent remarks in the main text will further illuminate the various uses of the term, but, in order to eliminate a potential source of confusion, a few supplementary comments concerning the second, constructive sense of the term are included here. The statement that a given set of rules "generates" a particular sentence means that they provide at least one way of explicitly deriving or constructing that sentence. The statement does not, however, mean that there may not be other ways of constructing the same sentence; nor does it imply that speakers or listeners necessarily follow the same derivational process that the linguist utilizes. In Chomsky's words, such statements "say nothing about how a speaker or hearer might proceed, in some practical or efficient way to construct" or decipher utterances (Aspects, p. 9). Thus, Chomsky denies that the differences in the directional orientation of generative grammars imply anything about the order of psychological computational processes occurring in the mind or brain of a language user as he speaks or hears utterances. In particular he denies that the top-down orientation of his approach implies that speakers or hearers utilize the same strategy in actual practice: "I see no plausibility at all to the assumption that the speaker must uniformly select sentence type, then determine subcategories, etc., finally, at the last stage, deciding what to talk about; or that the hearer should invariably make all higher-level decisions before doing any lower-level analysis" (Aspects, p. 197).
a rule may be interpreted to mean that in the derivation of an utterance, the leftmost symbol may be rewritten as the string of symbols to the right of the arrow, in the specified order. In addition, a notational device called the "parenthesis convention" enables a phrase structure rule to indicate that certain symbols are optional: if a symbol on the right-hand side of an arrow is enclosed in parentheses, it may be either included or omitted during the rewriting process.

The following examples should help to clarify the nature of such rules. Among the various phrase structure rules developed by Chomsky are, for example, the following: 1. S → NP + VP, 2. NP → (Det) + N, and 3. VP → V + (NP).

The first rule means that the symbol S, which stands for "sentence," may be rewritten as the string NP + VP, which stands for the concatenation of a noun phrase and a verb phrase. The second rule means that an occurrence of the

17 See Chomsky, Aspects, p. 66 and Baker, pp. 36-37.
19 As indicated above (p. 234), the exact formulation of generative rules often varies from one of Chomsky's books to the next. Many of the examples given in this section will consequently represent a synthesis of formulations contained in Syntactic Structures and Aspects. The foregoing rules, for instance, represent a synthesis based on pp. 26 & 110 of Syntactic Structures and pp. 68 & 72 of Aspects.
symbol NP may be rewritten as either N, which stands for "noun," or as Det + N, the concatenation of a determiner and a noun. The third rule means that an occurrence of the symbol VP may be rewritten as either V, which stands for "verb," or as V + NP, the concatenation of a verb and a noun phrase.

Given the information that the is a member of the paradigmatic class of determiners, man and ball are members of the class of nouns, and hit is a member of the class of verbs, the foregoing three rules can be used to generate the sentence The man hit the ball. The generative process involved in producing that sentence may be represented by a "sequential derivation" like the following one: (a) application of rule one replaces the initial symbol S with the string NP + VP, (b) application of rule two expands that string into [Det + N] + VP, (c) application of rule three yields [Det + N] + [V + NP], (d) reapplication of rule two yields [Det + N] + [V + [Det + N]], and (e) "lexical insertion" rules can be used to replace each of the "non-terminal" symbols in that string with appropriate "formatives," i.e., non-phonological specifications of actual words.

---

20 The above information would be contained in the lexicon, which is included, along with the set of phrase structure rules and a set of transformational rules, in the syntactic component of a Chomskian grammar (Aspects, p. 84).
thereby yielding the sentence The man hit the ball.\textsuperscript{21}

Phrase structure rules thus enable a linguist to generate a sentence by expanding the initial symbol $S$ into successively longer strings of symbols, which can finally be replaced with formatives. It should be noted that such rules often allow a linguist to derive a particular sentence in more than one way. In fact, the foregoing derivation does not represent the only means of generating The man hit the ball. After the first step, steps (b) and (c) could have been done in reverse order, expanding $NP + VP$ into $NP + [V + NP]$, and then expanding that into $[Det + N] + [V + NP]$.

As Chomsky observes, the use of brackets in the derivation of a sentence serves to indicate the boundaries between immediate constituent groupings.\textsuperscript{22} For example, the brackets in the above derivation indicate that the immediate constituents of the string $[Det + N] + [V + [Det + N]]$ are the

\textsuperscript{21} Chomsky, Syntactic Structures, pp. 26-27 and Aspects, p. 84. Chomsky uses the term "sentence" to refer to "strings of formatives rather than strings of phones" (Aspects, p. 16). It is the task of the phonological component of a grammar to associate each sentence or string of formatives generated by the syntactic component with a phonological representation (Aspects, p. 16). Although brackets do not appear in the Syntactic Structures version of the above derivation, they are used throughout Chomsky's later work; see, for example, Aspects, pp. 12-3. Their function will be discussed shortly.

\textsuperscript{22} Chomsky, Aspects, p. 12.
two strings **Det + N**, which is a noun phrase, and **V + [Det + N]**, which is a verb phrase, and that the immediate constituents of the latter are **V**, which is a verb, and **Det + N**, which is another noun phrase. In fact, the paradigmatic category of each constituent grouping is frequently indicated by labeling the brackets in the following way:

\[ \mathrm{NP}^{\text{Det + N}} \] + \[ \mathrm{VP}^{\text{V + [NP^{Det + N}]} } \].

Such a "labeled bracketing" is said to specify a sentence's "phrase marker."\(^{23}\) The phrase marker of a sentence can, alternatively, be specified by means of a "derivational tree" such as the one in figure number one on the next page.\(^{24}\)

One of the most important features of phrase structure rules is that, unlike the sentential formulae employed by Bloomfieldian linguists, they are interactive, e.g., rules two and three interact with rule one to further expand the

---

\(^{23}\) Chomsky, *Aspects*, pp. 64-65.

\(^{24}\) Chomsky, *Aspects*, pp. 64-65. Notice that a derivational tree and a labeled bracketing convey less information than a sequential derivation because they do not always reveal the specific order in which the rules have been applied. While several different sequential derivations may, for example, be used to generate *The man hit the ball*, only one labeled bracketing and only one derivational tree fit the utterance. The foregoing fact prompts Chomsky to observe that a derivational tree "retains just what is essential . . . for the determination of the immediate constituent analysis" of a sentence (*Syntactic Structures*, pp. 27-28). Much the same could be said about the use of labeled bracketings.
The man hit the ball.

FIGURE # 1
symbols that appear to the right-hand side of the arrow in that rule. It is the interactive character of such rules that makes a phrase structure grammar "generative" in the constructive sense of the term; because they can interact, thereby expanding an initial symbol into a string of symbols ready to replaced by formatives, such rules enable a linguist to generate--to construct--a given sentence. Furthermore, because they can interact in different ways, a limited number of phrase structure rules can be used to generate a variety of different sentence-level patterns. The foregoing three rules can, for example, be used to generate not only the pattern \([\text{Det} + \text{N}] + [\text{V} + [\text{Det} + \text{N}]]\), but also the following five sentential patterns: 2. \(\text{N} + [\text{V} + \text{N}]\), 3. \(\text{N} + [\text{V} + [\text{Det} + \text{N}]]\), 4. \([\text{Det} + \text{N}] + [\text{V} + \text{N}]\), 5. \(\text{N} + \text{V}\), and 6. \([\text{Det} + \text{N}] + \text{V}\). Thus, a phrase structure grammar can specify six different sentential formulae by listing only three short rules, whereas a Bloomfieldian grammar would have to actually list all six formulae. Moreover, the

---

25 The statement that a language consists of rules that can interact to generate sentences might be taken as an indication that languages are dynamic and therefore constitute systems in the sense explicated in Chapter Two (see above, pp. 51-52). However, that statement is really an abbreviated way of saying that people can use such rules to generate sentences; it is language-users, not language itself, that possess the dynamic properties in question. In keeping with common practice, the present study will, however, continue to speak of "language-systems"; see fn # 76, p. 126.
addition of each new rule to a phrase structure grammar usually results in the specification of more than one new sentential pattern; the number of rules and the number of formulae do not increase at a one-to-one ratio.

Because phrase structure rules can be used to specify a wide variety of different sentence-level patterns without having to actually list each of those patterns in the manner of a Bloomfieldian grammar, they represent a significant advance in the traditional taxonomic effort to specify the permissible syntagmatic combinations of paradigmatic classes of units. Givón has suggested that formal modeling of linguistic phenomena allows

a restatement of the facts at a tighter level of generalization. Such a restatement has obvious, common sensical advantages over a mere list of the facts:

1. Maximum clarity
2. Maximum economy
3. Maximum generality

Because phrase structure rules enable linguists to model linguistic phenomena at a much tighter level of generalization than is possible through the use of sentential formulae, generative grammars possess the foregoing advantages to

26 Givón, p. 5.
a greater degree than do Bloomfieldian grammars.

The degree to which generative grammars exhibit the foregoing virtues is, moreover, considerably magnified by their incorporation of an additional type of rule: the recursive rule (in the case of Chomskian grammar, the recursive phrase structure rule). A recursive phrase structure rule is a rule that can be applied to its own output an indefinite number of times in the derivation of a single sentence. Such rules enable generative linguists to model certain sorts of linguistic phenomena in a very economical way.

Among the various syntactic constructions found in English, as well as other languages, there are, for example, those which consist of two or more main clauses joined together by coordinating conjunctions like and. Interestingly enough, there appears to be no upper limit to the number of coordinated main clauses that may be contained in a single sentence.

27 Although Chomsky preferred to handle recursion by means of transformational rather than phrase structure rules at the time he wrote Syntactic Structures, in Aspects and all of his later work only phrase structure rules exhibit recursive properties (Aspects, p. 137; see also, Newmeyer, p. 26).

28 Lyons, Semantica, p. 389. Recursive rules were first developed by mathematicians and logicians. Chomsky and other early generative linguists then adapted them to the task of describing natural languages. For a discussion of the subject, see Newmeyer, p. 36.
grammatical sentence. It is, of course, likely that a speaker would find a sentence composed of a large number of coordinated main clauses unusual—the sort of thing that he would not normally say or expect to hear during the course of ordinary discourse—but he would most likely admit that it is nevertheless perfectly grammatical, i.e., syntactically well-formed. Thus, the English language does not contain a longest possible grammatical sentence; no matter how many main clauses have been strung together with coordinating conjunctions, it is always possible to add one more to the string.

In order to account for the potentially unlimited length of such constructions, it is necessary to employ a recursive rule—such as the following phrase structure rule: $S \rightarrow S + \text{and} + S$, where $S$ is allowed to represent both a sentence and a main clause. That rule allows the initial symbol $S$ in a derivation to be rewritten as $S + \text{and} + S$ and allows either or both of the $S$-symbols in that string to be similarly rewritten, and so on indefinitely, before any of the $S$-symbols are expanded into other symbols. Because the three phrase structure rules introduced earlier would allow each of the $S$-symbols in such a string to be expanded into

---

29 Chomsky, Mind, p. 118.
30 Baker, p. 277.
any one of six different patterns, the combinatorial possibilities inherent in a phrase structure grammar that contained only the four rules introduced so far would be enormous. Indeed, the addition of a single recursive rule enables a phrase structure grammar to specify an infinite number of different sentence-level patterns on the basis of a finite input.\textsuperscript{31}

The incorporation of recursive rules marks a crucial difference between generative grammars and their Bloomfieldian predecessors. Whereas a Bloomfieldian grammar could specify, albeit in cumbersome fashion, all of the sentential patterns that can be specified by a set of non-recursive phrase structure rules, it does not possess any means of specifying the set of sentential patterns captured by a grammar that contains recursive rules—it being impossible to actually list each and every member of an infinite set.

Thus, the rules contained in a generative grammar and the lists of sentential formulae contained in a Bloomfieldian grammar do not constitute mere notational variants of each other after all; the contrast between a "grammar of lists" and a "grammar of rules" marks a significant

\textsuperscript{31} Thus, it is the inclusion of recursive rules that accounts for the fact that the phrase structure component of a Chomskian grammar is "generative" in both the first and third senses of the term distinguished in fn # 16, pp. 242-43.
distinction. That distinction is, moreover, closely linked to several other important differences between the two approaches.

Before those differences are discussed, however, one feature of the generative approach warrants special emphasis. As the foregoing discussion has shown, phrase structure rules constitute a concise means of specifying the integrative relations between units located on adjacent levels within the linguistic hierarchy, e.g., the recursive phrase structure rule discussed above provides a concise way of specifying the relation between a conjunctive sentence and the clauses that it contains, the first rule discussed provides a way of specifying the relation between a main clause and the phrases it contains, and the second rule provides a way of specifying the relation between a noun phrase and the words it contains. Thus, identification of the hierarchical levels of linguistic structure is as essential to the generative approach as it is to the taxonomic approach; it would have been impossible to formulate a generative approach to language without having first identified a series of hierarchical levels of linguistic structure.

The difference in the formal devices that Chomsky and the Bloomfieldians utilize as a means of specifying the immediate constituent structure of sentences is closely linked to differences of opinion concerning three additional,
interrelated issues of fundamental importance: the extent of the linguistic capacities possessed by native speakers of a language, the nature of the underlying source of those capacities, and the proper function of grammatical theory.

According to Chomsky, observation of normal linguistic behavior indicates that an average native speaker of a language possesses certain creative linguistic capacities:

Having mastered a language, one is able to understand an indefinite number of expressions that are new to one's experience . . . and one is able, with greater or lesser facility, to produce such expressions on an appropriate occasion, despite their novelty . . . . The normal use of language is, in this sense, a creative activity. 32

A native speaker is also, according to Chomsky, capable of deciding whether each of an indefinite number of previously unfamiliar utterances is syntactically well- or ill-formed, i.e., grammatical or ungrammatical.33 (Although it may

32 Chomsky, Mind, p. 100.

33 See, for example, Chomsky, Syntactic Structures, pp. 13-15, Aspects, pp. 10-11, and Baker, p 4. It is important to note that the above statement allows for the fact that a speaker will not always be able to decide whether a given utterance is grammatical or not; it merely asserts that there is no upper limit to the number of utterances that he is capable of classifying as grammatical or ungrammatical. As Chomsky points out, there are degrees of grammaticalness (Aspects, pp. 11, 75-79, 148-53); and that allows for a continuum of "borderline" utterances which fall between the potentially infinite sets of definitely grammatical and definitely ungrammatical sentences.
well be necessary to posit additional types of well- and ill-formedness, all subsequent references to evaluative capacities and well-formedness in what follows will, unless otherwise specified, concern matters of syntax.\textsuperscript{34}

34 During a discussion of grammaticality in \textit{Aspects}, Chomsky draws a distinction between syntactically ill-formed utterances and utterances that exhibit "purely semantic (or 'pragmatic') incongruity" (pp. 76-77). Although Chomsky does not explicitly claim that the latter utterances exhibit a distinct type or level of ill-formedness, his remarks seem to suggest that such is the case. Other linguists have, in fact, gone on to posit a variety of different types or levels of well- and ill-formedness. In \textit{Semantics}, for instance, Lyons posits at least three separate types or levels: phonological, syntactic, and semantic (p. 379). The differences between the three levels may be illustrated in terms of the following three strings: (a) \texttt{bnrgout cblton}, (b) \texttt{the man the hit ball}, (c) \texttt{quadruplicity drinks procrastination}. The first string is obviously not even phonologically well-formed; English does not permit the combinations of consonants that appear at the beginning of the two "words" in that string. The second string is phonologically well-formed, but it is syntactically ill-formed for reasons having to do with the sequential arrangement of the words—English does not permit a verb to be immediately preceded by a determiner and immediately followed by a noun. Although the third string is phonologically and syntactically well-formed, it is not semantically well-formed; aside from the possibility of some sort of metaphorical interpretation, (c) cannot be assigned any coherent meaning. In addition, it might be useful to posit a level of pragmatic ill-formedness to cover a variety of other sorts of anomalies, e.g., anomalies that derive from the relation between an utterance and the context in which it is uttered.

Two additional points need to be made here. First, it may not always be easy to tell whether a particular utterance is ill-formed in one way rather than another. For example, while Chomsky claimed that the sentence \texttt{Colorless green ideas sleep furiously} is anomalous for purely semantic rather than syntactic reasons in \textit{Syntactic Structures} (p. 15), he later seems to suggest that it is also syntactically anomalous (\textit{Aspects}, p. 149). Indeed, the question of where
Unlike his Bloomfieldian predecessors, Chomsky believes, moreover, that a native speaker is able to produce, understand, and evaluate utterances based on unfamiliar syntactic patterns:

much of what we say in the course of normal language use is . . . not even similar in pattern—in any useful sense of the terms "similar" and "pattern"—to sentences of discourse that we have heard in the past. This is a truism, but an important one, often overlooked and not infrequently denied in the behaviorist [i.e., Bloomfieldian] period of linguistics . . . when it was almost universally claimed that a person's knowledge of language is representable as a stored set of patterns, overlearned through constant repetition and detailed training, with innovation being at most a matter of "analogy." The fact surely is, however, that the number of
to draw the boundaries between syntax, semantics, and pragmatics has led to some of the most hotly debated issues in current linguistics (see, for example, Janet Fodor, pp. 97-102 and Lycan, pp. 4-8).

Second, although the foregoing remarks might seem to suggest that the various levels of well-formedness exhibit a uniform sequential order wherein well-formedness on lower-levels constitutes a necessary condition for well-formedness on higher-levels, that does not always appear to be the case. While minimal well-formedness would seem to constitute a necessary condition for syntactic well-formedness, and syntactic well-formedness may be a necessary condition for semantic well-formedness, it would appear that an utterance could be pragmatically well-formed without being semantically or even syntactically well-formed. For example, a sentence fragment may be considered pragmatically well-formed, i.e., appropriate to context, if the context of utterance is a crisis.
sentences in one's native language that one will immediately understand with no feeling of difficulty or strangeness is astronomical; and that the number of patterns underlying our normal use of language and corresponding to meaningful and easily comprehensible sentences in our language is orders of magnitude greater than the number of seconds in a lifetime.\(^{35}\)

Thus, whereas most Bloomfieldian linguists supposed that a native speaker is capable of dealing with utterances that are innovative only in the sense that different words have been inserted into underlying syntactic patterns that are the same as—or, at most, analogous to—patterns he has previously encountered, Chomsky believes that a normal native speaker is capable of dealing with utterances based on patterns that are in no clear sense similar or analogous to those he has previously encountered.

Once the Bloomfieldian view concerning the extent of a native speaker's linguistic capacities has been rejected, their views concerning the source of those capacities must, of course, also be rejected; the creative aspects of normal language use simply cannot be accounted for in terms of a fixed repertoire of sentence-level patterns. Chomsky therefore suggests an entirely different hypothesis about the underlying source of a native speaker's linguistic capacities:

\(^{35}\) Chomsky, *Mind*, p. 12.
A person who knows a language has mastered a system of rules that assigns sound and meaning in a definite way to an infinite class of possible sentences. . . . we may say that [this system] generates an infinite set of "structural descriptions," each structural description being an abstract object of some sort that determines a particular sound, a particular meaning, and whatever [syntactic representations are needed] to mediate the relation of sound and meaning. . . . we use [our] knowledge [of these rules] totally without awareness or even the possibility of awareness, in producing [various] sentences or understanding them when they are produced by others.36

Thus, Chomsky maintains that a native speaker's ability to produce, comprehend, and evaluate the syntactic well-formedness of novel utterances derives from an ability to assign each of a potentially infinite number of sentences a structural description that specifies its phonological, semantic, and syntactic form. That ability derives, in turn, from the speaker's implicit knowledge of a particular system of rules—an internalized grammar of his language.37

36 Chomsky, Mind, pp. 103-04.

37 While Chomsky believes that implicit knowledge of an internalized grammar capable of generating a structural description for a potentially infinite set of sentences constitutes a necessary condition for normal language use, he does not believe that it constitutes a sufficient condition: "Surely the normal use of language requires access to other systems of knowledge and belief. . . . We use language against a background of shared beliefs about things within the framework of a system of social institutions. The study of language use must be concerned with the place of language
Chomsky's views concerning the source of a native speaker's linguistic capacities are, moreover, closely related to his views about the proper goals of linguistic study for he maintains that the primary task in studying a particular language is that of formulating "an explicit articulated theory that attempts to express precisely the rules and principles of the grammar in the mind of [an] ideal speaker-hearer." The linguist's grammar is thus to constitute a formal model of an ideal native speaker's internalized grammar, a model that will enable the linguist to mirror the effects of the speaker's internalized grammar (though it need not necessarily enable him to produce those effects in precisely the same way). For example, the linguist's grammar ought to enable him to simulate the native speaker's ability to generate a structural description for each of a potentially infinite number of sentences: "A fully adequate grammar must assign to each of an infinite

in a system of cognitive structures embodying pragmatic competence, as well as structures that relate to matters of fact and belief" (Rules, pp. 247-48). Thus, Chomsky maintains that the linguistic abilities of a native speaker derive from the interaction between a variety of different systems of knowledge and belief.

38 Chomsky, Rules, p. 220.

39 See above, fn # 16, pp. 242-43.
range of sentences a structural description indicating how this sentence is understood by the ideal speaker-hearer."^40

If it is to accomplish that task, a grammar must provide the linguist with some means of generating an infinite output from a finite amount of input; that crucial function is performed by the phrase structure component in a Chomskian grammar. Other components are, of course, necessary as well, but the phrase structure component is crucial because it is responsible for initially generating a potentially infinite set of phrase markers, thereby supplying the input, either directly or indirectly, to all the other components.41

As earlier remarks have indicated, Chomsky also believes that native speakers are capable of assessing the syntactic well-formedness of a potentially infinite set of sentences.42 How is the linguist's grammar to account for that capacity? In Syntactic Structures, Chomsky required that an adequate grammar for a given language L generate all of the grammatical sequences of L and none of the ungrammatical ones. One way to test the adequacy of a grammar proposed for L is to determine whether or not the sequences

[^40]: Chomsky, Aspects, pp. 4-5.
[^41]: Chomsky, Mind, p. 125; see also Janet Fodor, p. 113.
[^42]: See above, p. 255.
that it generates are actually grammatical
[in the opinion of] native speaker[s]. 43

Given such a grammar, a linguist could also simulate an
ideal native speaker's evaluative capacities: in order to
determine whether a given sequence is syntactically well-
formed or not, the linguist could simply check to see wheth­
er the rules of the grammar enable him to generate that
string. 44 Since the phrase structure component is needed in
order to initially generate a potentially infinite set of
strings, it would naturally play a crucial role in such a
simulation, just as it does in the simulation of the speak­
er's ability to assign structural descriptions to sentences.

43 Chomsky, Syntactic Structures, p. 13.

44 In Aspects Chomsky requires that an adequate gram­
mar generate not only the well-formed sentences of a lan­
guage and their structural descriptions, but also the
ill-formed or deviant sentences and their structural des­
criptions: "A descriptively adequate grammar must assign to
each string a structural description that indicates the man­
er of its deviation from strict well-formedness (if any).
A natural terminological decision would be to say that the
grammar directly generates the language consisting of just
the sentences that do not deviate at all . . . with their
structural descriptions. The grammar derivately generates
all other strings . . . with their structural descriptions" (p. 227). Because such a grammar would convey information
about the manner in which each ill-formed string deviates
from well-formedness, it would enable a linguist to simulate
the evaluative capacities of native speakers even more
closely. Of course, such a grammar would also have to be
much more complicated.
Whereas sentential formulae could be used to represent the stock repertoire of sentence-level patterns posited by the Bloomfieldians as the underlying source of a native speaker's linguistic capacities, more powerful formal devices are thus needed to construct a model of the internalized grammar posited by Chomsky. Moreover, even though the Bloomfieldians speculated about the source of a native speaker's linguistic capacities, they considered the proper task of linguistics to be the description of sentences—not the description of the underlying source which enables a native speaker to produce, comprehend, and evaluate sentences. In his *Methods in Structural Linguistics*, for example, Harris stated that "The overall purpose of work in descriptive linguistics is to obtain a compact one-to-one representation of the stock of utterances in [a given] corpus."\(^{45}\)

While Chomsky's goal of constructing a model of an ideal native speaker's internalized grammar requires the use of interactive rules, the description of the finite set of utterances contained in a given corpus could be achieved, albeit in a rather cumbersome fashion, through use of sentential formulae. Thus, the differences in the formal

---

devices utilized by Chomsky and the Bloomfieldians are symptomatic of larger differences of opinion concerning a whole set of crucially important issues: the extent of a native speaker's linguistic capacities, the underlying source of those capacities, and the primary goals of linguistic theory.

Since the goals of the Bloomfieldians were far less ambitious than are Chomsky's goals, they would, of course, be more easily attainable. As Lyons observes, "The generation of all and only the sentences of English, or any other language, might seem to be impossibly ambitious," but, as he goes on to point out:

this represents an ideal, which, even if it is impossible of fulfillment, is a goal toward which the grammarian of any language [can] continually work; and one grammar can be evaluated as better than another if, all other things being equal, it approximates more closely to this ideal.46

Thus, the goals of Chomskian grammar represent an ideal that can be useful in evaluating the relative merits of rival accounts—even if none of the accounts succeed in fully attaining those goals.

46 Lyons, Chomsky, p. 46.
LITERARY APPLICATION

The primary literary import of the foregoing remarks obviously concerns the possibility of developing an approach to literature that would roughly parallel Chomsky's approach to language, a research program that has attracted relatively few advocates and many opponents. As Pettit's brief survey of pro-generative work indicates, such work has been rather scarce, and the little work that has been done has for the most part consisted, not of attempts actually to develop a generative theory of literature, but of attempts to argue in favor of trying to develop such a theory. Moreover, such discussions have rarely included a detailed account of what is involved in the generative approach to linguistic study, and, as a result, they have tended to be rather vague about the nature and extent of the presumed parallel between the linguistic model and the approach to literature for which the author has argued.

In order to achieve greater clarity about such matters, the following discussion begins with a brief summary of the basic claims that underlie Chomsky's generative approach to language, based on the remarks contained in the previous

---

47 Pettit, pp. 45-6.
section. That summary is followed by a preliminary consideration of the literary applicability of those claims, and a group of roughly parallel literary claims is formulated. The remaining discussion is devoted to an extended consideration of the plausibility of the latter group of claims in light of the criticisms that have been advanced by anti-generative writers.

Chomsky's generative approach to language appears to be based on four main claims. First, Chomsky claims that normal native speakers of a natural language are able to produce, understand, and evaluate the syntactic well-formedness of an indefinitely large number of sentences based on an indefinitely large number of different syntactic patterns. Second, he claims that a speaker could not produce, understand, or evaluate the syntactic well-formedness of a sentence without assigning a structural description to it. Third, he claims that the primary goal in the grammatical study of a particular language ought to be that of formulating an account of the linguistic knowledge that underlies a native speaker's unbounded ability to assign structural descriptions to sentences. Fourth, he claims that an adequate account of that knowledge must enable a linguist to simulate the native speaker's unbounded ability to generate a structural description for each of a potentially infinite number of sentences based on an indefinitely large number of
different syntactic patterns and must, as a consequence, contain, among other things, a set of interactive formation rules, at least some of which are recursive.

Since more than one of the above claims contains a reference to syntactic phenomena, perhaps the literary applicability of the distinction between syntax, semantics, and pragmatics should be discussed at the outset. Unfortunately, very few previous pro-generative writers have even addressed this matter. The omission is perhaps understandable for, as Lyons observes, even within linguistics itself "The boundary between syntax and semantics has long been, and remains, the subject of dispute"; and much the same can be said about the boundary between semantics and pragmatics.48

Since there is a good deal of disagreement among linguists concerning the exact boundaries between the domains of syntax, semantics, and pragmatics, any attempt to identify their literary analogues is, of course, bound to be controversial. However, because the distinction between syntactic and non-syntactic properties of sentences underlies much of Chomsky's theory of language—e.g., the tripartite division of a grammar into a phonological, a syntactic, and a semantic component, which are respectively responsible for

48 Lyons, Semantics, p. 374; see also above, fn # 34, pp. 256-57.
determining the phonological, syntactic, and semantic representations that together constitute a complete structural description of a sentence—some attempt to identify those literary analogues must be made. Without such an attempt, the nature and extent of the parallels between a pro-generative critic's proposals and the linguistic theory he is using as a model will remain hopelessly vague. Indeed, the failure to make such an attempt constitutes one of the principal sources of vagueness in previous proposals.\footnote{In \textit{Structuralist Poetics}, for example, Culler argues in favor of a generative approach based on the notion that proficient readers possess implicit knowledge of literary conventions that can be used to generate structural descriptions for literary works (pp. 116-20). Unfortunately, Culler does not, however, provide any indication of what might be involved in assigning a structural description to a literary work; nor does he include an account of what is involved in assigning a structural description to a sentence. As a result, it is unclear whether a structural description of a literary work is to specify its semantic, perceptual, and syntactic form as a structural description of a sentence does, or whether it is to specify only one of those features. Moreover, since he does not discuss the distinction between syntax, semantics, and pragmatics—as it applies in either the domain of literature or that of language—it is unclear what the preceding statements would even mean. Without discussions of such matters, it is difficult to tell what his proposals amount to and whether they are plausible or not.} 

Despite disagreements concerning the exact boundaries between semantics and pragmatics, previous linguists and philosophers of language have tended to agree that the former should be primarily concerned with what might be called the "decontextualized" meaning of linguistic units—the
meaning that such units possess when considered apart from any particular context of utterance—and they have tended to agree that the latter should be concerned with the "contextualized" meaning of such units. According to Lycan, for example:

Semantics is the study of the meanings of words and sentences, in abstraction from the various speech contexts in which they might be called upon to serve. Pragmatics is the study of the uses of sentences in context, and of how gaps in sentences' abstract meanings may be filled in by context.  

A number of previous writers have, moreover, gone on to suggest that semantics will be particularly concerned with the referents and truth conditions that linguistic units have in abstraction from particular contexts of utterance and that pragmatics will be concerned with the role that various contextual factors play in the performance of different types of speech acts and in the use of sentences to convey various sorts of loose implications, secondary meanings, etc.  

50 Lycan, p. 4. As Lyons observes, it is not possible to screen out all aspects of context when talking about the meaning of a sentence; he therefore suggests that semantics be construed as the study of "the meaning of maximally de-contextualized sentences" (Lyons, Semantics, p. 590). Subsequent references to the meaning, referents, and truth conditions of "decontextualized" units are intended to carry the qualification suggested by Lyons.

51 See, for example, Lycan, pp. 18–25, 73–82; Devitt, pp. 68–74, and Lyons, Semantics, pp. 590–606.
Suppose, for example, that during the course of a movie Smith turns around and utters the following sentence to a man sitting behind him: *Every civilized person knows that it is rude to talk in the middle of a movie.* Although Smith's utterance constitutes a declarative sentence, it is clear, in view of the situation, that Smith is not simply making an assertion—he is issuing a request or command; he is not simply trying to convey an interesting piece of information to the man sitting behind him—he is asking the man to be quiet. His utterance thus implies more than is contained in its literal meaning. Such facts belong to the domain of pragmatics. A semantic characterization of the sentence uttered by Smith would not take its situational context into account; it would be more concerned with the conventional meaning of each word and with the truth conditions of the sentence apart from any particular context of utterance.

It appears that degree of contextual dependency might even provide a rough basis for distinguishing cases of semantic ill-formedness from cases of pragmatic anomaly. If the unacceptability of an utterance derives from the fact that it was uttered in an inappropriate context or from the fact that it seems difficult to interpret in light of background knowledge about the way the world is, relevant social conventions, etc., then the problem is most likely pragmatic. Lyons provides a good example:
Consider . . . an utterance-token like *We are having a fine summer here in Queensland this year* produced by someone in Edinburgh in December. It is grammatically and semantically well-formed, but situationally inappropriate; and it is for this reason that it is uninterpretable (except . . . under rather special circumstances).\(^{52}\)

By way of contrast, the difficulties in interpreting an utterance of *Quadruplicity drinks procrastination* would not derive from its relation to a given context of utterance, or from background knowledge of the world, social conventions, etc.; it is simply difficult to imagine what that sentence could mean, regardless of its utterance-context. As a result, it may be said to be semantically ill-formed.

Characterizations of syntax typically focus on the constraints that govern the ways in which words can be combined. Lyons, for example, characterizes syntax as the study of "the distribution of word-forms throughout the sentences of [a] language in terms of the permissible combinations of classes of word-forms."\(^{53}\) Unfortunately, such a characterization cannot be used clearly to distinguish syntax from semantics for, as Lyons himself points out elsewhere, "there is presumably a combinatorial aspect to

---

\(^{52}\) Lyons, *Semantics*, p. 578.

\(^{53}\) Lyons, *Semantics*, p. 375.
The problem can be illustrated by comparing the following two pairs of sentences: 1A. The tall man hit the ball, 1B. The tall man the hit ball; 2A. John signed his will and died, and 2B. John died and signed his will. The latter sentence in each pair can be produced by simply reordering the words contained in the former sentence, and the latter sentence in each pair is anomalous. Nevertheless, only 1B is syntactically ill-formed; the problematic status of 2B derives from semantic or pragmatic factors. There is, however, another important difference between the two cases. In order to explain why 2B is anomalous, a person would have to talk about matters of reference and truth; the anomaly derives from the fact that the activities referred to cannot be performed in that order—people cannot perform voluntary activities after they are dead. It would not, however, be necessary to talk about reference and truth in order to explain the problematic status of 1B; it is anomalous because a verb cannot be immediately preceded by a determiner and immediately followed by a noun.

Perhaps, then, syntax can be distinguished from semantics and pragmatics on the basis of whether the combinatorial constraints in question are bound up with specific semantic, as well as to syntax.  

54 Lyons, Semantics, p. 117.
information about reference, truth conditions, situational context, etc. This characterization does not, however, require that there be no connection whatsoever between syntactic phenomena and questions of reference, meaning, etc. If Lyons is correct in thinking that "there is an intrinsic connection between the meaning of words and their distribution," there presumably will be some sort of connection. 55 The characterization only requires that syntactic phenomena not be tied to specific—as opposed to general—information about the referents and meanings of words. Although such a characterization of the distinction is still exceedingly rough, it should provide some indication of how the distinction might be applied within the domain of literature.

If what has been said so far is on the right track, the distinction between syntax, semantics, and pragmatics might be applied within the domain of literature roughly as follows: pragmatics would be primarily concerned with the relation between a literary work and a particular situational context, either the context that obtained at the time when it was written or the context that obtained at the time of a particular reading or, in the case of drama, performance. A critic's statement that a particular work is no longer

55 Lyons, Semantics, p. 375.
relevant, that it is outdated, and so on might, accordingly, be considered to be a statement about the work's pragmatic well-formedness given the general contexts of reading that obtain at the time of the critic's utterance.

Literary semantics would be primarily concerned with matters of reference and truth. Statements about the referents of the percept-images contained in a play and statements about the truth conditions of various lines of dialogue would, for example, fall within the bounds of semantics. Cases of semantic ill-formedness might be said to arise when a work violates its own conventions of truth and reference. The attribution of physiologically incompatible attributes to one and the same character within a realistic work might, for example, constitute a semantic anomaly (though it might not constitute an anomaly within the context of an absurdist or dadaist work).

Finally, literary syntax would be concerned with the "parsing" of literary works of a given type or kind into units or segments whose combinatorial distributions can be accounted for without introducing specific information about reference, truth, appropriateness to situational context, etc. Dramatic works are, for example, customarily parsed into units like initial incident, complicating incident, crisis, climax, denouement, etc., which are characterized without introducing specific information about reference,
truth, or appropriateness to situational context, and it is possible to specify constraints on the way that such units can be combined. In fact, some of the units are even defined in terms of the sequential relations they bear to other units, e.g., the denouement is defined as that portion of a play which follows the climax or highest point of emotional tension. Violations of the combinatorial constraints associated with the genre or style of a particular work would count as cases of syntactic ill-formedness. The constraints associated with most dramatic genres and styles do not, for example, permit the climax or highest point of emotional tension to occur earlier than halfway through the play. A play whose climax occurred earlier than that would be found unacceptable on the grounds that it peaked too early; it would, in other words, be viewed as syntactically ill-formed.

Now that the literary applicability of the distinction between syntax, semantics, and pragmatics has been discussed, the possibility of formulating literary counterparts for each of the four claims that underlie Chomsky's generative approach to language may be considered. An exact literary

---

56 In order to capture the potentially infinite diversity of underlying patterns, the parsing process would, of course, have to be carried far beyond identification of units of the foregoing sort; those units are being mentioned for purely illustrative purposes and are by no means intended to constitute a complete list of the types of syntactic units contained in dramatic works.
counterpart of the first of those claims would assert that a person who possesses competence with respect to a particular type or kind of literary works is able to produce, comprehend, and evaluate the syntactic well-formedness of an indefinitely large number of such works based on an indefinitely large number of different syntactic patterns.

It is clear at once that one feature of the claim should be dropped. Even individuals who possess a high degree of competency with respect to a particular type or kind of literary work are often unable to produce works of that kind. In the realm of literature, unlike the realm of language, the ability to produce and the ability to comprehend do not normally go hand in hand. A generative theory of literature should, therefore, focus primarily on the activities involved in reading rather than writing.

A similarly modified literary counterpart of the second of Chomsky's claims would assert that a reader could not understand or evaluate the syntactic well-formedness of a literary work without assigning it a structural description. A structural description of a literary work would presumably specify its perceptual, semantic, and syntactic form just as a structural description of a sentence does.

An exact literary counterpart of Chomsky's third claim would assert that the primary goal in studying works of a particular kind or type ought to be that of formulating an
account of the literary knowledge that underlies a reader's unbounded ability to assign structural descriptions to such works. An exact literary counterpart of Chomsky's fourth claim would assert that an adequate account of that literary knowledge must enable a theorist to simulate a proficient reader's ability to generate a structural description for each of a potentially infinite number of such works based on an indefinitely large number of different syntactic patterns and must, as a consequence, contain, among other things, a set of interactive formation rules, at least some of which are recursive.

One feature of the fourth claim requires modification. It seems likely that a reader's implicit knowledge of literary conventions would not alone enable him to generate the perceptual and semantic representations included in a complete structural description of a literary work. As a result, a theorist's account of that knowledge should be expected to enable him to generate, at most, a syntactic description for each of an indefinitely large number of works based on an indefinitely large number of underlying patterns. As earlier remarks have indicated, Chomsky believes that a native speaker's linguistic abilities ultimately depend on the interaction of a number of different systems of knowledge and belief, and a reader's literary abilities may well depend on even closer and more complex
interactions between diverse systems of knowledge and belief.  

The remainder of the present chapter will be devoted to the task of assessing the plausibility of each of the above literary claims in light of the various arguments that have been advanced by anti-generative writers. In order to facilitate that task, the claims will now be restated in slightly modified form and will be numbered as follows:

1A. a person who possesses competence with respect to a particular type or kind of literary works is capable of understanding an indefinitely large number of such works based on an indefinitely large number of different syntactic patterns; 1B. a person who possesses competence with respect to a particular type or kind of literary works is capable of evaluating the syntactic well-formedness of an indefinitely large number of such works based on an indefinitely large number of different syntactic patterns; 2. a reader could not understand or evaluate the syntactic well-formedness of a literary work without assigning a structural description to it; 3. the primary goal in studying works of a particular kind or type ought to be that of formulating an account of the literary knowledge that partly underlies a reader's ability to assign structural descriptions to works of that

57 See above, fn # 37, pp. 259-60.
kind; and 4. an adequate account of that knowledge should at least enable a theorist to simulate a proficient reader's ability to generate a syntactic description for each of a potentially infinite number of such works based on an indefinitely large number of different syntactic patterns and must, as a consequence, contain, among other things, a set of interactive formation rules, at least some of which are recursive.

As it stands, generative assertion 1A is somewhat ambiguous and is, as a result, open to misinterpretation. The major source of potential error derives from the ambiguous character of the phrase "capable of understanding." It might be thought that the capacity in question consists merely in an ability to make some sort of sense—whatever it might be—of a literary work, but that is not the case. When a linguist says that native speakers are "capable of understanding" sentences, he does not simply mean that each speaker is capable of making his own sort of sense of the sentences, he means that they are capable of making roughly the same sort of sense of each sentence—that there is widespread intersubjective agreement in the way that they understand each sentence. Generative assertion 1A is intended to make the rather controversial claim that proficient literary critics and theorists possess a roughly similar capacity.
A second potential misinterpretation of generative assertion 1A could derive from a failure to appreciate the fact that the assertion makes a claim only about a proficient individual's capacities—not about his actual behavior. In particular, the assertion does not claim that such individuals are actually able to understand each and every work of a given kind they encounter—any more than Chomsky's corresponding linguistic assertion claims that native speakers of a particular language are able to understand each and every sentence of that language they encounter. In both the realm of literature and that of language even a proficient individual's knowledge is likely to be incomplete, and gaps in a person's knowledge will obviously render him incapable of understanding certain works or sentences. Moreover, in both realms performance factors like an individual's emotional state of mind, interruptions, etc. may prevent him from manifesting the knowledge he does possess.

What generative assertion 1A does imply is that there is no upper limit to the number of literary works of a particular kind that a proficient individual is capable of understanding. Thus, he is said to be capable of understanding an indefinitely large number of works based on an indefinitely large number of different syntactic patterns. An attempt to refute assertion 1A would, therefore, have to be based on one of the three following claims:
(a) although proficient individuals are obviously capable of making at least some sort of sense out of literary works, they are not capable of understanding any literary works in the strong sense implied by 1A, (b) although proficient individuals are, at least sometimes, capable of understanding a literary text in roughly the same way, each individual is capable of such understanding only with respect to a finite number of works, or (c) although a proficient individual is capable of understanding an indefinite number of literary texts in the strong sense, his capacity to do so is limited to works that are based on a finite set of underlying syntactic patterns.

Most anti-generative writers would probably choose the first strategy. For example, a deconstructionist critic like Miller, who believes that every literary work can be interpreted in an unlimited number of ways, might wish to advance something like the following argument.58

Generative assertion 1A claims that proficient literary critics and theorists possess a capacity to understand literary works that is roughly similar to the capacity for understanding sentences that linguists claim native speakers have, but that is not true. When a linguist talks about understanding a set of sentences, he implicitly assumes the

58 See, for example, Miller, "Steven's Rock."
existence of a communally shared interpretation for each member of that set. He assumes, in other words, that a majority of the interpretations assigned to each given sentence by native speakers of the language in question would be similar enough to allow abstraction of a "standard" consensus interpretation of it. Since he believes that to be the case, a linguist can posit something like the following operational definition of linguistic understanding: an individual can be said to understand a particular sentence if and only if his behavior, e.g., his use of the sentence, his statements about it, etc., indicates that he assigns it an interpretation sufficiently similar to its "standard" interpretation.

However, the critic might continue, any attempt to define literary understanding in terms of some sort of "standard" interpretations would be ludicrous. Even a cursory examination of critical commentaries would show that interpretations of the same work rarely—if ever—converge at all. Since literary critics virtually never advance similar interpretations of the same text, it is obvious that there are no normative standards that could be used to determine whether or not an individual "really" understands a given work. Hence, generative assertion 1A must be false. Critics can be said to be capable of understanding literary works in the weak sense of being able to assign them some interpretation or other, but they cannot be said to be
capable of understanding works in anything like the sense that the word assumes in the hands of a linguist because there are no communally shared interpretations that could be used to measure and define such understanding.

The foregoing argument certainly possesses an initial air of plausibility; critics do seem to be forever quarreling over matters of interpretation. Morris Weitz's *Hamlet and the Philosophy of Literary Criticism*, for example, furnishes ample empirical evidence of that.⁵⁹ Weitz surveyed interpretations of *Hamlet* advanced by over seventeen of the most distinguished critics who have written on the subject, and his survey leaves no doubt that their interpretations do not converge on a single "standard" one.

Does such evidence prove that proficient literary critics and theorists do not have a capacity roughly similar to a native speaker's capacity for linguistic understanding? It might if it were true that every sentence of a natural language can be assigned a single, standard interpretation, but that is, of course, not the case—for natural languages abound with ambiguous sentences, sentences whose most distinguishing feature lies in the fact that the interpretations assigned to them by native speakers do not converge on

---

a single standard interpretation. That fact does not, however, mean that native-speaker interpretations of an ambiguous sentence hardly converge at all, or that such a sentence can be assigned any interpretation whatsoever. Interpretations of an ambiguous sentence do converge; it is just that they converge on a set of interpretations rather than a single one. Moreover, that set defines a standard just as much as a single interpretation does in the case of an unambiguous sentence; an individual can assign an ambiguous sentence any one of a number of different interpretations and still be counted as understanding it—as long as the interpretation he selects bears a sufficient resemblance to one of the interpretations contained in the standardized set.

Earlier remarks conceded that anti-generative writers are justified in claiming that literary critics seldom assign a given literary work interpretations that converge on a single consensus interpretation, but are they equally justified in suggesting that interpretations of the same work hardly converge at all? Can a work really be interpreted in an unlimited number of ways? Is the interpretive process really guided by no standards of acceptability whatsoever?

The answers would appear to be uniformly negative. First of all, if there really were no standards that govern interpretation, if there were no distinction between an acceptable and an unacceptable interpretation, between
understanding and misunderstanding the same work, if members of the literary community did not even believe in the legitimacy of such distinctions, then, as Culler observes, "there would be little point to discussing and arguing about literary works and still less to writing about them." The entire corpus of interpretive criticism stands as proof, however, that the great majority of critics believe that there are certain standards of interpretive acceptability—and so, for that matter, does the vast system of academic departments devoted to the study of literary criticism. As Culler observes, most teachers presume that they are not simply introducing their students to a variety of particular literary works, but that their students are also gaining a "sense of how to read." They presume, in other words, that the students are also learning how to go about assigning literary works interpretations that fall within the limits of acceptability.

An anti-generative critic might argue that the foregoing observations only show that most members of the literary community believe that there are communally shared standards of acceptability, but they might be wrong; perhaps

60 Culler, *Structuralist Poetics*, p. 121.
61 Culler, *Structuralist Poetics*, p. 121.
most people think that everyone else accepts the same standards they accept, when, in fact, two individuals hardly ever share the same standards. Perhaps most members of the literary community believe, for example, that there are only a limited number of *Hamlet* interpretations that would receive any sort of widespread acceptance when, in fact, the large number of divergent interpretations that have already been proposed proves that they are mistaken. Indeed, Weitz's survey appears, at first sight, to support such a claim. Weitz himself repeatedly emphasizes the "multiple character of *Hamlet* criticism," and singles out interpretive or, as he calls it, "explanatory" criticism as "the area of most violent disagreement."\(^62\) Well, what does an examination of a particular interpretive dispute show?

Weitz suggests that one of the most—if not the most—frequently and hotly debated issues in *Hamlet* criticism concerns the question of why Hamlet delays so long before carrying out the task set for him by the Ghost.\(^63\) Hence, it should provide a good example. Weitz's survey shows that all of the following interpretations have been advanced at one time or another: Hamlet delays because 1. his moral

\(^{62}\) Weitz, p. 206.

\(^{63}\) Weitz, pp. 54, 208, 264-65.
scruples prevent him from performing the task (p. 4),
2. because he has a moral repugnance to the task (p. 4),
3. because he is too gentle for the task (p. 4, advanced by Goethe), 4. because his excessive intellectualism renders him inactive (p. 5, Coleridge), 5. because he suffers from melancholia (p. 5--A. C. Bradley, p. 61--Levin Schucking, and p. 84--Lily Campbell), 6. because undertaking the task would reactivate certain Oedipal fixations which he has long repressed and which are too horrible for him to bear (p. 20, Earnest Jones), and 7. because he is not sure that the Ghost is telling the truth until the play-within-the-play scene in Act III, Scene 2, and the only real chance he has to do the deed thereafter is spoiled by the fact that he believes if he killed Claudius at prayer, Claudius's soul would not be sent to Hell (pp. 4, 9--J. A. Waldock and p. 52--E. E. Stoll). 64

The standard critical reaction to a list like the foregoing invariably focuses exclusively on the diversity it exhibits, thereby passing over its other features, some of which are well worth emphasizing. For example, the list shows that critical interpretations do sometimes converge: the interpretation based on melancholia has been endorsed by

64 The page numbers indicated in parentheses indicate page numbers in Weitz's book. In some cases, the names of the most famous critics who have advanced the view in question are also indicated inside the parentheses.
three separate critics and the last interpretation has been endorsed by at least two different critics.

Second, critics are usually so busy emphasizing the number of different interpretations contained in the list that they never stop to consider the number of conceivable interpretations that it does not contain, for example: Hamlet delays because he is afraid his action would offend Polonius, because he does not want to risk upsetting the Danish economy, because he would really just rather go back to school, because he is hoping that his inaction will prompt a return visit by the Ghost—and on and on—indefinitely. In sharp contrast to the truly unlimited number of interpretations that could be included in the foregoing list is the number contained in the list complied from Weitz's book—seven. In response to what is supposedly the most frequently and hotly debated question about the most written-about play in all of dramatic literature only seven of the myriad possible explanations have been seriously advanced by the critics Weitz surveyed! Why did none of those critics advance one of the four interpretations preferred above? The reasons seem obvious: because no one would consider any of them to be even remotely acceptable, because critics do share some standards of acceptability within which they feel obliged to work, because, as Culler puts it, "certain expectations about [literature] and ways
of reading guide the interpretive process and impose severe limitations on the set of acceptable or plausible readings.\textsuperscript{65}

Were an anti-generative critic to concede the points made in the preceding paragraph, he might, however, still have some grounds for objecting to the suggestion that literary and linguistic understanding involve roughly similar capacities. He might, for example, point out the fact that the issue of Hamlet's delay represents only one of over eleven major sources of interpretive dispute isolated by Weitz, and, although none of those other ten issues have given rise to as many different positions as the issue of Hamlet's delay has, some of them have led to four or five different views.\textsuperscript{66} When most critics speak of "an interpretation" of a literary work, they are naturally referring not simply to a position adopted with respect to one single feature of the work, they are referring to a reading of the work as a whole. Thus, even if the number of different views towards each single issue is restricted by certain standards of acceptability, those views can be combined in various ways to produce a great number of different readings for the work as a whole. Indeed, if each of the eleven

\textsuperscript{65} Culler, \textit{Structuralist Poetics}, p. 127.

\textsuperscript{66} Weitz, p. 262.
major sources of dispute isolated by Weitz led to, on the average, five different acceptable positions, those positions could be combined to produce a total of 48,828,125 different readings of the play as a whole! Such an astronomical number, an anti-generative critic might maintain, far exceeds the number of acceptable interpretations that can be assigned to an ambiguous sentence. Thus, even if there are some standards of acceptability that restrict the class of permissible interpretations of a literary work to a finite set, that set is so much larger than the set of permissible interpretations assigned to any sentence that attempts to compare the capacities involved in literary and linguistic understanding should be rejected.

Several remarks could be made in response to such an argument. First, something could be said about possible incompatibilities between positions adopted with respect to different aspects of the same work. For example, one of the ten other major sources of dispute isolated by Weitz concerns critical attempts to identify Hamlet's predominant character traits, and he observes that one view on that issue, the one advanced by G. Wilson Knight, claims that

67 The above figure was obtained by taking the number of acceptable positions per issue, five, as a base figure. That figure was then raised to a power equal to the number of disputed issues, eleven. Thus, $5^{11} = 48,828,125$. 
callousness, cruelty, and brutality represent the character's major traits.\(^{68}\) Obviously, such an assessment of Hamlet's predominant traits is quite incompatible with Goethe's explanation of Hamlet's delay on the grounds that he is too gentle for the task.\(^{69}\) Given the assumption that a glaring internal inconsistency automatically disqualifies a reading from membership in the set of acceptable interpretations, the earlier total of possible interpretations must, therefore, be reduced by the number of possible readings that combine the two preceding positions, a number that is likely to run into the thousands. Furthermore, that represents only one of many cases where such incompatibilities arise. When all internally inconsistent readings are eliminated, the total number of acceptable interpretations will thus drop substantially.

Even if none of the various positions on different issues turned out to be incompatible, however, a second objection to the anti-generative argument could still be made. The argument assumes that the number of standard interpretations assigned to a sentence is always quite small. While that is certainly true in the case of most ambiguous sentences a person normally encounters, there is no reason to

\(^{68}\) Weitz, pp. 31-32, 263.

\(^{69}\) See the third interpretation in the list on p. 287.
believe that it must always be the case. If, for example, a sentence consisted of eleven coordinated main clauses, each of which was ambiguous in exactly five ways, it would give rise to a total of 48,828,125 acceptable interpretations—the same number assigned to Hamlet in the earlier calculation. Of course, speakers rarely encounter, much less concern themselves with such sentences. Nevertheless, the implicit knowledge that underlies the capacity for linguistic understanding appears to set limits on the number of interpretations that could be assigned to each grammatical sentence—whether the sentence in question is the sort normally encountered or not. Thus, even if the number of acceptable interpretations assigned to most literary works does far exceed the number assigned to most everyday sentences, that does not necessarily indicate that there is something fundamentally dissimilar about the capacities involved in each case. It may simply indicate that the sequences being assigned interpretations in the former case are a good deal longer and, as a result, more complex than those in the latter case—an observation that should come as no great surprise.

While the foregoing responses remove much of the initial air of plausibility that the earlier anti-generative argument might seem to possess, they do not, however, get at
the real, fundamental problem with that argument. That argu­
ment, like nearly everything else that has been said about
the topic thus far, implicitly assumes that critics' answers
to the sorts of questions that Weitz classifies under the
heading of interpretive or explanatory criticism constitute
just the sort of information that should be used in deter­
mining whether or not generative assertion 1A is plausible,
but that is a highly questionable assumption. In order to
see why, it will be necessary to differentiate between sep­
ate senses of the terms "interpretation" and "understand­
ing" as they are applied to sentences.

As earlier remarks have indicated, when Chomsky talks
about a native speaker's "interpretation" of a sentence, he
is usually referring to the speaker's structural description
of it--his tacit characterization of it in purely grammati­
cal, i.e., phonological, semantic, and syntactic, terms. 70
There is, however, another, less technical sense of the word
wherein an "interpretation" takes into account a variety of
other factors. A "non-grammatical" interpretation might,
for example, include hypotheses about what factors motivated
the speaker to utter a given sentence on a particular occa­
sion, hypotheses about what his utterance of the sentence on

70 See above, pp. 258-59.
that particular occasion implies about the society in which he lives, etc.

Naturally, a linguist would tend to ask a native speaker different sorts of questions depending on which sort of interpretation he is investigating. If he were interested in finding out about a native speaker's "grammatical" interpretation of a sentence, he would, for example, ask the speaker questions designed to reveal his intuitions about the sentence's immediate constituent structure, its paraphrase relations, etc. If he were, on the other hand, interested in finding out about the speaker's non-grammatical interpretation, he might ask the speaker to explain why he thinks the sentence was uttered by the person who said it, what possible social implications he thinks that has, etc. Whereas most generative linguists appear to believe that native-speaker answers to the former sorts of questions would exhibit a high degree of convergence in at least a substantial number of cases, it is unlikely that any linguist would wish to claim that the same would hold for the latter sorts of questions. When asked for a non-grammatical explanation of why a sentence contains a particular word, for example, even speakers who actually witnessed the speech act in which the sentence was uttered would be likely to offer a wide variety of answers, such as, for instance, the following: because the speaker was angry at the time,
because the speaker has an extensive vocabulary, because he likes to show off, because he comes from a particular socio-economic background, etc.

Thus, all of the earlier statements about the convergence of interpretations towards a standard or a standard set apply only in the case of grammatical interpretations; it may turn out that there are no uniform standards when it comes to non-grammatical interpretations. Likewise, earlier statements that native speakers understand sentences in roughly the same way must be limited to grammatical understanding; there would appear to be little evidence that speakers understand sentences in the same way where a very broad, all-encompassing sense of "understanding" is involved.

It would appear that the distinction between "grammatical" and "non-grammatical" interpretations can also be applied within the domain of literature. Just as a linguist usually asks native speakers questions concerned with immediate constituent structure and paraphrase relations when he wishes to find out about the grammatical interpretations they assign to sentences, a literary theorist might want to ask roughly parallel questions. In the case of Hamlet, for example, he might ask a critic to identify the major "joints" or "breaks" that divide Hamlet into major segments or immediate constituents, and he might ask the critic to
give a fairly complete synopsis or paraphrase of the play's plot.

Most literary critics and theorists would probably say that answers to such questions constitute no more than a preliminary step towards assigning an "interpretation" to Hamlet. They would focus instead on answers to the sorts of questions that Weitz groups under the heading of interpretive or explanatory criticism, answers that invariably involve a variety of non-grammatical factors. In other words, they tend to focus on the sort of information that would be relevant to determining a critic's "non-grammatical" interpretation. Indeed, when literary critics and theorists talk about an "interpretation" of a work, they almost always seem to have a non-grammatical rather than a grammatical interpretation in mind.

As earlier remarks have suggested, it would be a serious mistake to assume that since the non-grammatical interpretations assigned to sentences by native speakers do not converge, neither do their grammatical interpretations; and such an assumption would be equally unwarranted in the case of literature. In fact, most linguists suggest that the information contained in a speaker's non-grammatical interpretation of a sentence is of little use in attempting to

---

71 See, for example, Weitz, pp. 245-68.
determine his grammatical interpretation of it. Much the same may well be true in the case of literature. Thus, most of the information that was implicitly assumed to be crucially relevant to determining the plausibility of generative assertion 1A during earlier discussion turns out not to be so.

The plausibility of generative assertion 1A should accordingly be determined on the basis of whether most critics and theorists would offer roughly similar grammatical—not non-grammatical—interpretations of a given literary work. It should, in other words, be determined on the basis of their answers to questions about the location of the major joints or breaks in a play and questions about the proper synopsis of a play's plot. Would most critics and theorists offer similar responses to such questions?

Before examining the actual evidence, it is worth noting that educated common sense seems to yield an affirmative answer to both questions. It seems reasonable to suppose, as Bloomfield might have put it, that any proficient individual who concerns himself with the matter is sure to say that the major immediate constituent boundary, the major joint or break, in *Hamlet* occurs in Act III, Scene 2, when

72 See, for example, Chomsky's remarks on pp. 188-89 of *Rules*. 
the play-within-the-play exposes Claudius's guilt.\textsuperscript{73} It seems reasonable to suppose that they would also say that the major boundary between the beginning of the play and III, 2 occurs in I, 5, during Hamlet's encounter with the Ghost; and that the major boundaries between III, 2 and the end of the play occur in III, 3, when Hamlet fails to kill Claudius at prayer, in III, 4, when Hamlet kills Polonius, in IV, 7, when Claudius and Laertes join forces, and, of course, in V, 2, during the duel. It also seems reasonable to suppose that most critics would offer roughly similar synopses of Hamlet's plot.

The preceding remarks represent, however, no more than a hypothesis based on the sometimes misleading dictates of educated common sense. The real question is whether the corpus of critical literature contains enough positive evidence to confirm that hypothesis decisively. What, for example, does Weitz's survey show? It contains some evidence that critics tend to agree when they concern themselves with questions directly related to Hamlet's plot, questions which Weitz classifies under the heading of "descriptive" rather than interpretive or "explanatory" criticism. For example, as Weitz observes, the survey provides ample evidence that

\textsuperscript{73} Bloomfield's roughly parallel remark is quoted above, p. 238.
virtually all critics have offered an affirmative answer to the "grammatical" question "Does Hamlet delay?", which is, of course, to be distinguished from the explanatory question "Why does Hamlet delay?" However, evidence of the foregoing sort is rather sparse, and the survey contains almost no discussions of immediate constituent structure. Thus, although Weitz's survey contains some positive evidence, it does not appear to contain enough positive evidence to confirm the earlier hypothesis decisively. However, it does not contain any evidence that would disconfirm the hypothesis either, and the following remarks will show that the very lack of critical commentary about matters of immediate constituent structure and paraphrase relations constitutes a sort of negative evidence that provides a fair amount of support for the hypothesis in

74 Weitz, p. 240.

75 There are, however, a few exceptions to the foregoing statement. For example, as Weitz observes, Francis Fergusson does speculate about the immediate constituent structure of Hamlet. His account is similar to the one advanced above, except that instead of indicating a single major boundary, and then proceeding to identify sub-boundaries, he simply identifies a series of major boundaries (Weitz, p. 97). That should, however, come as no surprise. If a native speaker were simply asked to identify the major constituents of Poor John ran away, he might give the answer advanced by Bloomfield, identifying the largest constituent groupings first (see above, p. 238), or he might simply identify each of the four words as major constituents. The latter answer does not, of course, preclude his thinking that the first two and the last two words also form larger constituents.
question.

At the close of his chapter on descriptive criticism, Weitz observes that "descriptions are basic in *Hamlet* criticism since explanation and evaluation depend on them." In other words, descriptive passages, like those which bear directly upon a play's plot or its immediate constituent structure, are basic because they function as the evidence that critics use in arguing about the acceptability of various explanations and evaluations of literary works. Now it seems reasonable to suppose that, as in all other forms of argumentation, literary critics attempt to convince their audience that they should accept a conclusion, about which there could be some disagreement, on the grounds that it follows from certain evidence—about which there is little or no disagreement. So, when critics adduce as evidence remarks that bear directly upon the paraphrase of a play's plot, there is good reason to believe that there is general agreement about those paraphrase relations. Moreover, the very fact that Weitz classifies such remarks under the heading of "descriptive" criticism lends additional support to that supposition; whereas individuals often disagree about explanatory matters, they tend to exhibit widespread

---

76 Weitz, p. 244.
agreement about descriptive matters.

If, however, the foregoing remarks are true, why does Weitz's survey expose so few examples of agreement concerning matters of immediate constituent structure and paraphrase relations? Three reasons can be given. First, since such matters are included in critical texts only as evidence offered in support of particular explanatory or evaluative points, complete synopses and descriptions of immediate constituent structure are hardly ever included in such texts. Instead, critical essays usually contain a group of remarks, each of which bears on only one specific part of a work's plot or immediate constituent structure, remarks which are scattered throughout the text in which they occur. Second, critics seldom—if ever—refer to another critics' remarks about such matters, a phenomenon which may plausibly be attributed to the thoroughly uncontroversial character of such remarks.

Finally, many critics rely on an implicit understanding of such matters without ever explicitly mentioning them. For example, a critic might proceed directly to the question of why Hamlet delays without even bothering to address the question of whether he does in fact delay; the critic simply assumes that the answer to the latter question is too obvious to merit explicit mention. Naturally, most critics
do not wish to dwell on the obvious; they are anxious to move on to the business of explanation and evaluation. They are, in other words, anxious to push ahead to the kind of material that will allow them to say something original, a desire that has been deeply instilled in them by the institution of criticism itself. Indeed, one of the most basic tenets of critical practice can be stated, after the style of Thumper's Mother, as follows: if you can't say anything original, don't say anything at all!

Realization of the foregoing fact about critical practice provides several important insights into the relation between the corpus of critical literature and the study of literary competence. As earlier remarks have suggested, the study of literary competence represents an attempt to explain, among other things, the phenomenon of rough agreement when it comes to grammatical understanding of literary works. As such, that study will attempt to locate data that reflect such agreement, and the corpus of critical literature would seem to be the natural place to look for such data. That is not, however, the case. Because the practice of literary criticism is guided by the maxim of originality, critics tend to avoid spending much time agreeing with what previous critics have said about a particular work; to disagree with what previous critics have said about a given
work is to say something original, but to dwell on points of agreement, except briefly in passing, is to reiterate what has already been said— it is to violate the maxim of originality. Thus, the corpus of critical literature is likely to provide very little positive evidence that can be used in the study of literary competence.

Of course, realization of the role that the maxim of originality plays in critical practice also helps to explain why certain matters, e.g., those having to do with a work's implications and explanation, receive a great deal of attention in the critical literature while others, e.g., those having to do with immediate constituent structure and paraphrase, are hardly discussed at all. The maxim of originality encourages critics to focus on the sorts of issues with respect to which there may be substantial disagreement, and it encourages them to ignore matters about which there could be little room for disagreement. As a result, the absence of discussion about a subject, even though it only constitutes a sort of negative evidence, may often represent the strongest—often the only—evidence in the critical literature that there is substantial agreement about that subject. Thus, the noticeable lack of critical commentary about immediate constituent structure and paraphrase relations constitutes negative evidence that there is, in fact,
There is, then, at least some positive and fairly strong negative evidence that the grammatical interpretations assigned to a given work by a variety of proficient critics and theorists would, at least in some cases, 

77 Although they do not bear directly on the relation between the corpus of critical literature and the study of literary competence, a few additional remarks concerning the effect of the maxim of originality on literary practice seem to be warranted in order to shed some light on a doctrine held by some deconstructionists. Some deconstructionists attribute the presence of overwhelming discord in the corpus of critical literature to the nature of literary works themselves. Miller, for example, suggests that critical disagreement derives from the fact that a literary work has, as he puts it, no single "ground," no source of unity and coherence ("Steven's Rock," p. 333). As a result, claims Miller, every literary work is "incapable of being encompassed in any single coherent or homogeneous interpretation" ("Steven's Rock," p. 333—incidentally, Miller would presumably claim that the foregoing holds true with respect to both non-grammatical and grammatical interpretations). Thus, Miller argues that the presence of widespread critical discord derives from the fact that literary works are fundamentally unstable, deconstructive entities. The foregoing remarks in the main text suggest that there may, however, be a more plausible explanation than the one advanced by Miller: the primary cause of critical discord lies not in the nature of literary works, but in the fact that the maxim of originality tends to discourage critics from writing about matters of widespread agreement and encourages them to focus, instead, on matters that allow for substantial disagreement. Since the entire practice of literary criticism, from a publisher's decision about what to put in print down to the aspiring student's choice of paper topics is guided by that maxim, it is small wonder that the corpus of critical literature exhibits a noticeable lack of unanimity—it could hardly be otherwise.
converge enough to allow abstraction of a standardized set of such interpretations—which supports the contention that there is such a thing as the capacity to understand literary works, in the strong sense, after all. Anti-generative attempts to refute generative assertion 1A by proving that there is no such capacity may therefore be rejected. As earlier remarks have indicated, there are, however, two other strategies that a critic might wish to employ in arguing against assertion 1A. He might wish to argue that although proficiency with respect to a particular kind of literary work entails a capacity to understand at least some works of that kind, proficient individuals do not possess a capacity to understand a potentially infinite number of works of that kind; or he might wish to argue that they possess a capacity to understand a potentially infinite number of such works, but only so long as the works are based on a finite number of underlying syntactic patterns.

Neither of those strategies have, however, been pursued by anti-generative writers, and it is not difficult to see why. As Culler observes, most anti-generative critics firmly insist that the creation and interpretation of literary works are by nature totally free and unrestricted
activities. As a result, such critics would naturally not be inclined to argue that a person's interpretive abilities are restricted to a finite set of works or to works based on a finite set of underlying patterns.

There are, however, more compelling reasons for thinking that arguments based on either of the two other strategies should be rejected. If a person who is proficient with respect to a given type of work were capable of understanding only a limited number of such works or only works based on a limited number of underlying patterns, everytime he encountered an unfamiliar work or a work based on an unfamiliar pattern, he could be expected to be quite baffled. However, that does not appear to occur; a person who is able to advance acceptable grammatical interpretations of two or three works of a particular kind, say neo-classical tragedies or romantic comedies, is rarely baffled by unfamiliar works of that kind or works based patterns that are dissimilar to those exhibited by the works he already understands. Since all three avenues of argument against generative assertion 1A appear to be unsuccessful, that assertion may be tentatively accepted as plausible.

---

78 Culler, *Structuralist Poetics*, pp. 241-42.
Generative assertion 1B, the claim that a person who possesses competence with respect to a particular type or kind of literary works is capable of evaluating the syntactic well-formedness of an indefinitely large number of such works based on an indefinitely large number of different syntactic patterns, exhibits much the same sort of ambiguity as does assertion 1A: it could be construed as making the extremely weak claim that proficient individuals are able to make some sort of evaluative assessments of literary works, or it could be construed as making the stronger claim that proficient individuals are capable of making evaluative judgments that converge to a high degree.

Anti-generative critics would not, of course, wish to deny the weak claim; it is a matter of obvious fact that individuals are capable of making some sort of evaluative judgments about literary works. It is, instead, the plausibility of the stronger claim, which is indeed entailed by generative assertion 1B, that anti-generative authors have wished to deny; and their major argument closely resembles the major argument advanced against assertion 1A: individuals could not possess the capacity in question because evaluative judgments of literary works do not converge the way evaluative judgments of sentences do.
Even Culler, who attempts to defend the generative approach, concedes the foregoing point:

In linguistics it is not difficult to identify facts that an adequate grammar must account for: though one may need to speak of 'degrees of grammaticalness' one can produce lists of sentences which are unquestionably deviant. . . . In the study of literature, however, the situation is considerably more complex. Notions of 'well-formed' or 'intelligible' literary works are notoriously problematic.

Pettit agrees, adding that the apparent lack of consensus concerning literary well-formedness constitutes an insurmountable problem for generative theory:

If the assigning of structure is to be represented in a generative grammar then this condition at least must be fulfilled: that it is intuitively clear when a text has structure and when it fails to have structure, when it is allowable and when it is unallowable. The corresponding condition is met by sentences because his intuition tells the native speaker when a sentence is grammatical and when it is ungrammatical. . . . there is [however] no intuition by which a reader can definitely say that a text lacks structure and is unallowable.

79 Culler, Structuralist Poetics, p. 122.

80 Pettit, p. 47. Pettit's references to a text's "lacking structure" are somewhat baffling. As earlier remarks have indicated (see above, pp. 30-31), a text could not possibly fail to have a structure; having an ill-formed structure is not the same as having no structure at all.
Thus, anti-generative writers like Pettit reject generative assertion IB, claiming that there is no consensus within the literary community concerning matters of literary well-formedness.

As was the case with the main objection advanced against generative assertion 1A, Pettit's objection possesses an initial air of plausibility; critics do seem to be forever quarreling about not only the interpretation, but also the evaluation of literary works. In fact, Weitz's survey of *Hamlet* criticism would, once again, appear to furnish empirical evidence in support of that statement.\(^8\)

Subsequent remarks will, however, show that much of the apparent plausibility of the claim that there is little widespread agreement derives from a failure to recognize several crucial features of actual critical disputes about the well-formedness of literary works; but first something should be said about Pettit's contention that lack of critical consensus concerning the well-formedness of literary works would completely undermine the generative approach.

Even if Pettit were right in thinking that there is no consensus about the well-formedness of literary works, there

---

\(^8\) Weitz, pp. 269-84.
is reason to believe that he has misjudged the import that would have for generative theory. In their book *A Generative Theory of Tonal Music*, which constitutes by far the most elaborate generative theory of an art form to date, linguist Ray Jackendoff and composer-musicologist Fred Lerdahl argue that the notion of grammaticality is not as fundamental to the generative approach in linguistics as many people have thought—and that it certainly need not play a fundamental role in generative theories of other sorts of phenomena:

The early work in [generative linguistics], such as Chomsky 1957 and Lees 1960, took as its goal the description of "all and only" the sentences of a language, and many were led to think of a generative grammar as an algorithm to manufacture grammatical sentences. . . . [However], it was pointed out by Chomsky and Miller (1963), and it has been an unquestioned assumption of actual research in linguistics, that what is really of interest in a generative grammar is the structure it assigns to sentences, not which strings of words are or are not grammatical sentences. The same holds true for our theory of music. It is not intended to enumerate what pieces are possible, but to specify a structural description for any tonal piece; that is, the structure that the experienced listener infers in his hearing of the piece.  

If Jackendoff and Lerdahl are right, then critics like

---

Pettit are mistaken in thinking that the plausibility of developing a generative approach to literature depends crucially on the question of whether there is widespread agreement about the well-formedness of literary works. It is not, however, clear that critics like Pettit are even right in thinking that there is little widespread agreement about literary well-formedness.

As mentioned above, much of the initial plausibility of such claims derives from a failure to recognize certain important features of critical disputes about the well-formedness of literary works. For example, neither Pettit nor Culler point out that such critical disputes often concern matters of semantic or pragmatic well-formedness rather than syntactic well-formedness. Weitz observes, for instance, that evaluative disagreements about Hamlet involve a wide variety of different concerns, including

originality, metaphysical or moral truth, true-to-life characterization, individuality, variety, and consistency of character, power of evocation, subtlety of exhibition of passion, ability to create and sustain illusion, and organic integration of the various characters, language, plot, and theme. 83

83 Weitz, p. 271.
In light of what was said earlier about the literary application of the distinction between syntax, semantics, and pragmatics, it is clear that most of the concerns cited by Weitz, e.g., concerns over true-to-life-characterization, metaphysical or moral truth, originality, etc., fall within the domains of semantics and pragmatics. Since generative assertion I.B only requires that there be widespread agreement about matters of syntactic well-formedness, critical disagreements over matters of semantic and pragmatic well-formedness do not have any real bearing on its plausibility.

A closer look at evaluative quarrels would also show that even when critics disagree about the syntactic well-formedness of a work, they are usually arguing about how a well-formed work fares when compared with other unquestionably well-formed works. In other words, they usually assume, although they often fail to say so explicitly, that the work in question falls within the range of well-formed works, and they then proceed to argue about degrees of grammaticalness within the class of works that exhibit syntac-

---

84 See above, pp. 273-75.
tically well-formed patterns.\textsuperscript{85} Most of the critics Weitz discusses do not, for example, wish to claim that \textit{Hamlet} exhibits a positively ill-formed pattern; they wish, instead, to claim that the pattern exhibited by \textit{Hamlet} is more or that it is less well-formed than patterns exhibited by other syntactically well-formed plays, either ones also written by Shakespeare or ones written by other authors.\textsuperscript{86}

Exceptions to the foregoing generalization can, however, be found. For example, Weitz draws attention to "Voltaire's series of attacks against Shakespeare's violation of the rules of drama, especially the three unities ... and his mixture of the comic and the tragic."\textsuperscript{87} Voltaire's criticisms are of considerable interest because they represent an attempt to claim that at least some of

\begin{quotation}
\textsuperscript{85} The fact that most critics spend little—if any—time acknowledging that a particular work does indeed fall within the class of well-formed works and a great deal of time discussing exactly where, within that class, the work falls is, of course, in keeping with the maxim of originality. Since the overwhelming majority of critics think that \textit{Hamlet}, for instance, is, in general, well-formed, dwelling on that point would allow a critic to say very little of an original nature. However, there is a good deal of room for disagreement and originality when it comes to the question of exactly where, within the class of well-formed works, \textit{Hamlet} should be placed.

\textsuperscript{86} Weitz, pp. 269-84.

\textsuperscript{87} Weitz, p. 146.
\end{quotation}
Shakespeare's plays are ill-formed because they fail to follow the rules fundamental to one particular style or "dialect" within the "language" of drama, those of neoclassicism. A great number of cases in which critics have quarreled about whether a work even belongs to the class of well-formed works would appear to parallel the foregoing case; they typically involve critics who refuse to acknowledge the fact that there are a number of different literary "dialects" and that works belonging to different dialects may legitimately exhibit quite different underlying patterns.

Disputes of a similar nature have, of course, been known to arise among speakers of differing linguistic dialects. Such disputes tend to occur more frequently in the realm of literature, however, because the "dialects" of literature change more often and more abruptly than do linguistic dialects. As a result, a new literary dialect, e.g., romanticism, may spring into existence all at once, leaving many critics in the position of a speaker who has suddenly stumbled upon an entirely unfamiliar dialect. Like

88 For a discussion of the sorts of hostilities that can arise between speakers of different linguistic dialects, see Bolinger's remarks concerning the tension between speakers of "standard white English" and speakers of "non-standard black English" (p. 365).
such a speaker, critics are likely to be baffled by the unfamiliar dialect and are, therefore, likely to assess works written in the dialect as ill-formed.

Recognition that there are a plurality of different literary dialects, each of which may employ different syntactic patterns, suggests that there is a problem inherent in Pettit's discussion of well-formedness much more basic than any of the problems mentioned so far. His claim that there is "no intuition by which a reader can definitely say that a text lacks structure and is unallowable" implies that it ought to be clear that some texts are unallowable—without qualification. Comparison with the case of language shows, however, that such a requirement is far too strong. Native speakers do not have intuitions that a given sequence is ungrammatical in an unqualified sense; they have intuitions that a sequence is an ungrammatical sentence of English, or French, etc. Judgments about grammaticality are, in other words, always to be understood relative to a particular language or dialect.

Thus, the plausibility of generative assertion 1B should not be determined on the basis of whether competent readers have strong intuitions that a text is unallowable without qualification, but, instead, on the basis of whether they have strong intuitions that a text is ungrammatical
with respect to a particular style or genre; and surely that is the case. Just as any competent English speaker knows immediately that a sequence of some other language, of French or Japanese, etc., does not constitute a well-formed sentence of English, any critic who possesses proficiency with respect to, for example, romantic comedies knows immediately that *Oedipus*, or *The Little Foxes*, or, to take an even more extreme example, Tristan Tzara's *The Gas Heart*, does not constitute a well-formed romantic comedy. It is clear, moreover, that such judgments do not concern semantic or pragmatic features of the plays. In explaining why *Oedipus* is not, for example, a well-formed romantic comedy, a critic would not be likely to mention matters concerned with situational context, or reference, or truth, etc. The issue is, instead, one of syntax; *Oedipus* exhibits an underlying syntactic pattern that is perfectly well-formed relative to the dialect of classical tragedy--it is simply not a pattern that is well-formed relative to the dialect of romantic comedy.

Once the relativity necessarily inherent in all judgments of syntactic well-formedness is recognized, it becomes clear, moreover, that individuals who are proficient with respect to a particular type of literary work possess a capacity to tell whether each of an indefinitely large number of
works based on an indefinitely large number of underlying syntactic patterns constitutes a syntactically well-formed example of that type. Of course, there are bound to be borderline cases, but that is true in the case of language as well. Thus, there appear to be compelling reasons for believing that generative assertion 1B is plausible.

Few anti-generative critics have attempted to deny the plausibility of generative assertion 2, the assertion that a reader could not understand or evaluate the syntactic well-formedness of a literary work without assigning a structural description to it. Even the weak capacity to interpret a literary work in any way whatsoever and the corresponding weak interpretation of the capacity to issue evaluative judgments about syntactic well-formedness seem to presuppose that a reader has assigned the work in question some sort of structural description. Indeed, even anti-generative writers like Pettit agree that "just as with a sentence, a [literary] text can be given meaning, can be interpreted, only when it is assigned a structure."\(^89\) Little more need, therefore, be said about assertion 2.

\(^89\) Pettit, p. 47.
Although few of the current anti-generative critics have explicitly argued that theories of literature should positively avoid any attempt to identify the underlying source or sources of literary comprehension and evaluation as generative assertion 3 claims they should, it is by no means inconceivable that a critic might wish to do so. A writer who favored an "intuitionist" view of literary comprehension and evaluation might wish to advance something like the following argument: although an individual can come to know the meaning of a literary work, and he can come to know whether a work is syntactically well- or ill-formed, he does not acquire such knowledge by performing a series of cognitive operations that are ultimately based on some pre-existent system of internalized knowledge. According to an intuitionist, the comprehension and evaluation of literary works occurs, instead, by means of "immediate apprehension," i.e., an individual simply "apprehends" the meaning or the value of a literary work in the same sort of "unmediated" way that is involved in, for example, his coming to know that a red object is in front of him--he simply grasps the object of knowledge in an immediate way. Beyond the foregoing observation, the nature of literary understanding and evaluation is taken to be fundamentally inexplicable, and the source or sources of those capacities is taken to be
fundamentally ineffable. Attempts to explain such processes by locating some definite body of knowledge that constitutes their ultimate source will inevitably fail, and it is, therefore, pointless to expect a theory of literature to undertake such a task.\textsuperscript{90}

The foregoing position appears to be objectionable primarily on the grounds that it fails to provide a persuasive explanation; it amounts to a claim that the phenomena in question can be explained only by positing the existence of a mysterious process that is itself fundamentally inexplicable. Thus, after one rather questionable move, the explanatory process is brought to a stop. If every other possible explanatory hypothesis were to prove fruitless, critics and theorists might wish to accept such a claim— but only as a last resort. Thus, it seems reasonable to accept generative assertion 3's claim that there is some underlying system of literary knowledge as the most plausible hypothesis at

\textsuperscript{90} For a discussion of various writers who have advanced intuitionist accounts of aesthetic comprehension and evaluation, see Monroe C. Beardsley \textit{Aesthetics from Classical Greece to the Present} (New York: Macmillan Co., 1966), pp. 178-83, 231-79). Such accounts typically differ with respect to a number of details concerning the nature of "intuition," "immediate apprehension," etc.; the foregoing remarks in the main text thus represent no more than a vague, generalized sketch of the intuitionist position.
present.

One more objection to assertion 3 merits comment. Earlier remarks have indicated that assertions 1A and 1B are concerned only with the rather limited capacities of grammatical understanding and syntactic evaluation, and it is the explanation of those capacities and the related capacity to assign structural descriptions to literary works which assertion 3 claims should constitute the primary goal of literary study. Anti-generative critics might claim that in focusing attention on the explanation of how it is possible for readers to perform activities like assigning works acceptable plot synopses and immediate constituent structure descriptions generative assertion 3 draws attention away from the really important, "higher" capacities that constitute the major focus of literary criticism. Formulating an acceptable synopsis of a work's plot and a description of its immediate constituent structure are, they might argue, activities so simple that nearly everyone is capable of performing them; they are activities too simple to need explanation. The really pressing problem, they might continue, is how to explain an individual's capacity

\[91\] See above, pp. 295-97.
to formulate a non-grammatical interpretation of a literary work.

Part of the foregoing objection has considerable merit: formulating an explanation of the capacity for non-grammatical interpretation is certainly a pressing problem. As earlier remarks have indicated, the process of non-grammatical interpretation is governed by two opposing constraints. On the one hand, there appear to be certain standards of acceptability that restrict even the set of permissible non-grammatical interpretations of a work. On the other hand, the maxim of originality obliges a critic to formulate a non-grammatical interpretation of a work that has never before been advanced. Thus, a critic is obliged to be original without venturing beyond the boundaries of communal acceptability. That is the real challenge of criticism, and anyone who has written critical essays knows that it is not an easy challenge to meet. The explanation of what enables critics to meet that challenge is, at present, far from clear, and work on that topic is certainly important.

As regards the claim that generative assertion 3 focuses attention on activities that are too simple to merit explanation, however, a comparison with language suggests

\[^92\] See above, pp. 288-89.
that they may not be nearly as simple as they seem. Although grammatical comprehension and syntactic evaluation of sentences would appear to be simple activities, the explanation of those activities has turned out to be enormously complex.

Moreover, it may well turn out that progress in explaining the activities involved in non-grammatical interpretation will depend heavily on progress in explaining the "simple" capacities involved in grammatical interpretation and syntactic evaluation. It is of interest to note what Jackendoff and Lerdahl have to say about such matters as they arise in the case of music:

Music theory that is oriented toward explicating masterpieces tends to address primarily those aspects of [music] that are complex, ambiguous, or controversial. But such discussion takes for granted a vast substrate of totally "obvious" organization that defines the terms in which artistic options or questions are stated. . . . Throughout this study we come to grips with such musically mundane matters as a basis for understanding the more complex phenomena that an "artistic" theory deems worthy of interest.

Uninteresting though such an enterprise may at first seem, it has proved to us to yield two important benefits in the understanding of music. First, one comes to realize how intricate even the "obvious" aspects of musical organization are . . . These aspects only seem simple because their complexity is unconscious and hence unnoticed. Second, one can begin to see how
artistically interesting phenomena result from the manipulation of the parameters responsible for "obvious" intuitions. 93

It seems probable that much the same will hold true in the domain of literature.

In light of what has been said above, it does, however, seem advisable to modify the original statement of generative assertion 3 in one respect. As originally stated, the assertion claims that the primary goal in studying works of a particular kind or type ought to be that of formulating an account of the literary knowledge that partly underlies the ability to assign structural descriptions to works of that kind, and which therefore also underlies the abilities to understand and syntactically evaluate them. 94 Since explanation of the capacities involved involved in non-grammatical interpretation presumably also constitutes a major goal, assertion 3 should only claim that formulation of the sort of account in question constitutes a primary goal of literary theory. In that form, assertion 3 may be accepted as plausible.

As Culler's discussion of recent opposition to generative theory indicates, much of that opposition seems to be

93 Jackendoff and Lerdahl, p. 7.
94 See above, p. 278.
based on a deep-seated hostility towards its rule-oriented character. Thus, most anti-generative critics would presumably wish to attack that feature of generative assertion 4—the claim that an adequate account of the literary knowledge that partly underlies a reader's ability to assign structural descriptions to works of a particular type or kind should at least enable a theorist to simulate the reader's ability to generate syntactic descriptions for each of a potentially infinite number of such works based on an indefinitely large number of different syntactic patterns and must, as a consequence, contain, among other things, a set of interactive formation rules, at least some of which are recursive.

Widespread opposition to rule-oriented theories of literature is not, of course, new to the canon of critical discourse. Such opposition can be traced back at least to the advent of Romanticism when, for example, Victor Hugo complained that the rule-oriented character of neo-classical dramatic theory had resulted in an unending series of "wretched quibbles with which mediocrity, envy and routine

95 Culler, *Structuralist Poetics*, pp. 241-43.
The very idea of a rule-oriented theory of literature has been anathema to practically every critic since that time and for much the same reason: rule-orientation is thought to entail a prescriptive attitude that will inevitably end up restricting the creative freedom of authors and readers.

The foregoing line of reasoning is, for example, reflected in John Dewey's attack on the enterprise of poetics:

the notion of fixed classes and that of fixed rules inevitably accompany each other. If there are, for example, so many separate genres in literature, then there is some immutable principle which marks off each kind and which defines an inherent essence that makes each species what it is. This principle must then be conformed to; otherwise the "nature" that belongs to art will be violated and "bad" art will be the result. Instead of being free to do what he can with the material at hand and the media under his control, the artist is bound, under penalty of rebuke from the critic who knows the rules, to follow the precepts that flow from the basic principle. Instead of observing the subject matter, he observes the rules. Thus, classification sets limits to perception. If the theory that underlies it is influential, it restricts creative work.

---


Interestingly enough, Culler observes that a number of relatively traditionally oriented writers currently oppose generative theory for the very reason that it represents a form of poetics. In view of Dewey's remarks, it is easy to see why: if generative theory is a form of poetics, then it must be rule-oriented, and if it is rule-oriented, it is presumed to constitute a prescriptive attempt to stifle the creative freedom of both artists and readers.

Of even greater interest is Culler's observation that avant-garde critics, like those who belong to the Tel Quel and deconstructionist movements, base their opposition to generative theory on the following, rather similar argument:

Texts can be read in many ways; each text contains within itself the possibility of an infinite set of structures, and to privilege some by setting up a system of rules to generate them is a blatantly prescriptive and ideological move.

As the foregoing quotation suggests, the argument advanced against generative theory by such avant-garde critics is simply a variation on the arguments that have been used against rule-oriented theories since the advent of

98 Culler, Structuralist Poetics, pp. 242-43.

99 Culler, Structuralist Poetics, p. 242; see also Norris, pp. 32, 90-99.
Romanticism; because it is rule-oriented, generative theory is presumed to entail a prescriptive attitude that will inevitably end up stifling the creative freedom of both authors and readers. Hugo urged his contemporaries to "take the hammer" to rule-oriented theories of literature—to kill the vile monsters—and contemporary writers, those of avant-garde as well as those of more traditional affiliations, see generative theory as an attempt to bring the monsters back to life—an attempt that must be stopped.100

Interestingly enough, some linguists have also exhibited a tendency to link the use of rules with a prescriptive orientation. For example, Saussure criticized the Port Royal Grammar on the grounds that "it is normative and assumes the role of prescribing rules, not of recording facts."101 Saussure's tendency to posit a sharp dichotomy between the formulation of rules and the description of facts is further manifested in his discussion of syntagmatic structure, where he speaks exclusively in terms of "general types" and "regular patterns," scrupulously avoiding any

100 Hugo, p. 368.

101 Saussure, p. 82. Chomsky has, incidentally argued that the charge of prescriptivism is totally unfounded in the case of the Port Royal Grammar (Mind, pp. 14-15).
Such remarks suggest that Saussure held something like the following view: a linguist who speaks of regular patterns is simply recording the fact that certain patterns tend to recur in the utterances of native speakers, but a linguist who speaks of rules automatically implies that the utterances of native speakers not only do, but that they should exhibit those—and only those—patterns that accord with a particular set of rules; he is, in other words, putting himself in the position of dictating, or prescribing, proper usage—regardless of whether he intends to do so or even realizes that he is doing so.

As earlier remarks have indicated, the Bloomfieldian linguists also exhibited a predilection for talking in terms of distributional patterns and sentential formulae rather than in terms of rules. Moreover, as Lyons points out, the Bloomfieldians were extremely anxious to avoid the slightest trace of prescriptivism. In view of the fact that Bloomfield himself repeatedly associated normative grammar with the use of rules, citing the Port Royal Grammar as an example, it does not seem entirely unlikely that

---

102 Saussure, pp. 124-25.
103 See above, pp. 239-40.
104 Lyons, Chomsky, p. 37.
the Bloomfieldian's hostile attitude toward rule-orientation was based on a view much like Saussure's.\textsuperscript{105}

The fact that Saussure and Bloomfield both cite the \textit{Port Royal Grammar} as an example of the connection between rule-orientation and prescriptivism is of additional interest because the paradigm example of the alleged connection in literary theory, neo-classicism, was developed during the same era, the seventeenth-century. There may, in fact, have been some indirect contact between the founders of the two seventeenth-century theories by way of the French Academy; the Academy was very much involved in formulating both literary and linguistic standards of correctness, and it may quite possibly have exerted some influence on the Port Royal Grammarians.\textsuperscript{106} Thus, it would appear that the monster of linguistic theory and the monster of literary theory were veritable cousins.

The foregoing similarity does not, moreover, constitute the only interesting parallel between the history of the two

\textsuperscript{105} Bloomfield, pp. 6-8. If the Bloomfieldians did in fact associate rule-orientation with prescriptivism, that might help to explain their initial hostility towards generative grammar: in addition to various other factors, their views on the prescriptive effects of rule-orientation may have led them to see the manifestly rule-oriented nature of Chomsky's work as a departure from a purely descriptive approach to language.

disciplines. As Lyons indicates, widespread dissatisfaction with the rationalistic character of theories like the Port Royal Grammar surfaced during the period extending from the latter part of the eighteenth- through the early part of the nineteenth-centuries, the era which also marked the beginning of the romantic rebellion against neo-classicism. 107 Thus, the monsters of literature and linguistics were not only born during the same era, they were also slain during the same era. Moreover, like the former event, the latter event appears to have involved some contact between the two fields. As Lyons goes on to observe, the romantic critics' revolt against the rationalistic, rule-oriented character of neo-classicism exerted a good deal of influence on the linguistic work done at the time. 108 Thus, not only were the monsters cousins, they were slain by blood-brothers.

As previous remarks have indicated, both Saussure and Bloomfield exhibited a tendency to associate rule-orientation with prescriptivism, and the remarks contained in the last paragraph suggest that their tendency to do so may, perhaps, have been tied to a tradition dating back to the eighteenth-century, as the corresponding tendency among literary

107 Lyons, Theoretical, pp. 22-23.
critics most certainly is. Whereas the vast majority of contemporary literary critics continue to associate the two characteristics, however, contemporary linguists have ceased to do so.

As Lyons indicates, the distinction between descriptive and prescriptive grammar can be drawn even in cases where both approaches are rule-oriented, the distinction consisting in a difference "between the description of rules that are actually followed by native speakers and the prescription of rules that, in the opinion of the grammarian, they ought to follow, in order to speak 'correctly.'" Thus, the distinction appears to have very little to do with rule-orientation. It is, instead, based on a linguist's attitude toward the relation between his grammar and the behavior of native speakers. A descriptive linguist believes that his grammar should mirror the linguistic attitudes and behavior of native speakers; a prescriptive linguist believes that the linguistic attitudes and behavior of native speakers should mirror his grammar.

Of course, some of the linguistic attitudes that a descriptive linguist wishes to describe are prescriptive in

109 Lyons, Chomsky, pp. 36-37.
character, e.g., native speaker judgments concerning syntactic well-formedness constitute prescriptive judgments about which strings of words should be considered to be well- and ill-formed. Hence, there is a sense in which even a descriptive linguist can be said to be concerned with prescriptive matters. However, a descriptive linguist makes no attempt to pass judgment on the legitimacy of those prescriptive attitudes; his grammar is only supposed to constitute a formal description of those attitudes—not an endorsement or a condemnation of them. By way of contrast, a prescriptive linguist explicitly endorses the attitudes reflected in his grammar and insists that all linguistic behavior should accord with those attitudes; behavior that does not accord with the rules of the grammar is condemned. Thus, whereas a descriptive linguist tends to assume that there is something wrong with his grammar whenever there is a discrepancy between the grammar and the behavior and intuitions of native speakers, a prescriptive linguist tends to assume that there is something wrong with the behavior and intuitions of native speakers. Moreover, a prescriptive linguist's hostile attitude toward deviations from the standards embodied in his grammar often apply, as Lyons observes, not only to current deviations, but also to all future deviations in the
form of linguistic change. Thus, a prescriptive linguist often believes that not only contemporary, but even future native speakers should attempt to mirror his grammar.

The foregoing observations are of particular interest because they so aptly characterize the attitude adopted by most French neo-classical critics. The Abbé D'Aubignac claimed, for example, that the neo-classical rules of drama were based upon the dictates of reason rather than practices and attitudes accepted by members of any particular literary community, adding that his occasional citation of ancient authors was only meant "to show with what agreeable artifice they kept to these rules." Moreover, since they were based on the dictates of reason, D'Aubignac believed the rules to be universally binding, "for reason, being alike the world over, does equally require everybody's submission to it," and he went on to insist that the rules "ought never to suffer an eclipse." Thus, in true prescriptivist fashion, D'Aubignac maintained that his theory need not attempt to mirror the attitudes and behavior of "native

---

110 Lyons, Theoretical, pp. 42-44.
111 François Hédelin, Abbé D'Aubignac, The Whole Art of the Stage, in European Theories of the Drama, anonymous trans., p. 93.
112 D'Aubignac, p. 93.
speakers" of literature; he believed, instead, that those speakers--future as well as present--should endeavor to make their attitudes and behavior mirror his theory. Naturally, such a view led to an intolerant attitude toward all deviations from the standards of neo-classicism. Voltaire's criticisms of Shakespeare, to which reference was made earlier, furnish an excellent example of such intolerance.\textsuperscript{113}

As the foregoing discussion of neo-classical criticism has shown, it provides a paradigm of the way rule-orientation can be used to stifle literary creativity, but rule-orientation need not be used for such purposes. Indeed, far from confirming the widespread assumption that rule-orientation and prescriptivism necessarily accompany each other, neo-classical criticism can be used as an example of what a rule-oriented theory of literature, like generative theory, should--and can--avoid.

Generative theory should, for example, resist the neo-classical tendency to privilege the rules of one particular literary dialect. It should recognize the fact that there are a variety of genuine literary dialects, that different dialects are likely to be based on different rules, and that works written in one dialect should not be condemned for

\textsuperscript{113} See above, p. 313.
failing to accord with the rules of other dialects. Within the language of drama, for example, the rules needed to formulate structural descriptions of comedies might differ from those needed to formulate descriptions of tragedies, and those needed to formulate descriptions of classical tragedies might differ from those needed to formulate descriptions of romantic tragedies. Attempts to formulate the rules relevant to the description of works associated with relatively new dialects should be viewed with special caution. It is often difficult to perceive the underlying patterns and regularities inherent in innovative works, and even when an individual believes that he is beginning to discern them, his results should be viewed with caution.

Just as generative theory should refrain from adopting a prescriptive stance with respect to current dialects, it should also avoid the neo-classical tendency to dictate, to prescribe, the norms of future behavior. In particular, the rules contained in a generative theory should not be viewed as a guide that future readers or writers should feel compelled to follow. They should instead be viewed as an attempt to describe the attitudes and underlying knowledge that come into play when current readers comprehend and evaluate literary works of a given type, and their plausibility should be determined on the basis of how well they account for the intuitions that competent readers actually
have about such works.

One more point should be made with respect to the widespread belief that the rule-oriented nature of a generative theory would automatically bring with it a prescriptive attitude that will inevitably result in attempts to stifle literary creativity. Towards the beginning of Aspects, Chomsky points out that although some pre-twentieth-century writers were well aware that linguistic processes are in some sense creative, they lacked the formal devices needed to represent such creativity—"the technical devices for expressing a system of recursive processes were simply not available until much more recently." Since the rules employed in earlier grammars could only be used to specify a limited number of different syntactic patterns, it is not difficult to understand why linguists like Saussure and Bloomfield viewed them as restrictive and why such linguists came to associate the very idea of rule-orientation with prescriptivism. As earlier remarks have indicated, however, the rules employed in a generative grammar can be used to specify a potentially infinite number of different syntactic patterns. Thus, such rules do not possess the

---

114 Chomsky, Aspects, p. 8.
115 See above, p. 253.
restrictive features of earlier rules, and there is no longer any reason for automatically associating the use of rules with a prescriptive attitude; a grammar that contains such rules can begin to explain—rather than deny—the existence of linguistic creativity.

Likewise, it is not difficult to understand why critics might accuse earlier rule-oriented theories of literature of being restrictive and why the use of rules came to be inextricably associated with prescriptive attempts to stifle literary creativity. Given that the sort of rules that would be contained in a generative theory of literature could be used to generate an infinite number of literary patterns no two of which would be the same, there is, however, no reason for the continuing tendency to think that any rule-oriented theory must constitute a prescriptive attempt to stifle literary creativity. Far from being an attempt to stifle or deny such creativity, generative theory takes literary creativity as a given and attempts to account for its existence.

Although most anti-generative writers have focused on the alleged connection between rule-orientation and prescriptivism, Pettit has advanced an argument against generative assertion 4 that has nothing to do with the issue of prescriptivism. As earlier remarks have indicated, Pettit accepts the claim that literary comprehension presupposes a
capacity to assign structural descriptions to literary works, but he argues that the capacity to assign structural descriptions does not warrant adoption of a generative approach:

just as with a sentence, a narrative text can be given meaning, can be interpreted, only when it is assigned a structure. This conclusion does not entail however . . . that the reader who interprets a text is master of a generative grammar of narrative. This proposition follows only on the further premise that such structure can only be assigned generatively: that is, in the way represented by a generative grammar.\footnote{Pettit, p. 47; see also above, p. 317.}

Pettit chooses, however, to reject that additional premise on the grounds that a proficient individual's ability to assign a structural description to a text is based on implicit knowledge that must be modeled in the manner of Gestalt psychology rather than in generative terms:

\begin{quote}
Structure . . . is assigned to a narrative text as a gestalt perceived there, not on generative grounds. If we put a spread of dots on a blackboard, the eye will almost certainly see a pattern in them— one overall distribution of shapes. . . . This is in accordance with the law of prägnanz discussed in Gestalt psychology: roughly the law that in any set of data the mind sees a good form— one that is simple, continuous and bounded, for instance—and a comprehensive form— one that takes a high account of the data. . . . This law operates in our perception of
\end{quote}
dots on a board but does not dominate that perception in the way a recursive procedure does: it often allows oscillation between two or three 'pregnant' forms and it never fails to present some form. It would seem to be a law of this kind that operates in our perception of structure in a narrative text. We can always opt as readers for one or another organization of the text.

Thus, Pettit claims that, like a spread of dots on a blackboard, a single narrative sequence can often be assigned two or more equally satisfactory, equally 'pregnant', structural descriptions. He takes that fact as evidence that structural descriptions are assigned to literary works on the basis of gestalt criteria like the law of prägnanz, and he assumes that the use of such criteria automatically rules out the possibility of representing literary knowledge in generative terms.

Pettit is no doubt correct in thinking that a single narrative sequence can often be assigned two or more equally satisfactory structural descriptions, but his other claims appear to be incorrect. In their theory of tonal music Jackendoff and Lerdahl develop an entire system of generative rules that, according to them, "constitute an explicit statement of the Law of Prägnanz as it applies to musical

117 Pettit, p. 48.
Moreover, they argue at some length that it should be possible to do likewise in the case of visual perception; and there is no reason to think that the same would not hold true in the case of literature. Thus, Pettit's suggestion that there is some sort of incompatibility between a generative and a gestalt approach appears to be fundamentally misguided.

Even if Pettit were right in thinking that the two approaches are incompatible, however, the mere fact that it is often possible to assign a literary work more than one plausible structural description would not necessarily imply that literary knowledge must be represented in gestalt rather than generative terms. In *Syntactic Structures* Chomsky cited several examples of structural ambiguity or, as he called it, "constructional homonimity" in English. He observed, for example, that the sentence *I found the boy studying in the library* can be assigned two mutually exclusive and equally legitimate interpretations: *I found the boy while he was studying in the library* and *While I was*

---

118 Jackendoff and Lerdahl, p. 304.
119 Jackendoff and Lerdahl, pp. 302-07.
studying in the library, I found the boy.\textsuperscript{121} He went on to observe that the ambiguity of the sentence derives not from the ambiguity of an individual word or phrase, but from an ambiguity in its structural organization; the former interpretation assumes that the phrase \textit{studying in the library} modifies the subject of the main clause, whereas the latter assumes that it modifies the direct object.\textsuperscript{122}

Instead of assuming that the existence of structural ambiguity in language rules out the possibility of assigning structural descriptions on generative grounds, however, Chomsky proved that it is possible to formulate rules capable of automatically generating the various legitimate descriptions that can be assigned to such ambiguous sentences.\textsuperscript{123} Thus, contrary to what Pettit seems to suggest, generative rules also allow perception to "oscillate" between two or more

\textsuperscript{121} Chomsky, \textit{Syntactic Structures}, pp. 88-89.

\textsuperscript{122} Chomsky, \textit{Syntactic Structures}, pp. 88-89. As Lyons observes, the terms "structural ambiguity" and "constructional homonimity" serve to distinguish a sentence whose ambiguity derives from its organizational structure from cases of "lexical" ambiguity or homonimity, cases in which the ambiguity derives from the multiple meanings of one or more of the words contained in a sentence, e.g. \textit{I found the boy at the bank} (Lyons, Chomsky, p. 62).

\textsuperscript{123} Because of the complexities involved in the generation of multiple descriptions for ambiguous sentences, no examples will be furnished here, but see Chomsky, \textit{Syntactic Structures}, pp. 88-89, where a variety of examples are given.
equally plausible structural descriptions, just as the laws of Gestalt psychology do.

In view of the fact that the use of recursive rules in a generative theory would presumably enable theorists to begin explaining—rather than denying or stifling—the creativity involved in the activities of reading, and in view of the fact that gestalt criteria can be incorporated in an account based on such rules, there would appear to be no compelling reasons for rejecting generative assertion 4.

It is worth pointing out, moreover, that the potential value of a generative theory based on such rules would not be exhausted by its usefulness in accounting for the literary knowledge which presumably underlies various literary capacities.

In addition, such rules could serve a descriptive function; they could be used as a concise means of specifying the underlying patterns exhibited by all the works of a particular kind or type written to date. Moreover, they could serve a predictive function. Baker observes that generative linguists view "grammars as yielding predictions about judgments that speakers will make."\(^\text{124}\) Similarly, the rules contained in a generative theory of literature could

\(^{124}\) Baker, p. 5.
presumably be used to make predictions about the judgments a proficient individual will make concerning the immediate constituent structure, plot synopsis, and syntactic well-formedness of literary works.

One additional use is suggested by Chomsky's claim that a grammar based on generative rules can be used to help clarify the status of sentences whose well-formedness is unclear:

we may assume for this discussion that certain sequences of phonemes are definitely sentences, and that certain other sequences are definitely non-sentences. In many intermediate cases we shall be prepared to let the grammar itself decide, when the grammar is set up in the simplest way so that it includes the clear sentences and excludes the clear non-sentences. This is a familiar feature of explication. 125

Lyons illustrates Chomsky's point by way of the following example:

Since we do not know in advance whether The house will have been being built is a grammatical sentence or not, we can formulate the rules of the grammar to include all the definitely acceptable sequences and to exclude all the definitely unacceptable sequences and then see whether these rules exclude or

include such sentences as *The house will have been being built.*

Likewise, the use of generative rules might help to clarify the status of works whose well-formedness with respect to a given dialect is unclear. Given a set of rules that could be used to generate syntactic descriptions for each extant work of that dialect, and which could not be used to generate descriptions for works that clearly do not belong to the dialect, verdicts concerning borderline cases could be obtained by seeing whether the rules can also be used to generate syntactic descriptions for those works.

Of course, such verdicts should not be viewed as final. Consideration of such a verdict might well prompt members of the literary community to overturn it. The rules could then be revised in such a way that the pattern underlying the work in question would be generated by the rules. Indeed, such an example underscores the clarificational value of generative rules—for in such a case it would be the initial verdict obtained by using the rules that prompts members of the literary community to clarify their intuitions about the pattern involved. Moreover, as new works are written, and a

---

126 Lyons, Chomsky, p. 47. Lyons indicates, incidentally, that the system of rules advanced in *Syntactic Structures* does generate the sentence in question (p. 47).
given dialect undergoes changes, the rules contained in a generative theory of that dialect will, of course, also have to be revised.

Thus, generative rules could be useful in at least four different ways; they could be used for descriptive, predictive, clarificational, or explanatory purposes. The degree to which such rules might be useful in accomplishing the former three purposes does not, moreover, depend upon their explanatory usefulness. Even if it were assumed that proficient individuals have no knowledge, implicit or otherwise, of such rules—which would, of course, mean that they would be of no use in trying to explain the underlying sources of literary comprehension and evaluation—the rules might still constitute the most efficient means of specifying the patterns contained in a given corpus of extant literary works, they might still constitute a fairly reliable means of predicting current intuitions concerning matters of grammatical interpretation and syntactic well-formedness, and they might still constitute a valuable means of further clarifying those intuitions. Thus, a generative theory of literature could be broadly useful, even if it failed to provide an adequate account of the psychological sources that underlie various literary capacities.
There is, however, one major obstacle that must be overcome if such a theory is to be developed. As earlier remarks have indicated, development of a generative theory of language presupposes that the hierarchical levels of linguistic structure have already been identified. It seems reasonable to expect that much the same will hold true in the domain of literature, but, as remarks in Chapter Three indicated, Structuralists have had little luck in identifying hierarchical levels of literary structure. The remaining two chapters of the present work will be devoted to the task of showing that Aristotle's Poetics provides a firm foundation for the description of the hierarchical structure of drama and thus provides a means of overcoming what is perhaps the major obstacle standing in the way of developing a generative approach to drama.

127 See above, p. 254.

128 See above, pp. 227-32.
THE STRUCTURALIST ENTERPRISE AND ARISTOTLE'S POETICS
VOLUME II

DISSERTATION

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

By
Harry Mitchell Flower III, B.A., M.F.A.

*****

The Ohio State University
1986

Reading Committee:
Donald R. Glancy
Alan Hausman
Alan Woods

Approved By

Donald R. Glancy
Adviser
Department of Theatre
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>V. THE QUALITATIVE PARTS OF TRAGEDY ACCORDING TO ARISTOTLE</td>
<td>347</td>
</tr>
<tr>
<td>The Qualitative Parts as a Group</td>
<td>357</td>
</tr>
<tr>
<td>Spectacle</td>
<td>387</td>
</tr>
<tr>
<td>Melody and Diction</td>
<td>404</td>
</tr>
<tr>
<td>Thought</td>
<td>440</td>
</tr>
<tr>
<td>Character</td>
<td>460</td>
</tr>
<tr>
<td>Plot</td>
<td>465</td>
</tr>
<tr>
<td>Summary</td>
<td>476</td>
</tr>
</tbody>
</table>

| VI. ARISTOTLE'S CONCEPTION OF CAUSALITY AND THE QUALITATIVE PARTS OF TRAGEDY | 479 |
| Aristotle's Theory of Causality | 483 |
| Aristotle's Theory of Imitation | 495 |
| The Causal Interrelations Between the Qualitative Parts | 511 |

| VII. CONCLUSION | 543 |
| APPENDIX A | 546 |
| APPENDIX B | 574 |
| BIBLIOGRAPHY | 590 |
Chapter V

The Qualitative Parts of Tragedy According to Aristotle

The closing remarks of the previous chapter suggested that a careful examination of Aristotle's Poetics might be of considerable aid in the development of a Structuralist theory of drama. Such an examination may profitably begin with a brief consideration of what previous Structuralists have said about the matter.

Structuralist authors have tended to exhibit a rather curious attitude towards the Poetics. They typically acknowledge the work in praiseworthy terms and then proceed to say little or nothing more about it. Scholes, for example, observes that "Attempts to describe basic narrative structures are as old as Aristotle, and it is frequently said that little progress has been made since his time."¹ During his own survey of previous attempts to describe narrative structures, however, Scholes says virtually nothing about Aristotle's work.² Elsewhere, during a

¹ Scholes, pp. 92-3.
² Scholes, pp. 93-117.
discussion of Russian formalism, Scholes remarks that "the achievement of the Russian formalists, like that of Aristotle, will be permanent because it will have to be incorporated in any later poetics of fiction."\(^3\) That remark follows an extended discussion of how the work of the Russian formalists has been and could be incorporated into the Structuralist framework, but again, there is no discussion of Aristotle's work.\(^4\)

Scholes' remarks bespeak a rather puzzling attitude. If Aristotle has in fact contributed more to the study of narrative structure than any other single author, a thorough examination of his work from a Structuralist perspective would seem to be of considerable value to the Structuralist enterprise. Moreover, if that enterprise is to incorporate Aristotle's contributions, as Scholes suggests it must, such an examination would appear to be indispensible. Nevertheless, there has to date been only one extended treatment of Aristotle's *Poetics* from a Structuralist perspective, a discussion contained in Richard Hornby's *Script into Performance: a Structuralist View of Play Production*.\(^5\)

\(^3\) Scholes, p. 158.

\(^4\) Scholes, pp. 74-91.

Since Hornby's book contains the only extended treatment of the matter to date, the following discussion will begin with a critical consideration of his remarks.

Several chapters after his introductory chapter, which deals with what Hornby takes to be the primary characteristics of Structuralism, Hornby devotes a chapter entitled "What Aristotle Might Have Said" to a discussion of the relevance of Aristotle's work to the Structuralist enterprise. Hornby spends the first portion of his discussion identifying and praising what he sees as Structuralist tendencies in the Poetics. For example, he applauds Aristotle for adopting a descriptive rather than prescriptive approach to genre, emphasizing the differences between Aristotle's approach and the approach adopted by the French neo-classical critics. He also cites Aristotle's remarks concerning artistic unity with approval, remarking that for Aristotle a good work of art contains both unity and complexity. This apparent contradiction is resolved by the idea of the one whole action, or "generating principle," underlying the work.

---

6 Hornby's chapter on Structuralism covers pages 10-39; the chapter on Aristotle covers pages 68-91.

7 Hornby, pp. 69-71--see also Hornby's remarks concerning the distinction between descriptive and prescriptive criticism on pages 22-23.
There is complexity on the surface but unity beneath it. This is clearly a Structuralist idea . . . .

Finally, he praises Aristotle's definition of plot as "the arrangement of the incidents," observing that the definition is Structuralist "in stressing a pattern made by the parts rather than the parts individually." 9

Having identified the more Structuralist aspects of the Poetics, Hornby proceeds to identify what he believes to be the most harmful and least Structuralist aspects of the work:

Here then is the problem with Aristotle in the Poetics--his tendency to analyze, to break into parts, is not balanced by a tendency to synthesize, to see the parts in relation to the whole again. He does not follow through on a key sentence in his definition of unity: "a whole, the structural union of the parts being such that, if any one of them is displaced or removed, the whole will be disjointed and disturbed". . . . Instead of exploring how this "structural union" comes about, so that, for example, spectacle can become a valid and necessary part of Tragedy, Aristotle tends to rank elements, which only underlines their separation. He separates text from performance. He separates Tragedy into six parts--the three internal ones, namely plot, character, and thought, and the three external ones, namely diction, song, and spectacle. He separates plot into further parts, such as reversal, recognition, and denouement. 10

---

8 Hornby, p. 72.
9 Hornby, p. 79.
10 Hornby, pp. 74-5.
The preceding passage is followed by an extended discussion in which Hornby elaborates on what he believes to be the excessively atomistic character of Aristotle's division of tragedy into six parts—what are sometimes called the "qualitative" parts of tragedy.\textsuperscript{11}

According to Hornby, "Aristotle sees [spectacle] as something that is merely \textbf{added} to the text."\textsuperscript{12} Moreover, Aristotle failed to realize that character and thought are "always a matter of interrelationship with a larger whole."\textsuperscript{13} Worst of all, according to Hornby, Aristotle's attempt to divide plot into discrete parts, and his attempt to classify one of those parts, namely discovery, into five types, which fall prey to his perpetual "penchant for ranking," virtually obscure his insightful definition of plot as "the arrangement of the incidents."\textsuperscript{14} Having thus discussed what he believes to be the insurmountable shortcomings of

\textsuperscript{11} The term "qualitative" is, for example, used by R. S. Crane in his essay, "The Concept of Plot and the Plot of Tom Jones," in \textit{Critics and Criticism}, ed. R. S. Crane (Chicago: University of Chicago Press, 1952), p. 617. The word is used primarily to distinguish the nature of plot, character, thought, diction, melody, and spectacle as parts of tragedy from the nature of the so-called "quantitative" parts, e.g., the episodes. More on the subject will be said in a subsequent discussion—see below, pp. 385, 418-19.

\textsuperscript{12} Hornby, p. 75.

\textsuperscript{13} Hornby, p. 78.

\textsuperscript{14} Hornby, p. 79.
the six elements identified by Aristotle, Hornby dismisses them, offering six new ones of his own making in their stead.\footnote{Hornby, p. 80. Since Aristotle's work constitutes the major focus of the present chapter, Hornby's alternative six elements will not be discussed—see Hornby's own discussion of them on pp. 80–90.}

Before other possible interpretations of Aristotle's six elements are discussed, the rather widespread popularity of misreadings similar to Hornby's ought to be mentioned. Far from representing an unusual view, Hornby's interpretation appears to belong to the mainstream. For example, Orsini asserts that Aristotle

found organicity only in one of the parts into which he analyzed tragedy, i.e., only to the plot or mythos, and did not apply organicity to the unification of all the parts into a single whole, but limited himself to making an enumeration of them, i.e., mere taxonomy.\footnote{Orsini, pp. 37–38.}

Moreover, in a later passage, Orsini suggests that the six parts "exhibit no order or systematic unity, but are in the end merely arranged," and he goes on to conclude that "Goethe put it in a nutshell: Aristotle's enumeration of the parts of tragedy is like saying that a man is made up of soul, body, hair, and clothes. Mere taxonomy, and not very
good at that."\(^{17}\)

As Orsini's remarks indicate, the view that the qualitative parts of tragedy represent a disorderly collection of almost unrelated fragments dates back at least to the time of Goethe, and it appears to have held sway, with perhaps a few possible exceptions, ever since that time. Fortunately, however, it is not the only view that has been advanced. The "Chicago School Critics," so called because they were all associated with the University of Chicago at one time or another, tended to advocate a very different view of Aristotle's six elements, a view admirably described by R. S. Crane in a passage worth quoting at length:

For Aristotle, concerned with the construction of poetic wholes that afford "peculiar pleasures" through their imitations of different species of human actions, the four terms [i.e., the first four of the six qualitative elements] had designated the essential elements upon the proper handling and combination of which, relatively to the intended over-all effect, the quality of a tragedy or epic necessarily depends. They are distinct parts in the sense of being variable factors in the complex problem of composing works which, when completed, will produce their effects, synthetically, as organic wholes. Hence it is that in the Poetics they are treated, not discretely as co-ordinate topics, but hierarchically in a causal sequence of form-matter or ends-means relationships in which plot is the most inclusive or architectonic of the four, subsuming all the others as its poetic matter; in which character, while subordinate

\(^{17}\) Orsini, pp. 77-8.
materially to plot and effect, is similarly a formal or organizing principle with respect to thought and diction; in which thought, while functioning as matter relatively to character, incident, and effect, is the form which immediately controls the choice and arrangement of language in so far as this is employed as a means to imitative rather than ornamental ends; and in which diction, though necessarily having a form of its own by virtue of its rhythmical, syntactical, and "stylistic" figuration, is the underlying matter which, as significant speech, at once makes possible all the other "parts" and is in turn, mediately or immediately controlled by them. 18

Several decades later, Kenneth A. Telford, who is also associated with Chicago School Criticism by way of Richard McKeon, made some quite similar remarks:

Diction, thought, character, and plot . . . constitute an ascending hierarchy of more inclusive forms in which the lower syntheses are means to the attainment of the higher and in which the higher are the organizations or unities of the lower. In this hierarchy plot is the ultimate organization or form, while diction is the ultimate matter, or that which is ultimately organized. 19

Indeed, Telford's remarks plainly constitute a reaffirmation of the view espoused by Crane, a view that will hereafter be

18 Crane, pp. 617-18. Among the members of the original "Chicago School" were all of the authors who contributed essays to Critics and Criticism, namely Crane, W. R. Keast, Norman Maclean, Richard McKeon, Elder Olson, and Bernard Weinberg.

referred to as the "Chicago School" view.

As the preceding passages indicate, the Chicago School view is quite different from the one espoused by Hornby and Orsini—what might be called the "atomistic" interpretation; where the latter posits a haphazard collection of minimally related fragments, the former posits a systematically ordered hierarchy of synthetically integrated elements. It is also obvious that the Chicago School view provides a much more Structuralist reading of the Poetics. Indeed, Crane's references to an "architectonic" hierarchy wherein lower levels are "subsumed" by the higher ones suggest a possibility that possesses considerable importance for the Structuralist enterprise—the possibility that Aristotle may have succeeded in identifying a series of hierarchically organized levels of dramatic structure analogous to the hierarchically organized levels of linguistic structure. The remaining two chapters of this study constitute a detailed examination of that possibility.

In order to get a clear idea of how the qualitative parts of tragedy are related to each other, it will, of course, first be necessary to ascertain the nature of the parts themselves. The remainder of the present chapter is therefore be devoted to a discussion of the qualitative parts themselves; the relations between the parts are discussed in Chapter Six.
As each of the parts is discussed during the remainder of this chapter, a good deal of attention will be paid to what previous writers have said about them. For two reasons, special attention will, moreover, be paid to what Gerald F. Else has said in his book *Aristotle's Poetics: the Argument*. First, Else's voluminous 635 page work constitutes the most extensive, detailed English commentary on the *Poetics* to date, and it is widely regarded as definitive. Second, if many of Else's key contentions about the nature of the qualitative parts are correct, the parts cannot possibly be related to each other in a hierarchical manner. Therefore, although consideration of Else's arguments will entail a number of rather technical discussions, it is imperative that they be considered.

Although the primary reason for including discussions of Else's views in this chapter is that they must be confronted by any heavily Structuralist interpretation of the *Poetics*, it is worth pointing out that the discussions also constitute an important contribution to the scholarly study of the *Poetics*—aside from their relation to Structuralist issues. In spite of the fact that later commentaries on the *Poetics* often contain views that conflict with those

---

advanced by Else, they virtually never contain detailed ar-
guments designed to show what is wrong with his views. In
fact, such commentators rarely even point out the fact that
their views conflict with his.21 As a result, the real
plausibility of the various rival views, including those ad-
vanced by Else, remain largely undetermined. Since it con-
stitutes the first real attempt to assess the plausibility
of Else's views in detail, the following discussion will
thus be useful even aside from its connection with Structur-
alist concerns.

THE QUALITATIVE PARTS AS A GROUP

Before discussions of the qualitative parts on an
individual basis begin, the overall character of the parts
as a group must be discussed. In his 1957 commentary, Else
opposed what had by then become the traditional view of the
qualitative parts as a group, insisting that they "are not
parts of the poem, or of the collective entity called

21 For example, although Telford expresses views that
frequently conflict with those espoused by Else, he never
even refers to Else's commentary, so naturally his commen-
tary contains no attempt to show that his interpretation is,
in fact, superior to Else's--see the subsequent remarks on
pages 365, 390.
'tragedy' (i.e., the aggregate of all poems so called), but specifically of the art or composition of tragedy.22

During a subsequent discussion, he offers a somewhat more detailed summary of his interpretation, remarking that the six parts were thought of by Aristotle as constituent elements of the art of tragedy, or what comes to the same thing, moments in the activity of building a tragedy. Thus, they really mean respectively:

1. The visualization of characters.
2. The composition of verses (dialogue).
3. The composition of songs.
4. The structuring of the action.
5. The expression of character mainly through speeches.
6. The expression of ideas through speeches.23

In other words, the six terms used by Aristotle refer, according to Else, to six separate activities that are together involved in the making of a tragedy.

Four years later, Telford reasserted the traditional view, maintaining that the six terms "refer to formal parts of the work itself."24 Thus, the first question to be


23 Else, pp. 279-80.

24 Telford, p. 82.
discussed is that of whether the six parts of tragedy are parts of "tragedy" as a finished product, or parts of "tragedy" as a process which results in that product, or whether they are, perhaps, parts of "tragedy" in both of the foregoing senses. The answer will, of course, possess substantial implications as regards the usefulness of the Poetics to Structuralism; while the levels of linguistic structure discussed earlier are no doubt correlated with various activities that are involved in the production and comprehension of sentences, they also constitute structural features of the linguistic product—of the sentences themselves. Thus, if Else is correct in thinking that the qualitative parts of tragedy do not constitute structural features of the dramatic product, the analogy with linguistics will break down at the outset.25

25 The answer to the preceding question is not, moreover, unrelated to the question of whether the Chicago School interpretation represents a plausible reconstruction of Aristotle's original conceptualization of the qualitative parts. As subsequent discussion will show, the Chicago School interpretation maintains that when the parts are placed in a certain order, the lower ones constitute the material causes of the those above them (see below, pp. 521-23). If, however, the parts represent processes, it would be difficult to see how they could be so related; they could perhaps be related to one another as efficient or final causes, but not as material causes (see below, pp. 484-85). Unfortunately, Else never discusses the relation between his views and those of the Chicago School critics— even though he mentions Critics and Criticism in a footnote, which indicates that he must have been familiar with their views (Else, p. 4 fn # 12).
An exhaustive list of the various pieces of evidence adduced by Else in support of his position would be difficult to assemble because the sundry strands of his defense are scattered throughout his extensive commentary. Nevertheless, a rough list of the primary pieces of evidence upon which he bases his case can be provided. First of all, he attempts, during the opening pages of his commentary, to show that the Greek word for tragedy, "tragodia" [τραγῳδία], should almost always be translated as "the tragic art," in the sense of a compositional process, rather than as "tragedy" in the sense of a concrete object or collection of such objects. He does, in passing, note that the Greek word is "potentially ambiguous"—that it could mean "a tragedy," i.e., a single play, and probably also 'tragedy' in the 

---

26 Else, p. 6. Throughout the remainder of the present chapter, the initial reference to a key term in the Poetics will routinely include citation of not only the English translation of the Greek term, but will also include citation of the original Greek term itself, enclosed in brackets, and citation of the English transliteration of the original Greek term, sometimes, though not always, enclosed in parentheses. Subsequent references to the term will vary, sometimes including only the English translation, sometimes including only the English transliteration of the original Greek term, sometimes including both of the foregoing, and sometimes including the original Greek term along with both of the foregoing. The original Greek version of a term will, however, never appear by itself; it will always be accompanied by both the English transliteration of the term and the English translation.
sense in which we customarily speak of 'Greek tragedy'.

However, Else focuses attention on the fact that the first occurrence of the word in the text of the Poetics is embedded in a construction that clearly refers to a process.

Else then attempts to utilize that fact to motivate a general rule for translating the term:

Naturally he [i.e., Aristotle] does not drag these cumbrous phrases into the further discussion; but whenever he writes τραγῳδία [tragodia] in the Poetics, unless the context clearly demands another meaning, we are to read it as 'the tragic art'.

Several pages later, Else presents his second major piece of evidence. He draws attention to the opening remarks of the Poetics where Aristotle provides a brief list of major

27 Else, p. 6.

28 Poetics, 1: 1447a14. Whenever a remark in the main text refers to a passage in the Poetics but does not involve a direct quotation from a particular English translation—and whenever the text contains a quotation from the original Greek, editions of which can be found in both Butcher and Else—the adjacent footnote will merely cite the relevant line numbers. All remarks involving direct quotation from an English translation will, however, be accompanied by a footnote indicating the particular translation used. Because it represents the most literal of the major English translations, the Telford text will, aside from exceptional cases, be used throughout.

29 Else, p. 6.
topics he intends to cover, noting that the list specifically mentions the parts of the "poetic process" (poiesis, ποιήσις), but makes no mention of the parts of the poem or product, a fact which Else says "will have great concrete importance when we come to Aristotle's derivation of the 'parts' of tragedy in Chapter 6."^30

Else's last three arguments all concern material contained in chapter six of the Poetics. The third argument is based on his contention that the introduction of several of the parts in chapter six involve the use of "action-words." For example, the term usually translated as "melody" or "music,"" melopoiia" [μελοποιία], carries a sense of activity—of "making." Moreover, when explaining what he means by the term usually translated as "diction," "lexis" [λέξις], Aristotle describes it as "ton metron synthesin" [τὸν μέτρον σύνθεσιν], which literally means "the synthesis of the meters."^31 Likewise, Aristotle's explanation of what translators usually call "plot," mythos [μῦθος], refers to it as

^30 Else, p. 9 and the Poetics 1: 1447a8-13.

^31 Poetics, 6: 1449b34-35. Else translates the phrase in question as "the composition of the verse," but since the plausibility of Else's interpretation is partially dependent on the question of whether his looser translation manages to capture Aristotle's meaning, the main text will utilize Telford's more literal translation.
"ton synthesin ton pragmaton" [τὴν σύνθεσιν τῶν πραγμάτων], which literally means the "synthesis of the doings" or "incidents." According to Else, the introduction of the preceding three parts in such action-oriented terms indicates that they must refer to "parts' or moments of the poetic art, the making of the poem, not parts of the poem itself."  

Fourth, Else claims that Aristotle's first serial enumeration of the six parts in lines 1450a8-10 must be assumed to assert—not that every tragedy has six parts, as some translations like Butcher's do—but that "the constituent elements of the tragic art as a whole are six in number." Else offers the following remark in support of his claim:

Aristotle proclaims a few lines below (a25) that "tragedy can exist without character(s)," and in a30 the possibility of a plotless play seems to be envisioned. Thus it is not true that every tragedy has all six parts, or at least has them developed beyond what the meteorologists would call a 'trace.' What does have all the parts is the concept or art of tragedy, "tragedy as a whole."  

Else then argues, on similar grounds, that the immediately

32 Poetics, 6: 1450a4-5.
33 Else, p. 237.
34 Else, pp. 238, 245, and Butcher, p. 25.
35 Else, pp. 245-46.
following clause in the text of the *Poetics* should also be
assumed to refer to tragedy as "a special kind of art." 36

Else's fifth, and last, major line of defense is also
the one he emphasizes most heavily. He maintains that the
passage in which Aristotle first introduces the six parts
constitutes a "systematic deduction" of those parts, a de-
duction which he summarizes as follows:

[Aristotle's] procedure has been in effect to
say: we have before us a genre which has just
been defined as an imitation of an action, etc.,
etc. Now what will be its necessary compon-
ents—that is, the necessary components of the
total art of writing tragedy (πάρησις τῆς
τραγωδίας)? Everything hinges on the fact that
the characters perform the action themselves,
through action. First, they will necessarily
look one way rather than another. Second, they
must speak part of the time and sing part of the
time; for those are the means they have of
performing the action. Next, the action they
perform will have a certain structure (σχέδος)
which we call the plot or fable ('myth'), and
since it is to end in happiness or unhappiness
for them (=στοιχεία), we shall need to know the
moral and intellectual springs that lie behind
it; these are the substantive things they [i.e.,
the dramatis personae] place before us. These,
then, are the six 'parts' of the art of trage-
dy. 37

The preceding quotation does not make explicit the connection

36 Else, pp. 238, 246.
37 Else, p. 247.
between the things the **dramatis personae** must do and the things the poet must do, but it is fairly easy to see what Else has in mind. If the **dramatis personae** must look a certain way, speak, sing, exhibit certain moral and intellectual qualities, and participate in a structured action, then the poet must engage in a group of corresponding activities that will enable them to do so—he must create their appearances, compose dialogue and songs for them, create certain moral and intellectual attributes with which to endow them, and he must create a structured action in which they can participate. Thus, the foregoing group of activities constitute, according to Else, the true referents designated by Aristotle's six terms.

In spite of the fact that Telford's commentary was published several years after the appearance of Else's, it contains no mention of, much less a response to, the arguments contained in the earlier work. Moreover, Telford's

---

38 Other portions of Else's discussion make it clear that the foregoing is in fact what he has in mind—see especially pp. 279–80.

39 In fact, Telford's commentary virtually never refers to any other interpretations of the *Poetics*; it is simply intended to be an explication of the text, not a defense of that reading against rival interpretations. Thus, the comparative merits of the various rival readings remains to be assessed.
commentary does not include any remarks specifically intended to support his view that Aristotle's terms must refer to parts of the play itself. Nevertheless, the commentary does include some remarks that suggest a plausible line of defense. During his discussion of Aristotle's general approach to the problems of productive art, Telford includes the following observations:

By Aristotle's approach the problems of making are stated and solved in terms of the requirements of the thing that arises from this process, the product or work of art. For the process of making and the product made are essentially the same in the sense that they have the same formula or definition, since the product is simply the end of the process of making, its fulfillment or completion, and the nature of anything is what it is when its genesis is completed. The process, however, cannot be investigated as such. Its nature can only be determined by examination of its perceptible result, the product.  

Thus, the productive process can be investigated only insofar as it leaves its traces, the traces of its definition, in the product made and in that product's definition, which suggests that the parts of the process can, likewise, be identified only insofar as they can be inferred from the parts of the finished product. Identification of the parts of the product made, the play itself, must, therefore,

---

40 Telford, pp. 61-2.
constitute the first and foremost step in ascertaining the parts of the poetic process. The preceding argument clearly does not claim that Aristotle's six terms never refer to parts of the productive process; what it does claim, unlike Else's interpretation, is that the six terms also refer—and refer primarily—to parts of the play itself.

Since neither author's commentary contains a response to the arguments advanced by the other, they will be of little help in the task of attempting to assess the relative plausibility of their respective positions. Perhaps discussion should begin with an assessment of whether any, all, or, perhaps only some of the arguments advanced by Else are successful. Else's first argument rests on two claims: 1. the context makes it clear that the first occurrence in the Poetics of the Greek word for tragedy must refer to a process or activity, and 2. that fact is sufficient to motivate a general policy whereby successive occurrences of the term are, aside from exceptional cases, translated similarly. The first of the two preceding claims is unquestionably true; the initial occurrence of the term is embedded in a whole list of what are clearly all types of artistic

41 See above, pp. 360-61.
processes. The second claim is, however, questionable. In view of the fact that Else himself admits the term could be used in a variety of ways, and in view of the fact that the term occurs in a number of contexts within the Poetics itself that clearly refer to tragedy as a product (in fact, those occurrences appear to outnumber the ones where the term clearly designates a process), such a general rule cannot be accepted without a good deal of additional motivation.

Else's second argument is based on the supposition that since Aristotle explicitly says in his opening remarks that he is going to discuss the parts of the poetic process, but says nothing about discussing the parts of the poetic product, all subsequent references to the "parts" of tragedy must refer to the former and not the latter. However, the preceding argument once again involves questionable assumptions. Aristotle's opening remarks represent a preview of only some of the major topics he intends to cover, as is made quite evident by the fact that the brief list ends with his express intention to speak of the foregoing and "of any

42 Poetics, 1: 1447a14-17.

43 Passages that contain occurrences of the word that clearly refer to tragedy as a product rather than a process can, for example, be found at 1451a8, 1451b20, 1453a20, etc.

44 See above, pp. 361-62.
other matters in the same line of inquiry"—which obviously covers a good deal of territory. Moreover, the closing statement of the Poetics contains a second list, a list of major topics Aristotle has by that time finished discussing, and that list contains an item, namely reference to an earlier discussion of "the Objections of the critics, and the Solutions in answer to them," that the opening list omits. Thus, while the earlier list suggests that the Poetics probably does contain a discussion of the parts of the poetic process, that list cannot be used to prove that Aristotle does not discuss the parts of the poetic product.

Else's third argument is based on his contention that several of the parts are introduced in action-oriented terms. His point is undeniably correct in the case of "melopoia," i.e., music or melody; it does carry a sense of activity—of making. However, Else's remarks about "lexis" (diction) and "mythos" (plot) are open to question. Just like its literal English counterpart, "synthesis," the Greek term "synthesin" can refer either to a process in which various things are brought together or to the combination—the union—of those things

45 Bywater translation, 1: 1447a11.
47 See above, pp. 362–63.
qua product. Thus, the "synthesis of the meters" might refer to either the process of combining the meters or to the resultant combination. The same would, of course, hold true for plot as the "synthesin of the incidents"—were it not for the fact that Aristotle elsewhere associates plot with the phrase "systasin pragmaton" [σύστασιν πράγματων], a phrase which Else himself renders as "a structure of events"—not the structuring of the incidents or events, but the resultant structure. Thus, contrary to what Else says, Aristotle sometimes speaks of plot in manifestly non-process-oriented terms.

The same holds true, moreover, for several of the remaining parts. For example, the word usually translated as "spectacle," "opsis" [ὤψις], literally means "appearance" or "that which is seen"—not the making of "that which is seen" but the appearance itself. Furthermore, the word usually translated as "character," "ethos" [ἦθος], means something like "moral disposition"—again, not the making of a moral disposition but the disposition itself. Thus, while

48 Even Else's looser translation of the term as "composition" exhibits the same sort of ambiguity; the "composition of the meters" might refer to either the composing process or to the resultant composition.

49 Poetics, 6: 1450a34.

50 In order to avoid confusion over the intended meaning of the term "character," the present study will refrain from using the term as a means of referring to dramatis
Aristotle's introduction of one of the parts, melopoiia (melody), is definitely action-oriented, two others, opsis (spectacle) and ethos (character), are manifestly non-action-oriented, another, lexis (diction), is spoken of in ambiguous terms, and yet another, mythos (plot), is spoken of sometimes in ambiguous terms and sometimes in clearly non-action-oriented terms.

Far from supporting Else's case, the foregoing terminological considerations would appear to support the idea that the terms are equally applicable to both the parts of the process and the parts of the product. Interpreters who insist that all the terms refer exclusively to parts of the product or exclusively to parts of the process inevitably face the difficult task of explaining away passages that clearly fly in the face of their interpretation, a task that is usually accomplished by either ignoring the obvious implications of the bothersome passages or by twisting and torturing those passages into saying things they were never meant to say. By way of contrast, the present interpretation allows Aristotle to utilize the six terms

personae, e.g., Oedipus is a character in two of Sophocles' finest plays. The term will be used whenever the Aristotelian notion of ethos is at stake, and it will also be used in contexts where its intended meaning is roughly synonymous with that of "nature," e.g., the character of tragedy as a dramatic genre. When references to dramatis personae are made, terms like "dramatis personae," "dramatic personage," "agent," etc., will be used.
with the sort of freedom and flexibility he apparently wished them to have.

Else's fourth argument claims that the first serial enumeration of the six parts must refer to parts of the tragic art if the remark is to be reconciled with Aristotle's later reference to characterless and plotless plays. However, that argument, once again, suffers from several weaknesses. First of all, it is not entirely clear that Aristotle's references to characterless tragedies are to be taken at face value. Butcher, for example, argues that such references cannot be pressed in a perfectly literal sense. The meaning intended probably is, that there may be a tragedy in which the moral character of the individual agents is so weakly portrayed as to be of no account in the evolution of the action.

Moreover, Butcher's suggestion would appear to receive further support from Aristotle's apparent reference to plotless tragedies several lines later:

one would produce this [i.e., the proper function of a tragedy] much more with a tragedy which, though wanting in these respects [i.e., wanting in terms of diction, thought, and character], has a plot or a construction of incidents.

---

51 See above, pp. 363-64.

52 Butcher, pp. 344-45.

53 Telford translation, 6: 1450a30-4.
The passage clearly involves a comparison of two sorts of tragedies; one of them, the more successful sort, has a plot—and the other sort? Well, the other less successful sort is apparently supposed to lack a plot. However, as the very remark in question goes on to say, plot is the systasis [σύστασις]—the construction, or, as other translations would have it, the "combination" or "structure"—of the incidents. Since it is clearly impossible for the incidents in a drama not to be combined or structured in some way or other, the statement in question would appear to make no sense at all—unless the two sorts of tragedy being compared by Aristotle differ not in respect of the presence or total absence of a plot, but rather as regards the nature of their respective plots. In other words, while the incidents in both sorts of plays have been combined or structured in some way or other, only one of the two sorts exhibits anything like an appropriate kind of structure—the other sort "might as well" have no plot at all. Given the apparent parallel between the preceding passage and the one in which Aristotle refers to "characterless" tragedies, it seems possible that both passages are meant to convey a similar meaning; even though Aristotle baldly says that some tragedies are characterless, he may, perhaps, have meant

54 Bywater uses "combination"; Else uses "structure."
that their portrayal of character is so badly handled or so faint as to amount to practically nothing at all.

Even if Else's literalist interpretation of Aristotle's references to characterless and plotless tragedies were the only one possible, his argument would, however, still not succeed in proving that the lines in question must refer to tragedy as a process rather than a product. As Else himself indicates, the second clause in question, that which occurs in line 1450a9, could be translated "tragedy, insofar as it is a distinct species (e.g., distinct from the epic), has six parts," or "tragedy, qua tragedy, has six parts," and there is no apparent reason why the immediately preceding clause in the text of the Poetics could not be translated similarly.\(^55\) However, the contention that the two clauses both contain references to tragedy as a species certainly need not carry the implication that the species in question is a kind of activity; it could just as well be a species of artistic product. Moreover, the latter reading would in no way conflict with the possible existence of literally characterless and plotless tragedies, assuming for the moment that such entities do, in fact, exist; just as deformed members of an animal species may lack parts essential to the species as a kind, individual tragedies might lack parts

\(^{55}\) Else, p. 246.
essential to tragedy qua kind of artistic product.

Finally, Else's argument based on Aristotle's deduction of the six parts must be considered. An assessment of that argument requires the evaluation of two separate claims. According to Else: 1. Aristotle attempts to prove that tragedy must necessarily have the six parts in question specifically because its action is carried out directly by the dramatis personae, and 2. that fact shows that the six parts must be parts of tragedy qua process rather than product.56

The first of Else's claims is certainly true. The very first sentence of the lengthy passage in which Aristotle introduces the six parts testifies to that:

Since agents [rather than the poet] produce the imitation, first, of necessity, [1] one part of tragedy will be the ornament of spectacle and second, [2] the making of melody and [3] diction.57

Thus, Aristotle says that the first three parts belong to tragedy "of necessity," and Aristotle even uses his customary word for necessity, "ananke" [ἀνάγκη]. The philosophical lexicon contained in the Metaphysics includes an entry that distinguishes between various senses of necessity, among which is hypothetical or conditional necessity--

56 See above, pp. 364-65.

57 Telford translation, 6: 1449b31-4. The bracketed material appears in the Telford text.
basically the necessity involved in utilizing certain means if certain ends are to be attained.  That would appear to be precisely the sort of necessity referred to in the foregoing passage; if the representation of an action is to be carried out by the *dramatis personae* themselves, then that representation must necessarily involve the three parts mentioned. Several lines later, Aristotle remarks that the *dramatis personae* "of necessity are of a certain kind by virtue of character and thought."  Once again, if the representation of an action is to be carried out by the *dramatis personae*, then it must necessarily contain two additional parts, namely character and thought. Last, tragedy must contain an embodiment of the action itself, and as Aristotle says, that is plot. Therefore, concludes Aristotle, "of necessity every tragedy has six parts."

Before the plausibility of Else's second claim is discussed, several additional points related to the preceding remarks should be made, points that will assume a fair amount


59 Telford translation, 6: 1449b34–5.

60 *Poetics*, 6: 1450a4.

61 Telford translation, 6: 1450a8–9.
of importance in later discussions. First, Aristotle not only claims that tragedy necessarily has all six parts, he goes on to insist that it has no additional parts:

\[
\text{two of these parts are the materials in which the agents imitate, one is how they imitate, three are the things that they imitate, and besides these there are no others.}^{62}
\]

As Else observes, the classification of the six parts according to means, object, and manner of imitation indicates the reason why they constitute an exhaustive list:

\[
\text{there cannot be any more; for (1) the characters [i.e., the dramatis personae] must present the action to us through the media of sight and sound: there are no others that could be used for this purpose; and (2) they must present to us what they do and why they do it: there is nothing else they could present.}^{63}
\]

However, an important feature of Aristotle's position has apparently been overlooked by all previous commentators, including Else: Aristotle's position allows for the possibility that each of the six parts may possess aspects that are not strictly necessary features of tragedy, features


63 Else, pp. 247-48. Unfortunately, quotations, like the preceding one, will sometimes utilize the term "character" when referring to the agents or dramatis personae. Such occurrences of the term in quotations will be routinely followed by an indication, enclosed in brackets, that such is the case.
that are exhibited by some—but not all—tragedies. For example, a case might be made to the effect that the visual appearance of the agents constitute the only aspect of spectacle that is strictly necessary to tragedy qua action carried out by the *dramatis personae*; some tragedies can, for instance, be done without supplying the audience with any visual representation of the play's environment, but it would be impossible to perform a tragedy without supplying the audience with visual representations of the *dramatis personae*—since they must perform the action themselves, they, at least, must be seen. Since Aristotle claims that a tragedy could possess no parts beyond the six he enumerates, the parts must, however, encompass not only those aspects that are strictly necessary in the sense that no tragedy could be without them, they must also encompass aspects which only some tragedies possess. Since at least some tragedies possess scenic elements, those elements must thus be encompassed by one of the six parts, viz., spectacle. Moreover, spectacle might not be the only one of the six parts to include aspects over and beyond those strictly necessary to tragedy qua action carried out directly by the

---

64 See the subsequent discussion of spectacle on pages 391, 393-94 for more on the preceding subject.
agents; others might, as well.65

Although the text of the Poetics contains sufficient support for Else's claim that Aristotle derives the six parts of tragedy from the idea of tragedy as the representation of an action carried out directly by the dramatis personae, that fact does not, as Else supposes, prove that the six terms refer to parts of tragedy qua process rather than product. In fact, Else's own summary of Aristotle's derivation, which was quoted in full earlier, refers not to the poet's activities, but to the various things the dramatis personae must do—the "substantive things they [must] place before us."66 Now, the dramatis personae are obviously in the play itself, and the substantive things they place before us—their appearances, their sounds, their words, etc.—must obviously be contained in the play as well. Thus, far from supporting his own interpretation, Else's summary of the derivation would actually appear to corroborate the Telford reading; in order to identify the various things the poet must do, it is first necessary to know what the finished product must contain, and it is appropriate to carry out the analysis in the preceding manner because of the natural correspondence between the parts of the product

65 The foregoing represents an extremely important point, and it will play a prominent role in several subsequent discussions—see below, pp. 391, 393-94, 433.
66 See above, p. 364.
and the parts of the process.

Thus, none of Else's arguments succeed in proving that the qualitative elements are not parts of tragedy qua artistic product. At most, they provide support for the contention that the elements are at least parts of the artistic process, which is perfectly compatible with their also being parts of the product.

Interestingly enough, Else has rejected as spurious or unauthentic portions of two passages in the Poetics that furnish much stronger evidence that the elements are indeed parts of the artistic process than any of the passages so far discussed. At the close of chapter 5 in the Poetics, Aristotle remarks that although tragedy and epic have some parts in common, the former has some additional parts that the latter does not have. The word Aristotle uses for epic in that passage, "epopoiia" [ἐποποιία], clearly means "epic-making," which indicates that he must also be speaking of tragedy qua activity; a comparison of the parts of epic

---

67 Else's reasons for rejecting the portions in question do not seem very convincing, but a discussion of their plausibility would require more space than present purposes allow—and, since proof of their legitimacy would only serve to establish more firmly a point with which Else already agrees, such a discussion is not really necessary. Naturally, the following remarks will, however, be accompanied by footnotes that cite the page numbers on which his arguments can be found.

68 Poetics, 5: 1449b16-20. Else does not reject any of the preceding material as spurious—see Else, p. 203.
qua activity and tragedy qua activity obviously makes better sense than a comparison of the former as an activity and tragedy qua product. In chapters 24 and 26, Aristotle returns to the unfinished business of specifying just which parts tragedy and epic have in common, again the term "epo-poia" is used, and his answer is: plot, character, thought, and diction. Thus, if legitimate, the passages in question provide convincing evidence that the qualitative parts do in fact constitute parts of the artistic process.

Thus far, discussion has indicated that the Poetics contains fairly strong evidence that the qualitative elements are parts of tragedy qua artistic process and that it does not appear to contain any evidence that the elements are not also parts of tragedy qua product. In fact, earlier remarks have shown that in spite of his intentions to do otherwise, Else has uncovered evidence that the six elements do also constitute parts of the product. Might the Poetics contain additional evidence that the elements are parts of the product? The answer would appear to be that it does.

First of all, the Poetics contains a number of remarks that would not make any sense unless the parts constituted

---

69 Poetics, 24: 1459b7-17 and 26: 1462a13-16. The aforementioned passages contain the key phrases which Else considers to be spurious—see his remarks on pages 598 and 643 respectively.
parts of the completed product. In fact, several examples can be drawn from the foregoing discussion—specifically, Aristotle's remarks about characterless and plotless plays. Whether those remarks are interpreted as statements that some plays are so badly constructed that one or more of the six parts is "as good as" missing or as statements that one or more parts may be literally missing, the parts in question, specifically plot and character, must be parts of the plays themselves—not just parts of the artistic activities which brought them into being.\(^70\) Even Else admits that the relevant occurrences of the term "tragodia" must be interpreted as references to tragedies qua completed product.\(^71\) He fails to realize, however, that since that is true, references to the parts of those tragedies must refer to parts of the finished products. It is, moreover, significant that the remarks in question occur not as incidental observations in a later chapter, but as an integral part of Aristotle's introduction of the six elements in chapter 6; the remarks cannot, therefore, be passed off as a brief recognition by Aristotle that the six terms can, by the way, sometimes be used to refer to parts of the finished play—\(\text{they}\)

---

\(^70\) *Poetics*, 6: 1450a24-34. See also above, pp. 372-75.

\(^71\) Else, pp. 6, 252.
constitute an integral part of his discussion of what he normally means when he uses the terms.

Aristotle's observation that plot, itself a part of tragedy, can, in turn, be broken into parts furnishes additional proof that the six terms refer to parts of the finished product. The passages in which Aristotle identifies the three primary parts of plot--suffering, discovery, and reversal--clearly indicate that they are parts of the play. For example, Aristotle defines "suffering," or "pathos" \([\pi\alpha\omicron\delta\omicron\upsilon\sigma]\), as a "destructive or painful action."\(^{73}\) It is, of course, the dramatis personae--not the poet--who undergo such suffering, and they do so in the play. In fact, throughout his discussion of the parts of plot, Else himself speaks of them as parts of the play.\(^{74}\) How, then, could plot not be a part of the poetic product when its constituent parts are so? The only conceivable way would be if suffering, discovery, and reversal were viewed as ingredients used in the poetic process much as, for example, flour, milk, and eggs might, in somewhat loose fashion, be

\(^{72}\) Poetics, 10-11: 1452a12-1452b13.

\(^{73}\) Telford translation, 11: 1452b13.

\(^{74}\) See, especially, page 344, where Else emphatically states that reversal is "a change of or in the events of the play"--the emphasis appears in the original.
called "parts" of the cooking process. However, even that suggestion will not work. In the preceding example, the ingredients, flour, milk, and eggs, exist before the cooking process even begins. However, Else's own discussions clearly suggest that the three parts of plot are parts of the finished product; in other words, by Else's own account, the three parts result from rather than precede the poetic process. How then could the three parts be parts of the plot if the plot is supposed to be the creative process? The foregoing would require the three parts to constitute parts of a process which temporally precedes their very existence, which is clearly absurd. Only one possibility is left; plot, like all of the other qualitative elements, is a part of the finished product.

The preceding remarks bring to a close discussion of whether the six terms used by Aristotle are supposed to refer to parts of the poetic process or parts of the poetic product. The evidence considered during the course of that discussion suggests that the earlier interpretation derived from Telford's remarks is, for the most part, correct; the terms refer to both the product and the process—primarily to parts of the poetic product and derivatively to the parts

75 Else, p. 344.
of the poetic process. Before discussion of the parts on an individual basis, starting with spectacle, begins, two more points that bear upon the character of the six parts as a group should be mentioned.

At the end of his introduction of the parts in chapter 6, Aristotle states that they are what give tragedy its particular quality, "poia" [ποία]. It is the reference in that remark to the six parts as responsible for the quality of tragedy, and the sharp contrast between that remark and the twice repeated assertion that a list of parts identified in chapter 12 have to do instead with quantity, "poson" [ποσόν], that has led to the distinction between the "qualitative" and "quantitative" parts. A satisfactory discussion of the difference between the two will have to be postponed, but one observation should be made now.

Aristotle's reference to the six parts as being responsible for the quality of tragedy follows hard upon the heels of his proof that they are necessary to tragedy as the imitation of an action carried out directly by the agents, and it would appear to be precisely the dramatic quality

---

76 Poetics, 6: 1450a8-10.
77 Poetics, 12: 1452b14-16 and 25-27.
78 See below, pp. 418-19.
of tragedy, rather than other of its qualities, for instance its serious nature, that depends upon their presence. Naturally, the specific character impressed upon the various parts would be sufficient to differentiate a tragedy from a comedy; for example, if a play's plot imitated a ludicrous action, that would be enough to mark the play as a comedy rather than a tragedy. However, the mere presence of the six parts per se would not be sufficient to differentiate comedy from tragedy. Their mere presence is, however, sufficient to differentiate tragedy from all non-dramatic imitative arts, e.g., painting, sculpture, music, and even epic poetry. The six parts are, thus, responsible for the quality of tragedy insofar as it is a form of drama.

The second remark that bears upon the character of the qualitative parts as a group occurs several lines after the one just discussed. Before beginning an extended comparison of plot and character, Aristotle includes one more comment concerning the parts as a group, and in that remark, he refers to them as "formative elements" or "forms," "eide" [ἐἴδη]. The implications of that comment cannot be fully assessed at present, but they will be explored during the

79 Poetics, 6: 1450a13. The Bywater translation contains "formative elements"; Telford's contains "forms."
course of subsequent discussion.  

SPECTACLE

As earlier remarks have already indicated, Aristotle opens the passage in which he introduces the six qualitative parts with a reference to what translations usually refer to as "spectacle":

Since agents [rather than the poet] produce the imitation, first, of necessity, [1] one part of tragedy will be the ornament of spectacle.

Although Aristotle makes ten other references to spectacle, most of those concern not the character of spectacle as such but rather its place in the poetic process and its contribution to the merits of tragedy as a whole. The fact that the Poetics does not contain a full explanation of what Aristotle meant by the term has, not surprisingly, led to disagreement over what he had in mind when using it. Specifically, the controversy concerns the question of whether Aristotle's word for spectacle, "opsis" [\(\omega\pi\sigma\)\(\varsigma\)], is supposed to refer just to the costuming of the dramatis personae or

---

80 See below, pp. 527-28.

81 Telford translation, 6: 1449b32-4.
to all of a tragedy's visual elements; and, again, Else and Telford come down on opposite sides of the issue.

Else maintains that "opsis" refers exclusively to the costuming of the dramatic personages, and he bases his case on five arguments. First, he argues that since all six parts are deduced from the idea that tragedy represents an action carried out by the dramatis personae, "the usual interpretation of ὁψίς as 'spectacle,' i.e., the staging or mise en scene in general" must be rejected in favor of "Bywater's suggestion that it means the 'look' or visible manifestation of the characters. Ὅ τῆς ὁψεως κόσμος is, then, the adornment of their appearance."82 Second, Else claims that Bywater's suggestion would explain Aristotle's use of the Greek term in both its singular and plural form "for the appearance(s) of the characters [i.e., the dramatis personae] can be considered either collectively (ὁψίς) [opsis] or distributively (ὁψεῖς) [opseis]."83 Else's third argument concerns Aristotle's remark, at the end of chapter 6, that the execution of opseis depends less on the art of

82 Else, pp. 233-34.
83 Else, p. 234. The Greek material, enclosed in parantheses, appears in Else; the transliteration of that material, enclosed in brackets, has been added.
the poet than on the art of the "skeuopoioi" [σκευοποιοῖο]--a term many translators believe referred to the costumer. According to Else, that remark virtually proves that "opsis" refers to the costuming.

Else's last two arguments both concern a passage contained in chapter 14, where Aristotle observes that although it is possible to produce fearful and pitiful effects through the use of opseis, to do so is less artistic than to produce those effects by means of the plot--besides which, opseis are matters that require the services of a choregos. Commenting on that passage, Else insists that there cannot be any doubt that under Greek conditions—in the open air, and with a stage-set which offered only limited possibilities of alteration for tone and mood—any emotional effect of the stage-arrangements must have come primarily from the masks and costumes. And indeed it is in connection with them, particularly in connection with the famous first staging of the Eumenides, that we hear of such effects.

Finally, Else claims that Aristotle's mention of the choregos

---

84 Poetics, 6: 1450b16-20 and Else, pp. 287-79.
85 Else, pp. 278-79.
86 Poetics, 14: 1453b1-9.
87 Else, pp. 409-10.
supports his interpretation because it "was associated definitely with the provision of masks and costumes, not with the stage building and other stage-arrangements."\(^{88}\)

In his commentary, Telford reasserts the traditional view, remarking that spectacle "is simply the visual aspect of the drama, the scenery, costumes, gestures, etc."\(^{89}\) However, he does not go on to mention, much less respond to, Else's arguments; nor does he offer any evidence to support his own position. Is there, nevertheless, anything that can be said in response to Else's arguments? The answer is yes.

First of all, there is the fact that the Greek term "opsis" was associated with not only the appearance of humans, but with the visual appearance of any concrete object—meaning in its broadest sense "that which is seen." Since Aristotle is using the various words for the six parts as technical terms, care must, of course, be taken not to give their literal meaning outside the province of the Poetics more weight than it deserves. Nevertheless, the fact that "opsis" was not normally restricted to the appearance of people, much less dramatis personae, must weigh against Else's reading.

\(^{88}\) Else, p. 410.

\(^{89}\) Telford, p. 86.
Next, although Else's first argument presents fairly strong evidence that Aristotle thought the visual manifestation of the *dramatis personae* to be the only necessary aspect of *opsis*, it certainly does not prove that Aristotle conceived of *opsis* exclusively in such terms. As earlier remarks indicated, Else argues that since Aristotle attempts to derive *opsis* from the idea of tragedy as an action carried out directly by the *dramatis personae*, it includes only their visual appearances. Part of what Else has to say is correct. During lines 1449b32-1450a11 Aristotle is indeed concerned with the six parts insofar as they are each necessary to tragedy *qua* drama, i.e., tragedy *qua* representation of an action carried out directly by the *dramatis personae*. Furthermore, Else appears to be correct in assuming that the visual manifestations of the *dramatis personae* are the only aspects of *opsis* that are absolutely necessary to tragedy *qua* drama; if the action is to be performed directly by the *dramatis personae*, then they at least must be seen—their fictitious environment may or may not be represented.

Interestingly enough, it would appear to be *opsis* *qua* visual manifestation of the *dramatis personae* that also constitutes one of the key differences between tragedy and epic.

90 See above, p. 388.
Late in the *Poetics* Aristotle twice indicates that tragedy—but not epic—possesses *opsis*, and he might be interpreted to mean that epic was not associated with the production of visual effects at all. However, a remark in chapter 26 shows that Aristotle cannot have had the preceding interpretation in mind. After observing that some people believe tragedy to be a lesser art than epic because it panders to a cruder audience through the use of gesture, Aristotle points out that

> the censure does not touch the art of the dramatic poet, but only that of his interpreter; for it is quite possible to overdo the gesturing even in an epic recital, as did Sosistratos, and in a singing contest, as did Mnasitheus of Opus.

Thus, epic recitations could apparently be accompanied by a fair amount of visual reinforcement. The speaker could even gesture as if he were now one of the *dramatis personae*, now another. What the rhapsode or reciter could not, however, do was to let the audience see separate concrete embodiments of the various *dramatis personae*, gesturing and moving about together, and that is a difference so fundamental that it

---

91 *Poetics*, 23: 1459b9-11 and 26: 1462a14-18. Else considers the references to *opsis* in both passages to be spurious, but he nowhere denies that Aristotle believed the epic form lacks *opsis*—see Else, pp. 594-98 and 643-44.

92 Bywater translation, 26: 1462a6-8.
effects the character of any and all additional visual effects as well.

Even if an epic were to be recited in front of elaborate scenery, the scenery would not have anything like the effect it would have in the performance of a tragedy; in the latter case, the concrete embodiments of the various dramatis personae would mark the scenery as part of their environment—as a part of opsis, whereas the same scenery would simply constitute an elaborately appointed speaking platform for the rhapsode—it would not constitute opsis. Thus, there would seem to be an additional reason for thinking that Aristotle was especially interested in the appearances of the dramatis personae; they not only constitute the only aspects of opsis that are necessary to tragedy qua species of dramatic artifact they also constitute an important means of distinguishing tragedy from the epic.

Although Else is probably correct in thinking that the appearances of the agents constituted the aspects of opsis that most interested Aristotle, that does not constitute sufficient grounds for restricting the referential scope of "opsis" to the costuming of the dramatis personae; after all, interest in one aspect of a phenomenon need not prevent a person from recognizing that the phenomenon possesses other aspects as well. Moreover, if Aristotle's list of the
six parts is to be exhaustive, as he claims it to be, then opsis must be allowed to include everything in a tragedy that can be seen. The visual manifestations of the dramatis personae may be what make any additional visual effects in a tragedy part of opsis—as opposed to the visual effects associated with epic which are not part of opsis—but that does not make those additional effects less a part of opsis.

Before the plausibility of Else's second argument is assessed, an additional point related to the foregoing discussion should be made. The preceding remarks have spoken of opsis in primarily performance-oriented terms, which might seem to lend credibility to Hornby's contention that Aristotle saw spectacle "as something that is merely added to the text"—as something that is not in the play as written, but which is manifested only under the conditions of performance.93 The preceding remarks have, however, spoken of opsis in performance-oriented terms only because such talk seems more concrete and tangible than talk about the spectacle inherent in the written play; it has not been intended as an indication that Hornby's view is correct. In fact, the text of the Poetics makes it quite clear that

93 Hornby, p. 75.
Hornby's view is mistaken. For example, if Hornby were right, why does Aristotle explicitly claim that spectacle constitutes a necessary part of tragedy in the same breath that he declares the other five elements to be necessary? It would appear that in order to reconcile his claim that spectacle arises only when a play reaches production with the aforementioned passage, Hornby would have to claim that Aristotle thought all six parts arise only when a play reaches production—which is out of the question.

Moreover, if Hornby were right, and Aristotle thought that spectacle manifests itself only after a play has left the poet's hands, why does Aristotle caution would-be poets against "offenses against the sense perceptions that necessarily attend upon the poetic art" in chapter 15? That remark clearly suggests that, while those who are directly involved in the performance process may have more to do with the execution of the sensual aspects of tragedy, the poet too has his responsibilities, responsibilities which he must necessarily discharge before the play reaches even the first stage of the rehearsal process. Where those responsibilities concern the visual aspects of tragedy, they fall under

---

94 Poetics, 6: 1449b32-3.
95 Else translation, 15: 1454b16.
the heading of opsis. Thus, as Else observes, opsis is not exclusively connected with

performance on a stage, by living actors, but [is] a feature inherent in the drama as such. It is inherent in the drama of Othello that Othello is a Blackamoor, or that of Lear that Lear is a kingly old man. The concept of these characters [i.e., dramatis personae]--and it is a concept implicit in the idea of a drama--is that they will 'look' or appear in such and such a way: as a Moor, not a white man, a king, not a commoner, etc. Moreover, the characters [dramatis personae] must be thought of ("seen," i.e., by the poet: 17. 55a 24, 27) as being in certain places, moving in certain directions, being on or off stage at certain times, etc. These are necessities that the dramatist cannot dodge, a condition which he has laid upon himself by the act of writing a play, and which must necessarily affect the way he writes it and the way a reader--independently of any actual performance--will visualize it.\(^\text{96}\)

Else's remarks would appear, moreover, to hold true for aspects of opsis in addition to those connected directly with the visual manifestations of the dramatis personae.

In Euripides' Electra, for example, the agents make repeated references to the immediate proximity of a hovel in which Electra lives with a husband whom her mother and stepfather have forced her to marry.\(^\text{97}\) Moreover, those

\(^{96}\) Else, p. 234.

\(^{97}\) See, for example, lines 207-08 and 251-52.
references make it clear that the presence of the hut represents a visible symbol of Electra's enforced poverty and disgrace at the hands of her mother and stepfather. As such, the presence of the hut constitutes one of the key forces that goad Electra on, urging her to seek revenge. Since it constitutes such an integral part of the play's action, the hut must, therefore, appear onstage—even if the "stage" in question exists only in the imagination of a reader. It is, likewise, inherent in Aeschylus' *Agamemnon* that the first scene takes place at night—whether the play is taking place on a real stage or only in the imagination of a reader. Indeed, it is nighttime even when the play is performed on an outdoor stage in broad daylight. The audience may sit in bright sunshine, but onstage the watchman complains that his make-shift bed is "stricken with night and drenched with dew," and suddenly he sees a flare "burning from the blackness." In the play it is darkest night, and that darkness is a part of the play's *opsis*—onstage or off.

---

98 See, again, lines 207-08 and 251-52.

Else's second argument claims that interpreting "opsis" as a reference to the costuming would help to explain Aristotle's use of the term in both its singular and plural forms since the appearance(s) of the agents can be spoken of either way.\textsuperscript{100} However, the entirety of a play's visual appearance(s) can be just as easily spoken of in either singular or plural terms. Even in English, a person may speak of either a play's "visual aspect" or of its "visual aspects." Moreover, the correspondence between a particular ancient Greek and a modern English expression as regards the formation and use of their plural forms hardly constitutes proof positive for the purposes of translation and interpretation; after all, ancient Greek and modern English do not always allow a person to treat words with the same basic meaning in exactly the same ways grammatically.

Else's third argument is based on Aristotle's remark that the \textit{opseis} depend more on the art of the \textit{skeuopoioi} than the poet.\textsuperscript{101} However, the exact duties of the \textit{skeuopoioi} are far from clear. Butcher treats the term as a reference to "the stage machinist."\textsuperscript{102} In fact, Else himself acknowledges that the term may have referred to either

\begin{itemize}
\item \textsuperscript{100} See above, p. 388.
\item \textsuperscript{101} See above, pp. 388-89.
\item \textsuperscript{102} Butcher, p. 31.
\end{itemize}
"the 'property man' or stage costumer." Even if the duties of the person in question were, however, restricted to the costuming, Aristotle's remark does not say that the execution of opseis depends exclusively on him; it merely suggests that the execution of opseis is heavily dependent on him. Thus, like Else's first argument, the present one furnishes fairly plausible evidence that Aristotle tended to think of the appearances of the dramatis personae as the most essential aspects of opsis, but falls far short of proving that he conceived of opsis exclusively in those terms.

Finally, the two arguments that Else bases on Aristotle's remarks in chapter 14 must be evaluated. Else is no doubt right that the use of costumes and masks represented one of the most powerful visual means of producing fearful and pitiful effects, but it was certainly not the only one. The use of various pieces of stage machinery, e.g., the ekkuklema [ἐκκύκλημα], which was often used for the purpose of revealing dead bodies—like those of Agamemnon and Cassandra in the Agamemnon, and the mechane [μηχανή] or

103 Else, p. 278.
104 See above, pp. 389-90.
crane, which was used for miraculous entrances and exits—like Medea's exit at the end of Medea, cannot have left audiences completely unmoved.105

Last, Else's argument based on Aristotle's reference to the choregos must be considered. That argument falls prey to two major objections. First, even Else admits that his interpretation would require that the remark be given "some extension of meaning, since the choregia in the strict sense probably did not cover the actor's masks and costumes but only those of the chorus."106 Thus, a stretch in meaning is already required, and there is no apparent reason why the meaning could not be stretched to include the provision of not only masks and costumes, but the provision of whatever is necessary for the performance of a tragedy. Second, even if Aristotle did have the provision of only the costumes and masks in mind, his remark merely states that opseis require those provisions; it does not state that opseis require only those provisions.

As the foregoing discussions have shown, none of Else's arguments succeed in proving that Aristotle intended

105 For information on the aforementioned pieces of stage machinery, see, for example, Brockett, p. 39.
106 Else, p. 410.
"opsis" to refer exclusively to the costuming. At most, those arguments suggest that Aristotle believed the appearances of the *dramatis personae*—which properly include not only their dress but also the visual aspect of their activities—to be the most essential aspects of *opsis*, and that he was especially impressed by the contribution to *opsis* made by the costumes and the masks.

Besides providing an assessment of Else's arguments, the remarks contained in the foregoing discussions suggest that a number of rather subtle, but important distinctions are implied by Aristotle's remarks on the subject of *opsis*, distinctions that have not been recognized by previous commentators. First, the sense of "opsis" that refers to part of the imitative product must be distinguished from the sense that refers to part of the imitative process. The former sense refers to the entire visual aspect of the product—*to everything* in the product that can be seen; the latter sense refers to the composition of the product's visual elements. Next, parallel distinctions must be made within each of the two preceding senses; *opsis* as part of the product must be separated into *opsis* as part of the written play and *opsis* as part of the performed play, and, likewise, *opsis* as part of the process must be separated
into *opsis* as part of the poet's process and *opsis* as part of the production process. Finally, those aspects of *opsis* that are necessary to tragedy *qua* drama, aspects which are necessary to tragedy in both its performed and written form, must be distinguished from those that are not strictly necessary. For example, in both the written and the performed product, the appearances of the *dramatis personae* represent the most essential aspects of *opsis*; the nature of the product requires that at least some of the dramatic personages' visual manifestations, *e.g.*, their entrances and exits, be specified in performance and text alike, but either may fail to specify the fictional environment in which those personages live. While the visual manifestations of the *dramatis personae* represent the most essential aspects of *opsis* in both the written and performed product, the two differ, however, as regards the amount of those aspects they must necessarily specify; a written text may dispense with all but the entrances and exits and perhaps the specification of a few other visual manifestations, but the performed product cannot help but specify every possible facet of the dramatic personages' visual manifestations.

Similar observations can be made in the case of *opsis* as a part of the creative process. Again there are certain
aspects of opsis that constitute a necessary part of both the poet's process and the production process, viz., the specification of at least some of the agents' visual manifestations, and again the two processes differ as regards how much of the agents' visual manifestations must be specified. The poet need not—indeed, he could not—specify them completely; the members of the production group cannot avoid doing so—and that is why Aristotle says the execution of opsis is more heavily dependent on the activities of the skeuopoioi and his fellow workers than on those of the poet.

One final remark should be made before the present discussion of opsis draws to a close. Although Else's objection to the translation of "opsis" as "spectacle" is not entirely groundless since the term "spectacle" cannot normally be used when referring to a process, his translations of it as "the adornment of their physical appearance" or "the costuming" miss the mark by a far wider margin. A more adequate translation of "opsis" would probably be "visual composition" since that expression can be used to refer both to that which is composed and to the compositional process. Nevertheless, "spectacle" has become standard, and as long as its use as a technical term in the Poetics is stretched to include both product and process, no harm will come from
While previous commentators agree that the term usually translated as "diction," "lexis" [λέξις], was associated with at least some of a tragedy's language, they disagree about how much of that language was associated with it and about the intended meaning of the term usually translated as "melody" or "music," "melopoiia" [μελοποιία]. The vast majority of commentators have assumed that "lexis" refers to the entirety of a play's language and that "melopoiia" refers to the music that accompanied certain portions of Greek tragedy in performance, but Else disagrees. He has argued that "lexis" refers to the composition of only those lines in a Greek tragedy that were to be spoken and that "melopoiia" refers to the composition of "everything in the song[s], words and melody alike, not just the melody." In other words, Else believes that the two terms are linked to a segmentation of Greek tragedy into spoken dialogue versus song.

107 For a statement of the traditional view, see Cooper, pp. 33-34.

108 Else, pp. 236-37. All of the quoted material is italicized by Else.
Else presents two major arguments in support of his position. During the first argument, he claims that a series of remarks scattered throughout the early part of the Poetics establishes a sharp antithesis between speech and melody, an antithesis that reaches its fullest expression in a remark that immediately follows the definition of tragedy in chapter 6.109 Aristotle has just stated that tragedy includes speech that has been "sweetened" (hedysmenon [ἡδυσμένον]) with "each form [of pleasing ornamentation] used separately in the parts of tragedy," and he wishes to further clarify the meaning of that statement:

By speech made pleasing I mean speech which possesses rhythm, harmony, and melody. By forms used separately I mean that some parts are achieved through meter only while others are achieved through melody.110

Else argues that the subsequent introduction of "melopoia" and "lexis" should be understood as a subtle reference back to the preceding passage, the two terms representing the names of the processes involved in composing the two basic "parts" or types of pleasing speech: "melopoia" refers to the composition of that part in which melody is used—namely

109 Else, pp. 36-7, 179-81 and 223-24.
110 Telford translation, 6: 1449b28-32.
the songs; "lexis" refers to the composition of that part in which meter alone is used—namely the dialogue. 111

Else's second major argument depends on his interpretation of a remark that occurs towards the end of chapter 6. The remark in question asserts that of the remaining parts, melopoiia is the greatest of the "sweetenings" [hedysmaton]. 112 Else argues that since the remark contains the nominal form of the word Aristotle used earlier when talking about the parts of "sweetened speech," "hedysmenon," Aristotle must have considered melopoiia to be a part of sweetened speech. 113 Since something cannot be a part of sweetened speech unless it has words, reasons Else, "melopoiia" must refer to the composition of both the music and the text of the songs. 114

Instead of arguing in support of either Else's interpretation of "melopoiia" and "lexis" or the traditional one, subsequent remarks will attempt to argue in favor of yet another interpretation of the terms. Before discussion of that interpretation begins, however, the plausibility of

111 Else, pp. 236-37.
112 Poetics, 6: 1450b15-7.
113 Else, p. 277.
114 Else, p. 278.
Else's two major arguments must be assessed. Since Aristotle's remark about the parts of sweetened speech in chapter 6 definitely establishes a distinction between dialogue and song, Else's attempts to show that passages towards the beginning of the *Poetics* also establish that distinction need not be discussed; the distinction is assured regardless of whether Else's interpretation of the earlier passages is plausible.\footnote{115}

\footnotetext{115} Although the above discussion neither requires nor has room for a detailed discussion of Else's attempts to show that the distinction between dialogue and song is established during passages near the beginning of the *Poetics*, a few programmatic remarks about those attempts are offered here. Part of what Else says would appear to be correct. His claim that the remarks in question establish an antithesis, and his claim that speech represents one side of that antithesis both appear to be warranted. However, the passages in question repeatedly contrast the presence of speech—not with the presence of melody as he suggests—but, instead, with the absence of speech. For example, instead of contrasting the arts of instrumental music with arts in which speech without music is used, as Else would have it (see Else, pp. 36-7), the first series of remarks (those contained in chapter 1: 1447a18-1447b29) contrast the arts in which speech is used—whether it is used with or without melody—with those in which speech is not used—not only instrumental music but also dance. The whole passage is clearly intended to distinguish the literary arts, those which use speech, from the non-literary arts; the presence or absence of melody is extraneous to the contrast Aristotle is establishing. Similarly, the history of tragedy contained in chapter 4 (lines 1449a15-24) is not intended to highlight a movement from song to spoken dialogue as Else thinks (see Else, pp. 179-81); it is, instead, intended to recount a movement from an emphasis on non-literary elements in early tragedy—an emphasis on melody and dance—to an increasing emphasis on the literary aspects of tragedy.
The crucial question to be asked when considering the plausibility of Else's first argument is whether or not Aristotle's subsequent introduction of "lexis" and "melopoia" is supposed to refer back to the remark about the parts of sweetened speech, thereby linking the two terms to the earlier distinction between dialogue and song. Else bases his claim that the later remark is supposed to refer back to the earlier one on the fact that the Greek word "auten" [αὐτήν], which means "by itself" or "alone," is included in the later remark's statement that "lexis" refers to the synthesis of the meters. According to Else, Aristotle must be construed as saying that by "lexis" he means the synthesis of the meters by themselves or alone—as opposed to the combination of meters with music. If that is the case, argues Else, then Aristotle must have had the contrast between the parts of sweetened speech in mind; "lexis" must be tied to the parts of sweetened speech achieved through meter alone, i.e., dialogue, and "melopoia" must be tied to the parts achieved through melody, i.e., the songs.¹¹⁶

The preceding argument is, however, open to a number of objections. First, other translations suggest that the term

¹¹⁶ Else, p. 236.
"auten" does not play a contrastive role in the remark being considered. For example, Bywater's translation reads as follows: "by 'Diction' I mean merely this, the composition of the verses." Moreover, there are several reasons for being sceptical about the alleged connection between the two remarks in question. For one thing, in the earlier passage Aristotle indicates that one part of sweetened speech consists of meter used by itself through the use of the term "monon" [μόνον]—not through the use of the word "auten." For another thing, there is no apparent reason why Aristotle would not have used the terms "melopoia" and "lexis" in the earlier remark about the parts of sweetened speech if he did, in fact, intend the terms to represent technical names for the dialogue versus song distinction. Aristotle's liberal use of the terms for plot and character throughout early chapters, starting with a reference to plot in the second line of chapter one, constitutes ample evidence that he was not waiting until he could introduce all six of the terms together.\footnote{Poetics, 1: 1447a1.} It is, moreover, odd that the later remark contains no reference to the supposed connection between "lexis" and "melopoia" and the parts of
sweetened speech mentioned earlier; the Poetics contains numerous passages in which Aristotle explicitly refers back to remarks he has already made.\(^{118}\)

Although a fair amount of evidence that weighs against Else's first argument has already been presented, the most powerful piece of evidence has not yet been discussed. During chapter 12, the quantitative parts of tragedy are differentiated from each other precisely on the basis of whether the language contained in each part is spoken or sung:

The parts of tragedy which must be treated as elements of the whole [as forms—"eide"] have already been mentioned. We now come to the quantitative parts—the separate parts into which Tragedy is divided—namely Prologue, Episode, Exode, Choric song; this last being divided into Parode and Stasimon. These are common to all plays: peculiar to some are the songs from the stage and Commoi.

The Prologue is that entire part of a tragedy which precedes the Parode of the Chorus. The Episode is that entire part of a tragedy which is between complete choric songs. The Exode is that entire part of a tragedy which has no choric song after it.\(^{119}\)

Thus, Aristotle's division of sweetened speech in chapter 6 would appear to be linked not with his discussion of lexis

\(^{118}\) See, for example, Poetics, 1: 1447b24-5, 3: 1448a24-5, 5: 1449a33, 6: 1450b14, etc.

\(^{119}\) Butcher translation, 12: 1452b15-22.
and melopoia, but, instead, with his discussion of the quantitative parts in chapter 12; it is there that he elaborates on the division of tragedy into the spoken versus the sung parts.

How does Else attempt to block the obvious connection between Aristotle's statement about the parts of sweetened speech and the contents of chapter 12? He argues that all but the opening sentence of chapter 12 is spurious. His reasons for thinking it so are: 1. if it were not spurious, the relation between the first two sentences of chapter 12 would require the first sentence to refer to the six qualitative parts as parts of tragedy qua product, which his interpretation does not allow, and 2. he believes the identification of the quantitative parts represents a less organic and less systematic set of distinctions than is represented by the supposedly correct interpretation of "lexis" and "melopoia":

Instead of his organic division between the two species of 'sweetened language'--i.e., speech and song, defined by the absence or presence of melody--we find here a purely mechanical division based on no principle and carried through in the crudest possible manner. Of the four 'parts,' the first three (prologue, 'episode,' exodos) are obviously thought of as scenes, "separated" from each other by the choral elements.... The distinction among them is simply their position in the play: prologue at the beginning, exodos at the end, 'episodes' in between. Thus the list presents the absurdity of a
Thus, Else argues that the entire approach manifested in chapter 12 brands it as spurious.

As earlier remarks have shown, Else's contention that the six qualitative parts cannot refer to parts of tragedy qua product is mistaken.\textsuperscript{121} His first reason for thinking that chapter 12 is spurious may, therefore, be dismissed without further discussion. His second reason will, however, require more extensive consideration.

Since Else claims that the procedure employed in chapter 12 is crude and unsystematic, it may be wise to have a look at that procedure. Tragedy is first divided into parts on the basis of whether the various parts are spoken or sung: the prologue, episodes, and \textit{exode} are spoken—all of

\textsuperscript{120} Else, pp. 362-63.

\textsuperscript{121} See above, pp. 384-85.
the sung parts are loosely referred to as "choral parts," "chorikon" [χορικόν]. The sung parts are then differentiated on the basis of who performs them: there are the songs performed by the chorus alone (the parode and stasimon), those performed by the actors alone, and the commoi—songs performed by the actors and chorus together. The spoken scenes are then differentiated from each other on the basis of their relation to the songs performed by the chorus alone: the prologue is that scene which has no choral song before it, the exode is that scene which has no choral song after it, and the episodes are those scenes which are both preceded and followed by choral songs.

In addition to his claim that the foregoing procedure is less organic than the division of sweetened speech in chapter 6, Else raises two minor objections. First, he objects to the fact that the first four parts identified appear to put three already separated spoken parts on the same footing with what seems to be a genus term for the sung parts. Second, he objects to the inclusion of the actors' songs under the heading "choral parts." In fact, Else suggests that Aristotle wished to completely ignore the commoi and the songs from the stage, a maneuver that would

122 See above, pp. 411-12.
allow him to draw a cleaner distinction between the spoken dialogue on the one hand, and the choral odes on the other. 123

The first of the preceding objections does represent a problem. Aristotle is usually far too methodical to compile a list in which one or more genus terms are thrown together with non-genus terms. There are, however, reasons for thinking that he decided to make an exception in the present case. He was presumably trying to work with the standardized theatrical vocabulary of the time, a vocabulary that furnished him with technical names for the smaller parts but did not furnish him with specific names for either the spoken scenes as a group or for the sung portions as a group. How could Aristotle respect the traditional terminology and still draw the necessary distinctions? Perhaps he might stretch the term for the choral parts of tragedy to include all of the sung portions and set it over against the various names for the spoken scenes.

The preceding hypothesis would not only help to explain why a genus term is included in the same list as non-genus terms, it would also help to explain why the

123 Else, p. 224.
actors songs are included under the heading of "choral parts": Aristotle needed a term that would allow him to lump all of the sung portions of tragedy together, and since the musical portions of tragedy were mostly associated with the chorus anyway, it would not be too difficult to stretch the traditional term for choral parts, "chorikon," to include everything that was sung. At any rate, the preceding hypothesis seems equally—if not more—plausible than Else's suggestion that Aristotle wished to simply gloss over the commoi and the actors arias in favor of a cleaner distinction between dialogue and song. In chapter 6, Aristotle separates sweetened speech into the spoken versus the sung parts, which presumably includes all of the sung parts—not just the choral odes; and any supposed attempt to gloss over the fine points involved in that distinction would certainly be out of keeping with Aristotle's usual penchant for detail and fine distinctions. Else does not, moreover, cite any textual evidence in support of his suggestion.

Of more importance than the minor objections just discussed is Else's contention that the remarks in chapter 12 treat dialogue and song as discrete pieces of a play's text, and that such treatment does not seem nearly as "organic" as the introduction of the six parts in chapter 6.\footnote{See above, pp. 411-12.}
Else appears to be quite right about the preceding matters. The remarks in chapter 12 do indeed treat dialogue and song as discrete chunks of the play. However, by Else's own admission, Aristotle's reference to the parts of sweetened speech immediately following the definition of tragedy divides the play into equally discrete segments; Aristotle has, as Else flatly says, "made it clear that he regards the play as falling into just two alternating 'parts,' the verses (dialogue) and the songs." Thus, there is no reason for thinking that the remarks in chapter 12 represent a less "organic" division of tragedy than the one Else has in mind. It is true that the distinction between dialogue and song is not drawn as cleanly in chapter 12 as it is in the earlier remark concerning the parts of sweetened speech, but the distinction itself is manifestly the same in both cases—spoken dialogue versus song as discrete pieces of text.

The remarks contained in the preceding paragraph have shown that Else's distinction between dialogue and song is every bit as "inorganic" as the distinctions made during the course of chapter 12. Else's claim that "lexis" and "melopoia" refer to the parts of sweetened speech, or to

125 Else, p. 224.
the composition of those parts, therefore implies that the two terms are associated with a distinction that is, by his own admission, completely out of place in Aristotle's discussion of the qualitative parts. By way of contrast, the author of chapter 12—presumably Aristotle himself—completely avoids the preceding problem. In the first two lines of the chapter, the author explicitly notifies the reader that the ensuing distinctions, which include the division of tragedy into dialogue and song, are of an entirely different order than are those which differentiate the qualitative parts. The parts that should be treated as forms of the whole, i.e., the qualitative parts, have, says the author, already been discussed; it is now time to discuss parts of a different sort—the separate parts into which tragedy is divided. It is of particular interest that he utilizes the term "kechorismena" [κεχορισμένα] as a means of indicating that the quantitative parts are used "separately" because it contains the same word, "choris" [χωρίς], that was used in chapter 6 to indicate that the parts of sweetened speech are "separate" parts: choris, ke(choris)mena. There is, then, a clear connection established between the division of sweetened speech in chapter 6 and the remarks in chapter 12: they deal with the

126 See above, pp. 410.
same subject matter—the parts of tragedy that are used separately.

During the first two sentences of chapter 12, the author not only indicates that the ensuing parts are used separately, he also indicates that they possess quantity, "poson" [ποσόν]. Thus, the parts discussed in chapter 12 divide tragedy into a series of temporally discrete quantitative segments—separate "pieces of text," as Else would say. How do those parts differ from the six parts identified in chapter 6? The parts identified in chapter 6 do not divide a tragedy into temporally discrete segments. Thus, as Telford observes, the quantitative parts are parts of tragedy in a completely different sense than are the formal parts discussed in chapter 6. For acts, scenes, and episodes have a temporal duration, and one ceases when another begins. But the six formal parts are all parts of tragedy as a whole, for it is ridiculous to suppose that at a given point in a drama there is, for example, the end to a section devoted to plot and the beginning of a section containing diction or character. The formal parts can be imagined, if you will, as a vertical hierarchy of forms encompassing the entire drama. The quantitative parts will then be divisions in time along the horizontal sequence of incidents. However, Else's interpretation of "lexis" and "melopoia"

127 See above, p. 410.
128 Telford, p. 99.
directly violates the distinction between quantitative and qualitative parts; according to Else, the two terms are linked to segments of tragedy that, by virtue of their very definitions, cannot occur at the same time—the songs occur only when there is music, and the dialogue sections occur only in the absence of music. In other words, Else's interpretation turns lexis and melopoiia into quantitative rather than qualitative parts.

Else's interpretation would, moreover, face the same charge even if the contents of chapter 12 were written off as spurious. As Telford points out, Aristotle's discussions of the other four parts make it clear that they are not associated with temporally discrete segments. Even Else's discussions of the other four parts make that clear. In other words, although chapter 12 further clarifies the nature of the distinction between quantitative and qualitative parts, the distinction itself is presupposed by everything Aristotle says about the six parts of tragedy in chapter 6. Thus, Else's interpretation violates a distinction that is basic to Aristotle's entire conception of the six parts of tragedy. It is, moreover, important to note that since Else's interpretation rules out the possibility that the six parts constitute a hierarchy of levels, it would immediately rule out a Structuralist interpretation of them based on an analogy with the hierarchical levels of linguistic structure.
Perhaps it would be helpful to summarize briefly the preceding discussion of Else's first attempt to prove that "lexis" and "melopoiia" are tied to a distinction between dialogue and song. In his first argument, Else attempts to link the passage in which Aristotle introduces "lexis" and "melopoiia" and the passage in which Aristotle divides sweetened speech into dialogue and song. However, his primary evidence for the alleged link, the presence of the word "auten," which means "alone," falls far short of securing a clear connection between the two passages. More importantly, the very nature of the distinction between dialogue and song is completely foreign to Aristotle's discussion of the qualitative parts; regardless of whether chapter 12 is considered authentic, the distinction between dialogue and song does not fit in with Aristotle's discussion of the other four parts. Thus, Else's first argument must be considered a failure.

The plausibility of Else's second argument hinges on his interpretation of a remark in which Aristotle states that of the remaining parts, melopoiia is the greatest of the sweetenings, "hedysmaton."¹²⁹ Whereas other translators and commentators have assumed that the sweetenings being compared are melopoiia and opsis, Else argues that Aristotle

has "said nothing about ὁψὶς [opsis] being a ἁδυσμα [sweetening]; on the other hand he has said that ἁδυσμένος λόγος [sweetened speech] included the verse parts and the songs." \(^{130}\) If the remark does indeed represent a comparison of the parts of sweetened speech mentioned earlier, then, argues Else, "melopoia" must refer to the composition of both the music and the text of the songs; it could not be linked to a part or type of sweetened speech unless it also included the composition of the words.

The crucial problem with Else's interpretation is that it ignores the introductory phrase of the remark in question. Aristotle makes the remark after having completed his preliminary discussions of plot, character, thought, and lexis. \(^{131}\) "Of the remaining parts," he then goes on to say, melopoia is the greatest of the sweetenings. Now "the remaining parts" can only refer to melopoia and opsis—he has completed his discussion of all the others. Else's insistence that "the sweetenings" refers to melopoia and lexis turns the statement into nonsense: "of melopoia and opsis, melopoia is a greater sweetening than lexis." Surely it makes more sense to assume that "the remaining parts" and "the sweetenings" are coreferential; Aristotle is

\(^{130}\) Else, p. 277.

\(^{131}\) *Poetics*, 6: 1450a38-1450b15.
comparing the remaining parts—melopoia and opsis—to each other insofar as they are both sweetenings. It is true that Aristotle has not previously associated spectacle with the specific word "hedysma" (sweetening), but there is no reason that the two should not be associated. Aristotle is presumably using the term, which was primarily used in connection with various spices and sauces, to refer to the sensual aspects of tragedy, which clearly include spectacle. With the recognition that the remark in question has nothing to do with lexis, Else's second argument collapses. If the remark does not represent a comparison of melopoia and lexis, it cannot be linked to the earlier division of sweetened speech into song and dialogue and, therefore, cannot be used to establish a connection between melopoia and the text of the songs. Even if Else's second argument did not face the preceding problem, it would, however, face the charge that it turns the six parts into a confused mixture of quantitative and qualitative distinctions.

Before the present discussion of Else's interpretation is concluded, it may be worthwhile to consider an additional problem inherent in his reading. In his analysis of Aristotle's derivation of the six parts from the idea of tragedy qua drama, Else claims that lexis and melopoia are necessary because the dramatis personae must both speak and sing:
The drama is a presentation of men acting; and men 'act,' in Aristotle's sense, by speaking (and singing). Therefore the poet has to compose a text for them to speak and sing.  

The first part of Else's remark seems plausible; for Aristotle "action" requires rational deliberation, which, in turn, requires the use of language. Thus, if the action of a tragedy is to be carried out directly by the dramatis personae, they must have words to use, which means that lexis must necessarily be a part of tragedy. Furthermore, they must have some means of communicating the words; they could speak, or they could sing. However, there is no reason why they must both speak and sing. Thus, the division of tragedy into dialogue and song does not follow from the idea of tragedy qua drama the way the other parts do; it is simply not necessary that the dramatis personae both speak and sing in order to carry out the action themselves.  

Although the foregoing analyses of Else's two major attempts to support his interpretation of "lexis" and "melopoia" have raised a number of objections, two of those

132 Else, p. 235.
133 See below, pp. 443-44.
134 The notion that the six parts are necessary to tragedy insofar as it is a form of drama has been established above, see pp. 385-86.
objections stand out as crucial. First, his interpretation obliterates the essential distinction between quantitative and qualitative parts. Second, it destroys the hypothetical necessity involved in Aristotle's derivation of the six parts from the idea of tragedy qua drama. Even without the aid of the various other objections that have been raised, those two criticisms furnish enough evidence to refute Else's interpretation decisively.

The traditional interpretation of "lexis" and "melopoia," which assumes that the latter term refers to only the musical part of the songs, does not fare much better than Else's reading. The traditional view does have the advantage of not relegating lexis to discrete segments of a tragedy, but, like Else's reading, it nevertheless relegates melopoia to such divisions; at a certain point in a tragedy, the music begins--it lasts for a time--and then it stops. Later, there will be another stretch of music, and so on. Thus, the inappropriately quantitative character of melopoia is left intact. Furthermore, like Else's view, the traditional interpretation upsets the hypothetical necessity involved in Aristotle's derivation of the six parts from the idea of tragedy qua drama; there is no reason why the dramatis personae must sing and be

135 See above, p. 404.
provided with musical accompaniment in order to carry out the dramatic action. In addition, the traditional interpretation leaves the classification of a tragedy's spoken sound unclear. It is presumably supposed to fall under the rubric of *lexis*. That, however, leads to an extremely awkward picture of the distinction between *lexis* and *melopoia*. Whereas *lexis* includes the entire text of a tragedy—the portions that are spoken and those that are sung alike—it only includes some of the sounds produced when that text is delivered; it includes the spoken sounds, but the musical sounds belong to *melopoia*. Such an interpretation comes close to justifying Orsini's claim that the qualitative parts constitute the results of a taxonomic blunder.136

If neither Else nor the traditional view manages to advance a tenable interpretation of "*lexis*" and "*melopoia*," what could be left? The answer is quite simple: "*lexis*" refers to the entirety of a tragedy's language—both the spoken and the sung portions—insofar as that language is meaningful, and "*melopoia*" refers to all of the sounds made when that language is delivered—spoken and musical sounds alike. To put it another way, "*lexis*" refers to the

136 See above, pp. 352-53.
semantic side of a tragedy's language—to those aspects of language that remain constant regardless of whether the words are expressed in acoustic or orthographic form; the acoustic side of a tragedy's language belongs to "melopoia."

The preceding interpretation neatly manages to avoid the problems that plague other interpretations. For one thing, it manages not to upset the hypothetical necessity involved in Aristotle's derivation of the quantitative parts from the idea of tragedy qua drama. If the action is to be carried out by the dramatis personae themselves, they must deliberate, and rational deliberation requires the semantic resources of language. Thus, lexis must necessarily be a part of tragedy. Moreover, the agents must be able to give their thoughts overt expression, and since the only two means available to them—speaking and singing—both belong to melopoia, it also must necessarily be a part of tragedy. Second, the present interpretation restores to lexis and melopoia the thoroughly qualitative character they were meant to have. Neither occurs in discrete quantitative segments located here or there within a tragedy; like the other four parts, they are spread throughout the duration of a tragedy.

The present interpretation might also help to explain why Aristotle's extended treatment of lexis in chapters
19-22 contains virtually no discussion of the acoustic properties of poetic language. Butcher attributed the absence of substantive discussion to a general insensitivity to the acoustic aspects of language on Aristotle's part, suggesting that he came "perilously close" to advocating an approach to poetic language that "wholly ignores the effect of musical sound and looks only to the thought that is conveyed." However, the present interpretation of *lexis* and *melopoia* suggests that Aristotle may have concentrated primarily on the semantic side of poetic language during his treatment of *lexis*--not because he was insensitive to the musical qualities of language--but because he thought of *lexis* in semantic terms, the musical side of language being associated with *melopoia*. In fact, towards the end of chapter 6, *lexis* is even defined in terms that suggest it is associated primarily with the semantic side of language. "By diction," says Aristotle, "I mean, as was said before, the interpretation [of things] through language." Thus, "*lexis*" refers to the language of a tragedy *insofar as* that language constitutes a vehicle for the communication of meaning--

137 Butcher, p. 148.

An opponent of the preceding interpretation might well raise the following objection. "Melopoiia" is obviously based on the root word "melos," which serves as a sharp contrast to spoken dialogue in several of the passages contained in the Poetics. It is, therefore, unlikely that "melopoiia" would be used in such a way as to include the sound of spoken dialogue in its referential scope. The first part of the foregoing argument is certainly correct. As earlier remarks have indicated, Aristotle opposes the terms "melos" and "metron" in his division of sweetened speech, "By forms used separately I mean that some parts are achieved through meter only while others are achieved through melody."\(^{139}\) Certainly "melos" must there mean music as opposed to spoken dialogue. However, it is far from clear that the word should always be interpreted that way. In a passage from On the Soul, Aristotle observes that

```
Voice is a kind of sound characteristic of what has soul in it; nothing that is without soul utters voice, it being only by a metaphor that we speak of the voice of the flute or lyre or generally of what (being without soul) possesses the power of producing a
```

\(^{139}\) Telford translation, 6: 1449b30-1. See also above, p. 405.
succession of notes which differ in length and pitch [melos] and timbre. The metaphor is based on the fact that all these differences are found also in voice.\textsuperscript{140}

The context of the preceding passage makes it clear that Aristotle is talking about voice in general—not just the singing voice. Thus, "melos" must apply to the sound of spoken as well as sung language.

An opponent might, perhaps, argue that since the foregoing passage occurs in On the Soul, it is still possible that Aristotle intended the word "melos" to possess a more restricted range of application throughout the Poetics. However, the Poetics itself contains evidence to the contrary. Early in chapter 1, Aristotle lists a number of imitative arts which he says all utilize rhythm, speech, and harmony (harmonia [ʰαρμονία]).\textsuperscript{141} Later in the chapter, Aristotle observes that while some of those arts use only one or two of the preceding means of imitation, "Some arts use all of the things mentioned (by which I mean rhythm, melody [μέλεια], and meter)."\textsuperscript{142} The second list is clearly supposed to represent a repetition of the first one, which


\textsuperscript{141} Poetics, 1: 1447a21.

\textsuperscript{142} Telford translation, 1: 1447b24-5.
means that "speech" and "meter" are being used as synonyms, and likewise for "harmony" and "melody." Three chapters later, during a discussion of the history of tragedy, Aristotle remarks that iambic meter is more suited to tragedy than the dactylic hexameter used in epic because

the iambic is the most speakable of meters. There is a sign of this, for it is iambic we speak most in discourse with one another, while hexameters are seldom spoken even when we depart from the harmony [armonias] of common speech. 143

Thus, within the Poetics itself, "harmony," which can apparently stand in as a synonym for "melos," is explicitly applied to the sound of spoken language. Together the preceding passage and the passage from On the Soul constitute fairly strong evidence that, while the term "melos" is sometimes used to refer exclusively to the sound of musically intoned language during the course of the Poetics, it is, in other remarks, used to refer to the sound of spoken language as well—much as a person might today refer to the melodic or musical quality of a spoken passage. Given the preceding evidence, it seems possible that "melopoia" was intended to possess the wider of the two alternative readings of "melos." In view of the many problems

143 Telford translation, 4: 1449a25-8.
that have been shown to plague interpretations that assume melopoiia possesses a narrower scope, the foregoing suggestion not only seems possible but highly probable.

One potential problem remains. Twice during the later part of the Poetics, Aristotle says that the epic form lacks melopoiia. However, if melopoiia includes the sound of spoken as well as sung language, how could the recitation of an epic lack melopoiia? Before the preceding question is answered, it is worth noticing that, far from avoiding the preceding potential problem, interpretations that assume "melopoiia" referred to the sound of only sung language fare no better—or even make the problem worse. Although it is not known for certain if a rhapsode sang all of an epic during successive sittings, it does appear certain that at least substantial portions were sung—and accompanied by a lyre. Thus, if "melopoiia" referred to sung language, the epic would presumably have it as a part. Indeed, given the strong possibility that all of an epic text was sung, Else's view would seem to imply that it is lexis rather than


melopoia that epic lacks; if "lexis" refers to only those portions of a text that are spoken, and if the entire text of an epic was sung, then lexis ought to constitute no part of epic.

To return to the problem raised at the outset of the preceding paragraph, it may seem unclear how an epic could lack melopoia given that melopoia includes all of the sound made when the text of a tragedy is delivered. Interestingly enough, the most obvious solution closely parallels some observations made about spectacle during an earlier discussion. As that earlier discussion indicated, epic recitations were often accompanied by a good deal of visual reinforcement, but, unlike the performance of a tragedy, an epic recitation could not set separate visible manifestations of the various dramatis personae directly before the spectators, and that crucial difference is what marks all of the visual aspects of a tragic performance—and none of the visual aspects of epic recitation—as a part of opsis or spectacle.146 Much the same would appear to be true in the case of melopoia. Although a rhapsode could sing, and perhaps speak—could even sing or speak as if he were now one dramatic personage and now another—he could not place before his audience separate audible

146 See above, pp. 392-93.
manifestations of the various dramatis personae—he had no means of letting the separate dramatic personages sing or speak for themselves. A tragedy can, however, allow the agents to speak or sing for themselves, and that is why the sounds made when the language of a tragedy is delivered constitute part of melopoia, while the sounds made during an epic recitation do not.

The remarks contained in the preceding paragraph have suggested that spectacle and melopoia exhibit at least one interesting similarity, and it seems probable that the latter resembles the former in other ways as well. For example, although melopoia has been spoken of thus far exclusively in terms of vocal sound, and while that would appear to represent the only aspect of melopoia that is strictly necessary to tragedy qua drama, it is presumably not the only feature of tragedy that falls under the rubric of melopoia. The sound effects, the sounds made by the actors' bodies, and, of course, the instrumental music used during portions of a tragedy must be included under one of the six parts in order to render the list exhaustive, and melopoia represents the only possible choice. Thus, just as opsis includes certain aspects that are not strictly necessary to tragedy qua drama, so does melopoia.

Second, like spectacle, melopoia qua process falls partly in the hands of the poets and partly in the hands of
those who produce plays. Just as the poet cannot avoid specifying at least some of a tragedy's spectacle, e.g., the entrances and exits, he must also specify at least some of a tragedy's melopoia. He must, for example, specify which dramatic personage is to be speaking or singing at each point during the action. Naturally, each voice will possess its own distinctive sound; the male voices will, for example, sound lower than the female ones—even when men play the female parts. Thus, as the poet determines the order in which the agents speak and the length of each speech, he partly determines the play's music—its rhythms, its interplay of pitches, etc. In addition, the particular shape given to each line will partially determine the intonational pattern the actors must use. The preceding musical qualities are not, moreover, limited to performance situations; they are clearly in the written text. Like the visual manifestations of the dramatis personae, the aural manifestations are inseparable from the very idea of drama. Just as the dramatis personae must look one way or another, they must also sound one way or another, and that must necessarily affect the work of the poet. Furthermore, it must affect the way in which a person reads a tragedy—indeed, any actual performance; he must not only imagine that he sees the dramatis personae as they move about, he must also
hear them as they speak or sing.

While at least some aspects of melopoia must be incorporated into the text by the poet, the primary responsibility of executing melopoia must, however, fall to the actors, the musicians, and their fellow workers—just as the primary responsibility for executing spectacle falls to the skeuopoioi and his fellow workers. The poet cannot possibly specify every facet of a tragedy's melopoia; the people involved in producing a tragedy cannot avoid doing so.

The remarks contained in the foregoing paragraphs suggest that a set of distinctions exactly parallel to those drawn in an earlier discussion of spectacle can be made with reference to melopoia. First, the sense of "melopoia" that refers to part of the imitative product can be distinguished from the sense that refers to part of the imitative process. The former sense refers to the entire aural aspect of the product—to everything in the product that can be heard; the latter refers to the composition of the product's aural elements. Next, parallel distinctions can be made within each of the two preceding senses; melopoia as part

---

147 See the remarks on pages 401-03 for an analysis of spectacle that precisely parallels the following analysis of melopoia.
of the product can be separated into melopoiia as part of the written play and melopoiia as part of the performed play, and, likewise, melopoiia qua process can be separated into melopoiia as part of the poet's process and melopoiia as part of the production process.

Finally, those aspects of melopoiia that are necessary to tragedy qua drama can be distinguished from those that are not strictly necessary. In both the written and the performed product, the sounds the dramatis personae make as they deliver the play's text represent the most essential aspects of melopoiia; the nature of the product requires that at least some of the dramatic personages' aural manifestations, e.g., those determined by the alternation of speakers, be specified in performance and text alike, but either the text or a production may fail to include instrumental music or elaborate sound effects, etc. While the sounds made by the dramatis personae as they deliver a play's text represent the most essential aspects of melopoiia in both the written and performed product, the latter necessarily specifies those aspects more completely than does the former. Similar distinctions can be made in the case of melopoiia as part of the creative processes. The specification of the sounds the dramatis personae make as they deliver the text represents the most essential
aspect of melopoiia both for the poet and for those who produce the play, but the latter must necessarily specify those sounds more completely than the former.

It is questionable whether all of the foregoing distinctions can be applied to lexis. Lexis qua product can, of course, be distinguished from lexis qua process; the former sense of "lexis" refers to the words that make up the text of a tragedy insofar as they possess meaning, and the latter sense refers to the activities involved in composing a meaningful text. It is, however, questionable whether there would be any point to the distinction between lexis qua part of the written versus the performed play—or the parallel distinction between lexis qua part of the poet's process versus the performance process. The value of those distinctions will depend upon the sorts of identity conditions used in determining which performances and texts count as manifestations of the same play. If a performance and a written text count as manifestations of the same play only when they contain exactly the same words in exactly the same order—as some writers have argued, then the aforementioned

148 Insofar as the selection and arrangement of a play's words are concerned with matters of meaning, those activities fall under lexis; insofar as they are concerned with the acoustic qualities of the words, they fall under melopoiia.
distinctions will have little point.\textsuperscript{149} If both the performance and text contain the same words in the same order, the distinction between the \textit{lexis} in one as opposed to the other amounts to no more than the observation that in one case \textit{lexis} constitutes part of a performed product and in the other it constitutes part of a written product. Moreover, if those involved in the production process do not change the selection and arrangement of a play's words after it leaves the poet's hands, there is little point in trying to distinguish \textit{lexis qua} part of the poet's process from \textit{lexis qua} part of the production process; those concerned with the production process simply do not perform any activities that fall under the rubric of \textit{lexis}.

It is not, however, at all clear that the aforementioned conditions constitute the right ones for determining the identity of plays. It seems equally, if not more, plausible to claim that the effect of lexical changes on the identity of a play depends on how important the altered features are to the play's overall action, the motivation of the \textit{dramatis personae}, etc. (identifying all the factors that determine the relative importance of a lexical change

\textsuperscript{149} Such a theory of identity is, for example, advanced by Nelson Goodman in his book \textit{Languages of Art} (Indianapolis: Hackett Publishing Co. Inc., 1976), pp. 210-11.
would, of course, take a good deal of work). If something like the foregoing suggestion turned out to be correct, then there would be a point to the aforementioned distinctions. There would then be a substantive difference between the lexis of the written as opposed to the performed version of the same play, and the editorial activities performed by members of the production group would constitute a part of lexis qua process over and above the poet's activities.

Although Aristotle did not say anything in the Poetics to suggest that he took one or the other of the two preceding approaches to the identity of tragedies, there is reason for thinking that he may have recognized the problem. At the end of chapter 18 in the Poetics, Aristotle explicitly refers to the fact that choral odes were sometimes cut out of one tragedy and performed as part of a completely different play; he even seems to suggest that the actors were in the habit of doing the same with regard to sections of dialogue.150 What he would have said about the identity of the resultant performances is, however, unclear, and the present discussion will not attempt to settle the matter one way or the other.

Before the present discussion of "lexis" and "melopoia" is concluded, something should be said about the

150 Poetics, 18: 1456a28-32.
translation of the terms. While the English words "diction" and "melody" constitute problematic translations because they, like "spectacle," cannot normally be used to refer to processes, they have, nevertheless, become standard, and as long as they are thought of as technical terms capable of referring to parts of the poetic product and process alike, their use in subsequent discussion should cause no problem.

THOUGHT

Aristotle introduces the next two qualitative parts, thought (dianoia [\(\delta\iota\alpha\nu\omega\alpha\iota\alpha\)]) and character (ethos [\(\eta\theta\omega\alpha\)])}, together:

But tragedy is an imitation of actions acted by certain agents who of necessity are of a certain kind by virtue of character and thought, for it is because of these that we say that actions are of a certain kind, these two, character and thought, being the natural causes of actions, and it is by virtue of these that all agents happen to be fortunate or unfortunate.  

Else has rejected as spurious the portion of the preceding passage in which character and thought are said to be the natural causes of actions, or, as Bywater puts it, the

---

151 Telford trans., 6: 1449b37-1450a4. None of the major English translators, including Else, has objected to the translation of the Greek terms as "thought" and "character," so they will be utilized throughout the following discussion.
causes of actions "in the natural order of things."\textsuperscript{152} Since interpretation of the passage as a whole will depend in part upon the plausibility of Else's claim, it will be discussed first.

Else supports his claim with two arguments. First, he argues that if the clause in which character and thought are said to be the natural causes of actions were not spurious, "character and thought" would have to be construed as the antecedent of the phrase "\textit{kata tautas}" \([\kappa\alpha\tau\acute{\alpha} \tau\alpha\tau\varsigma]\)--"in accordance with" or "because of these"--in the last clause of the passage. That last clause would then assert that "it is because of character and thought that all agents happen to be fortunate or unfortunate." However, argues Else, the last clause must assert that all agents happen to be fortunate or unfortunate because of their actions--not by virtue of their character and thought.\textsuperscript{153}

The latter part of Else's argument is certainly correct. No more than twenty lines later, Aristotle explicitly states that is by virtue of their actions that agents are happy or miserable.\textsuperscript{154} If the preceding passage is to

\begin{itemize}
\item \textsuperscript{152} Else, p. 240.
\item \textsuperscript{153} Else, p. 240.
\item \textsuperscript{154} \textit{Poetics}, 6: 1450a20-1.
\end{itemize}
remain consistent with that statement, the antecedent of "kata tautas" must be "tas praxeis" [τὰς πρᾶξεις]—"the actions"—not "character and thought." However, it is not at all clear that the temporal intervention of the clause in question between "tas praxeis" and "kata tautas" obstructs the anaphoric dependence of the latter on the former. The claim made by the clause clearly constitutes a parenthetical observation embedded within the overall remark. Perhaps the reader is simply to bypass that parenthetical material in his search for the anaphoric antecedent of "kata tautas."

The passage might then be read as follows:

But tragedy is an imitation of actions acted by certain agents who of necessity are of a certain kind by virtue of character and thought, for it is because of these that we say that actions are of a certain kind (these two, character and thought, being the natural causes of actions), and it is by virtue of these [namely, the actions] that all agents happen to be fortunate or unfortunate.

Thus, the location of the clause in question does not brand it as obviously spurious as Else contends.

Else's second reason for considering the clause in question to be spurious is more directly related to the meaning of the passage as a whole. He argues that the doctrine in question, the doctrine that character and thought are, in the natural order of things, the causes of action, "is in fact not really Aristotelian. The two causes
of action, according to Aristotle (E. N. 6.2. 1139a31; De. An. 3.10. 433a10-30) 'are mind or reason . . . and desire.' The passages cited by Else in the preceding quotation do indeed state that reason and desire constitute causes of action. However, they also indicate that the foregoing is more or less another way of saying that thought and character are causes of action. In, for example, the passage from the *Nicomachean Ethics* cited by Else, Aristotle says that

> The origin of action—its efficient cause, not its final cause—is choice, and that of choice is desire and reasoning with a view to an end. This is why choice cannot exist either without reason and intellect [dianoias] or without a moral state [ethikes]; for good action and its opposite cannot exist without a combination of intellect [dianoia] and character [ethos].

Why is Aristotle's claim that desire and reason constitute the efficient causes of choice immediately followed by a statement that choice cannot exist without thought and character? The reason is that thought constitutes the natural manifestation of reason, and character constitutes one of the primary manifestations of desire— it manifests

---

155 Else, pp. 240-41.

that part of desire which has become habitual. The efficient causes of choice, which is, in turn, the efficient cause of action, may, therefore, be said to be either reason and desire or their more concrete manifestations, thought and character. Why are thought and character both necessary conditions for the occurrence of choice? Character is needed in order to provide a specific desire or preference, and thought is needed in order to connect that desire or preference with a rational plan of action. In other words, character is needed in order to determine which end is to be obtained, and thought is needed in order to calculate the various possible means of attaining that end. Together they result in a definite choice, which, in turn, results in a definite action.

Now that the Aristotelian credentials of the doctrine in question have been established, its bearing on the previously cited passage from chapter 6 of the Poetics may

157 Aristotle comments on the habitual nature of character in numerous passages—see, for example, Nicomachean Ethics, II: 1: 1103a14-8, where he even points out that the Greek word for character was derived from the word for habit. Subsequent remarks will suggest that the transitory parts of desire—the emotions—fall under the rubric of thought, and that they also constitute a necessary cause of action. However, the present discussion will concentrate on only the intellectual side of thought—along with character—as causes of choice and action—see below, pp. 456-57.
be examined. Aristotle begins the passage by observing that tragedy is an imitation of actions carried out by the agents who, he says, must necessarily be of a certain kind or quality by virtue of their character and thought. Why must they have the aforementioned qualities? The answer is suggested by the parenthetical observation embedded in the passage: character and thought constitute the natural causes of actions—they constitute necessary conditions for the occurrence of action. Since a tragedy is to imitate an action, and since that action is to be carried out by the dramatis personae themselves, those dramatic personages must necessarily possess character and thought; if they did not possess such attributes, there would be no action at all.

So far, discussion has suggested that the terms "thought" and "character" designate certain psychological attributes that the fictional personages in a tragedy must possess. More specifically, discussion has suggested that "character" is associated with a settled disposition to prefer some things over others, and "thought" is associated with the ability to calculate various means to be utilized in obtaining desired ends. Further consideration of "character" will be postponed until additional discussion of

158 See above, p. 440.
thought has taken place.

Aristotle's next remark about thought comes less than four lines after the passage that has just been discussed. "Thought," he says, "exists in whatever they say when demonstrating something or declaring some consideration."\(^{159}\) Near the end of the chapter he reiterates the preceding observation in a slightly amplified form, "Thought exists in those speeches in which the speakers demonstrate that something is or is not the case or declare something universal."\(^{160}\)

The preceding two passages further illuminate the nature of thought as a part of tragedy in several ways. First, they indicate the means utilized to convey thought. Analysis of Aristotle's first passage on thought has shown that the term "thought" designates a particular sort of mental ability that the dramatis personae must possess. How, it might be asked, is the poet to indicate that they possess the ability in question? Well, says Aristotle, the dramatis personae must be given speeches that serve as manifestations of that ability. So, for example, when a poet has placed an agent in a difficult situation, he may write a

\(^{159}\) Telford translation, 6: 1450a7-8.

\(^{160}\) Telford translation, 6: 1450b11-3.
speech in which the agent reasons about the various possible means of extricating himself from the difficulty. Such speeches serve to indicate that the *dramatis personae* possess means-ends rationality.

The second thing that the two passages show is that the ability in question is wider than was indicated by discussion of Aristotle's introduction of thought. That discussion only suggested that thought includes the ability to reason about the appropriate means to be used in attaining specified ends, but if thought is manifested whenever the *dramatis personae* demonstrate that something is or is not the case and declare something universal, the ability involved must be more extensive; it must be the ability involved in all forms of generalization and proof. It must, in other words, be the ability to think rationally.

Chapter 6 contains one additional passage on the subject of thought. Towards the later part of the chapter, Aristotle remarks that thought is being able to say what may be or is suitable and is that aspect of speeches which is the function of politics and rhetoric. For the ancient poets made their characters speak politically, the present ones make them speak rhetorically. 161

161 Telford translation, 6: 1450b5-8.
The first part of the passage more or less reiterates points that have already been made: "thought" refers both to the ability to discourse rationally, an ability which the *dramatis personae* must possess, and to those aspects of a tragedy's speeches that serve as manifestations of that ability. Aristotle then goes on to say that the business of manifesting thought through speeches properly belongs to the arts of politics and rhetoric, adding that the ancient poets made their dramatic personages speak politically, while the modern ones make them speak rhetorically.

What exactly is the difference between speaking politically and speaking rhetorically? According to Else, the two are "skewed" towards different points of view, the former "towards the 'ethical' or habitual, the unreflective, conditioned element in us, the other towards the reflective, self-conscious one: i.e., towards 'thought.'" The problem with Else's interpretation is that it turns one aspect of thought, that which involves "speaking politically," into character portrayal—it has to do with the "habitual," the "ethical" element in people. In effect, Else makes Aristotle say that only part of what falls under the rubric of

162 Else, p. 269.
thought, namely speaking rhetorically, really does fall under thought—the other part really falls under character. During an earlier discussion, Else has suggested that Aristotle "shows signs of wavering as to the delimitation between" thought and character, 163 but his interpretation goes far beyond mere wavering; if it were correct, Aristotle would be involved in a case of outright contradiction: some thought is not thought—it is character.

Luckily, there is an alternative to Else's problematic interpretation. Earlier discussion has suggested that thought is often concerned with the calculation of the various means that might be used to obtain a desired end.164 However, that end may, in turn, serve as a means to a further end, which may, in turn, serve as a means to yet another end, and so on until, as Aristotle observes in the Nicomachean Ethics, the series terminates in the ultimate end, which is happiness (eudaimonia [εὐδαιμονία]).165 Thus, the calculation entailed in means-ends reasoning may involve a lengthy chain of interlocking means to an ultimate end, a chain which might, perhaps, be divided into immediate

163 Else, p. 240.
164 See above, p. 444.
means/ends—those which come early in the chain—and posterior means/ends—those which come later in the chain. Naturally, a speech that manifests thought does not necessarily have to deal with the entire chain; some speeches discuss only the means to be used in achieving an immediate end. Such speeches will tend to concentrate on convincing listeners that a particular choice, when put into action, will be an effective means of obtaining the immediate end in question, and that is all. Other speeches may, however, be concerned with posterior as well as immediate means/ends, and those speeches will require discussion of additional matters like the question of whether a particular series of choices will eventually produce true happiness.

The preceding difference between types of speeches mirrors part of the difference in focus that distinguishes politics and rhetoric. In the second chapter of the Rhetoric, Aristotle defines rhetoric as "the faculty of observing in any given case the available means of persuasion."\[^{166}\] Thus, speaking rhetorically concentrates primarily on persuasion, and properly includes speeches in which the primary aim is to persuade listeners that a particular choice or action will be an effective means of

obtaining an immediate end. However, the art of politics is necessarily concerned with the attainment of true happiness, doing what is good, and so on.\footnote{Politics, trans. Benjamin Jowett, in The Basic Works of Aristotle, VII: 1-3: 1323a14-1325b32 and Nicomachean Ethics, I: 2: 1094a17-1094b11.} Hence, speaking politically will necessarily involve discussion of whether a series of means/ends will ultimately eventuate in the attainment of true happiness. In fact, in the Rhetoric Aristotle actually identifies "political speaking" as one of three main types of rhetoric, and goes on to point out repeatedly that the political speaker must concern himself with the connection between a potential plan of action and the attainment of happiness, the good, etc.\footnote{Rhetoric, I: 5-6: 1306b4-1363b4.}

If what has been said so far is on the right track, Aristotle's differentiation of the ancient from the modern poets is based not on the presence or absence of character development, as Else thinks, but on whether the speeches given to the *dramatis personae* tend to discuss the means needed for the attainment of more than just immediate ends. When faced with difficulties, the *dramatis personae* created by the ancient poets were made to concern themselves not only with the question of whether a potential series of
choices and actions would extricate them from their immediate circumstances, they were also made to concern themselves with whether those choices and actions would eventually lead to true happiness, which of course required some discussion of the nature of true happiness, etc. By way of contrast, the personages created by the modern poets were made to concern themselves only with the connection between potential actions and immediate ends. It is, of course, true that speaking "politically" tends to reveal a personage's settled preferences, i.e., his character, more fully than does speaking rhetorically, and insofar as Else's interpretation indicates that tendency, it is correct. Nevertheless, that tendency constitutes a side-effect of—rather than the basis for—the distinction between the two types of thought.

The Poetics contains one additional passage that bears directly on the nature of thought as a qualitative part of tragedy. Having engaged in a detailed discussion of plot and character in chapters 7-18, Aristotle returns to the subject of thought in chapter 19:

Let us assume what has been said about thought in the Rhetoric, for thought is more proper to that method. Whatever there may be in tragedy by virtue of thought, these are the things that ought to be rendered through speech. The parts of this are demonstrating,
solving, and rendering the passions (e.g., pity, fear, anger, and such), and moreover, maximizing and minimizing things.

The passage is important because it indicates that the scope of thought is even more extensive than previous passages had suggested it to be. The ability designated by the word "thought" not only includes the capacity to demonstrate and to solve—to find, for example, a means for obtaining a desired end—it also includes the capacity for maximizing and minimizing things and the capacity for rendering the passions.

What is involved in the capacity for maximizing and minimizing things is fairly apparent: it is simply the ability to make something look important or trivial, depending on the speaker's inclinations. The capacity for rendering the passions is, however, another matter. The Bywater translation speaks not of the "rendering" of passions or emotions, but of efforts to "arouse" emotion, and Butcher speaks of "the excitation of the feelings."


170 Where the Telford translation contains "solving," the Butcher and Else translations contain "refutation." However, the actual Greek term "luein" [λύειν] literally means "loosening," and it seems clear that Aristotle has in mind the "loosening" of a knotty problem in the dramatic situation by finding an appropriate plan for further action.
Furthermore, Else argues that a distinction must be made between the emotion a speaker feels as a natural response to his situation and the effect he deliberately tries to instill in his listeners, suggesting that only the latter properly falls under the rubric of thought:

the effect—as with the orator—has to be produced deliberately by him, in his speech... and through his speech alone...; whereas emotion is "revealed" (φαίνεσθαι), not "produced," and may be betrayed by action also.171

Even Telford construes "rendering the passions" in terms of "effecting another character's passions through speech."172

Thus, virtually all of the major commentators have assumed that thought includes the capacity to produce emotional effects in others but does not include a speaker's capacity to express his own emotions. That assumption appears to be based on Aristotle's twice repeated observation that the business of manifesting thought through speeches belongs to the art of rhetoric; Aristotle's discussion of emotion in the Rhetoric is exclusively devoted to the art of deliberately arousing emotions in others, and commentators have, therefore, tended to assume that Aristotle has the same sort

171 Else, p. 566, the emphasis is Else's.
172 Telford, p. 125.
of thing in mind when discussing thought in the *Poetics*.  

There are, however, grounds for disagreeing with the critical consensus about the matter. First, as earlier remarks have indicated, Aristotle states that the business of manifesting thought in speeches belongs to the art of politics as well as rhetoric, and there is no explicit indication that even those two disciplines completely exhaust the domain of thought as a part of tragedy. Second, during the course of chapter 17, Aristotle advises poets to

work out the incidents in gestures, for apart from nature itself those are most persuasive who themselves suffer the passions they are imitating. He who is himself distressed most truly distresses, and he who is angry most truly enrages.  

Aristotle seems to be saying that the poet will portray an emotion more effectively if he feels it himself, and the same would seem to apply where interactions among the *dramatis personae* are concerned; expression of a speaker's own emotional state will often constitute his most effective

---

173 See, for example, Else's remarks on pages 562-66.

174 See above, p. 447.

175 Telford translation, 17: 1455a31-33.
means of arousing the intended state in the other *dramatis personae*. Thus, in many cases a speaker's ability to express his own emotions and his ability to arouse emotions in others will amount to virtually the same thing; it would certainly be odd if the former ability belonged to thought while the latter did not.

Of much greater importance than the two preceding points is, however, the fact that the *dramatis personae* must possess emotions if they are to carry out the action themselves, and the fact that their emotional states must be conveyed primarily through the speeches. As earlier remarks have indicated, desire can be separated into two parts, the long-standing, habitual desires or preferences associated with character, and the more transitory desires commonly associated with the emotions. Those remarks also indicated that the habitual desires constitute a necessary condition for the occurrence of choice and action, but so do the more transitory desires. In the *Nicomachean Ethics* Aristotle says that

by states of character [I mean] the things in virtue of which we stand well or badly with reference to the passions, e.g., with reference to anger we stand badly if we feel it

176 See above, p. 444.
violently or too weakly, and well if we feel it moderately; and similarly with reference to the other passions.

The passage suggests that a character trait can be thought of as a habitual tendency to react in a particular way when an occasion provokes a certain kind of passion or emotional response. So, for example, a vindictive man possesses a tendency to react in a particular way when an occasion causes him to become angry; when angry, he tends to desire revenge. Thus, the "occasional" desire of anger triggers the "standing" desire associated with the character trait. Both are needed, along with the intellectual side of thought, for the occurrence of choice and action. The emotion is needed to bring the relevant character state into play, and together they determine a particular desire or preference, e.g., a desire for revenge; the intellectual side of thought is then required to connect that desire with a rational plan whereby the object of desire may be obtained. In combination, the three result in a definite choice, which can then be put into action.

Since emotion constitutes a necessary condition for the occurrence of choice and action, and, since the dramatis personae must carry out the action of tragedy themselves,

the agents must thus possess appropriate emotions at appropriate times. How is the poet to indicate that? His primary means will consist in giving the various dramatic personages speeches that serve as manifestations of whatever emotions are deemed appropriate. Even Else admits that "The emotions are often conveyed through speech," but he claims that the expression of those emotions cannot be included under the rubric of thought. However, thought represents the only one of the six parts that can serve as the primary locus for the expression of the agents' emotions; spectacle, music, and diction are obviously out of the question, the emotions are too transitory to be classified under character, and while a play's plot obviously includes the emotions expressed by the dramatis personae, as it includes everything else that is part of the play, it clearly will not do as the primary locus for those emotions. Since the agents must necessarily possess emotions, and since thought represents the only possible rubric under which the expression of those emotions can be classified, and since the phrase utilized by Aristotle, "to pathe paraskeuazein" [τὸ παθὲν παρασκευάζειν]—"the producing or rendering of emotions"—could be used in talking about the expression of a speaker's emotions as well as in talking about the arousal of emotions in listeners, there is good reason to assume that Aristotle included the capacity to express emotion as
Before further discussion of character begins, a brief summary of the conclusions reached in the preceding discussion and some previously unmentioned implications of those conclusions will be offered. The term "thought" can be used to refer either (1) to an ability that the *dramatis personae* in a tragedy must possess, or (2) to those aspects of a tragedy's speeches that serve as manifestations of that ability, or (3) to the process through which the poet endows his agents with that ability. When the term is used in the first of the preceding three senses, the ability in question includes the capacity to think and to feel emotions—especially the capacity to express those thoughts and emotions through speech. When used in the second of the preceding senses, the term refers to those aspects of a play's speeches that serve to manifest the foregoing capacities—namely those aspects of the speeches that indicate that the agents are busy proving something, solving something, increasing or diminishing the importance of something, or expressing or arousing emotion. When used in the last of the above senses, the term designates all of the activities that are involved in endowing the *dramatis personae* with the various capacities in question, e.g., the poet must decide whether an agent should at any given moment be occupied with proving something, solving a problem, altering the importance of
something, expressing or arousing an emotional response, or some combination of the foregoing—and he must then find the proper way of manifesting those activities through speech.

CHARACTER

Unlike the qualitative parts that have so far been discussed, virtually all commentators are in agreement about the key features of character, some of which have already been identified during the preceding discussion of thought. That discussion has, for example, suggested that character, like thought, is a psychological attribute that the *dramatis personae* of a tragedy must possess. More specifically, it is a settled disposition to prefer certain things over others when a situation occasions a particular type of emotional response.

There are, in addition to the passages from whence the preceding information was derived, three remarks in the *Poetics* that bear directly upon the nature of character. At the outset of chapter 2, Aristotle remarks that

Since imitators imitate agents, and these of necessity are either worthy or base (for differences between characters are nearly always consequent on this distinction alone, since it is in respect of badness or virtue that all men differ in their characters), the objects imitated are either better than we
are, worse than we are, or such as we are, just as with painters.\textsuperscript{178}

The heretofore undiscussed feature of character that is disclosed in the passage is its moral nature, a feature that is further illuminated in a passage from the \textit{Nicomachean Ethics}:

Now neither the virtues or the vices [i.e., the various states of character] are \textit{passions}, because we are not called good or bad on the ground of our passions, but are so called on the ground of our virtues and our vices, and because we are neither praised nor blamed for our passions (for the man who feels fear or anger is not praised, nor is the man who simply feels anger blamed, but the man who feels it in a certain way), but for our virtues and vices we \textit{are} praised or blamed.\textsuperscript{179}

There is, then, a difference between the occasional desires associated with the emotions and the standing desires associated with character states in addition to the fact that the former are transitory, while the latter are habitual and long-standing: moral praise or blame does not attach to the former, it \textit{does} attach to the latter. In fact, the moral quality of character is emphasized in Aristotle's definition of it in the early part of chapter 6; "characters," says Aristotle, "are the things by virtue of

\textsuperscript{178} Telford translation, 2: 1448a1-6.

\textsuperscript{179} \textit{Nicomachean Ethics}, II: 5: 1105b29-1106a2.
which we say the agents are of a certain kind [or quality]."  

The qualities Aristotle has in mind are presumably moral qualities—virtues and vices. Because of its connection with moral virtue and vice, Cooper sometimes refers to character as an agent's "moral bent," and Else refers to it as "moral orientation."  

The last passage in the Poetics that bears directly upon the nature of character occurs toward the end of chapter 6. Aristotle there remarks that character is that which reveals moral purpose [literally that which reveals "the choice" (proairesin—προαίρεσιν)], showing what kind of things a man chooses or avoids. Speeches, therefore, which do not make this manifest, or in which the speaker does not choose or avoid anything whatever, are not expressive of character.  

The passage suggests a number of important points. First,  

---

180 Telford translation, 6: 1450a5-7.  
181 Cooper, p. 34 and Else, p. 271.  
182 The preceding remark is, of course, not intended to imply that the Poetics contains no further remarks on the subject of character. To the contrary, the work contains a number of additional remarks; all of chapter 15 is, for example, devoted to a discussion of the goals that should be aimed at in character portrayal. The preceding remark is merely intended to imply that the Poetics contains only one more remark concerned with what sorts of things character encompasses.  
183 Butcher translation, 6: 1450b8-11.
it illuminates a heretofore unmentioned aspect of the relationship between character and choice. Earlier remarks have concentrated on the role played by character as an efficient cause of choice, but in the preceding passage Aristotle seems more interested in choice as a means of revealing character within the context of a tragedy. How is the poet to indicate that the *dramatis personae* in a tragedy possess characters? Well, since the primary function of character is to determine, in conjunction with thought, what sorts of choices an individual makes, the poet can arrange to have his dramatic personages make certain choices during the course of the play; those choices will serve as manifestations of the agents' characters. More specifically, the poet may give the various *dramatis personae* speeches in which a choice is made, speeches which may even go on to associate the particular choice being made with a general tendency or habit of choice. Such speeches will not only indicate that the *dramatis personae* possess characters, they will help to define precisely what kind of character each personage possesses. Thus, where a tragedy contains a choice, especially one which is discussed at some length, there is character; where nothing is chosen, where

184 See above, p. 444.
nothing is sought or avoided, character is not manifested.

Aristotle's suggestion that some speeches are not expressive of character is further illuminated by a passage from the *Rhetoric*. In book III, chapter 16 of that work, Aristotle discusses the inclusion of narrative sections in rhetorical speeches and makes the following remarks:

The narration should depict character; to which end you must know what makes it do so. One such thing is the indication of moral purpose; the quality of purpose indicated determines the quality of character depicted and is itself determined by the end pursued. Thus it is that mathematical discourses depict no character; they have nothing to do with moral purpose; for they represent nobody as pursuing any end. 185

The preceding passage not only indicates the sort of speeches that might fail to contribute to the delineation of character, it also reinforces the connection between character and desire or preference. Character is a sort of standing tendency to prefer or desire certain ends over others; where there is no desire to obtain an end, there is no character.

Before discussion of plot commences, the results of the foregoing remarks will be summarized and slightly extended.

Like the term "thought," "character" can be used to refer either (1) to a psychological attribute that the *dramatis personae* must possess, or (2) to those aspects of the speeches in a tragedy that serve to indicate that they possess that attribute, or (3) to the process through which the poet endows his agents with that attribute. When the term is used in the first of the preceding senses, it refers to a morally relevant, habitual disposition to prefer some things over others when circumstances occasion a particular kind of emotional response. When the term is used in the second of the preceding senses, it refers to those aspects of a tragedy's speeches that serve to delineate the agents' moral orientations—namely those facets of the speeches which clarify the nature of agents' choices and which indicate that the particular choices in question are part of a general tendency to choose certain things over others. When used in the last of the aforementioned senses, the term "character" designates all of the activities that are involved in endowing the *dramatis personae* with particular morally relevant, habitual preferences.

**PLOT**

Aristotle devotes an overwhelmingly large amount of the *Poetics* to discussion of various aspects of plot, including analyses of its component parts, the proper arrangement of
those parts, the various problems to be avoided in constructing plots, etc. The following discussion will, however, be concerned primarily with the task of identifying the basic characterizing features of plot, and will, therefore, have little to say about matters such as those just mentioned.

The first passage in the Poetics that bears directly upon the question of what the plot of a tragedy is occurs during the early part of chapter 6: "plot," says Aristotle, "is the imitation of action, for I mean by plot the synthesis of the incidents." That remark contains two important observations about plot. First, it states that plot is the imitation of action, the mimesis [μίμησις] of praxis [πρᾶξις]. A thorough explanation of the term "mimesis" (imitation) cannot be offered until Aristotle's doctrine of the four causes has been discussed, and will, therefore, have to be postponed. The approximate meaning that the term is intended to convey in the context of the preceding remark is, however, fairly clear. "Plot is the imitation of action" is roughly equivalent to saying that plot is that part of a tragedy that represents—that stands as the

186 Telford translation, 6: 1450a4-5.
187 See below, pp. 495-511.
embodiment of—an action. If the plot of a tragedy is to represent an action, it must, of course, exhibit the key features that an action exhibits. The next order of business must, therefore, be to ascertain the key features of praxis (action).

In a passage from the Politics Aristotle indicates that action may be mental and internal as well as physical and external:

Not that a life of action must necessarily have relation to others, as some persons think, nor are those ideas only to be regarded as practical which are pursued for the sake of practical results, but much more thoughts and contemplations which are independent and complete in themselves; since virtuous activity, and therefore a certain kind of action, is an end, and even in the case of external actions the directing mind is most truly said to act. . . . If this were otherwise, God and the universe, who have no external actions over and above their own energies, would be far enough from perfection. 188

In the Metaphysics Aristotle reiterates his suggestion that the most perfect being, the unmoved mover, is engaged in an exclusively internal, mental form of action; it is, he says, engaged in thinking, and "its thinking is a thinking on

Thus, action can, at least in some cases, occur without any external manifestation whatsoever. Can the opposite occur--can there be action that is purely external? The answer is decidedly negative. As earlier remarks have indicated, mind or reason constitutes a necessary condition of action, and the term "praxis" cannot, therefore, be applied in its full literal sense to activities that do not emanate from some sort of mental processes. Thus, there are two types of action that the plot of a tragedy could imitate, one being entirely internal and the other being partly internal and partly external.

Since the former sort of action could not possibly be carried out in the concrete way that is a necessary aspect of tragedy qua drama, the plot of a tragedy must, therefore, imitate the latter type. So, as Butcher observes, Aristotle's use of the term "action" within the context of the Poetics embraces not only the deeds, incidents, the situations, but also the mental processes, and the motives which underlie the outward events or which result from them. It is the compendious expression for all these forces working together towards a definite end.

189 Metaphysics, XII: 9: 1074b34.

190 See above, p. 444.

191 Butcher, p. 337.
If the plot of a tragedy is to represent such an action, it also must encompass both the outward doings and their internal sources.

The crucially important implication of the preceding remark is that the plot of a tragedy must constitute an all-inclusive culmination of the other five qualitative parts if it is to stand as a complete imitation of the appropriate kind of action. If it is to represent the external side of such an action, the plot must encompass spectacle, melody, and diction, and if it is to represent the internal side, it must encompass thought and character. Thus, plot is not something separate from the other five parts, it is the ultimate culmination of those parts in an overall structural unity.

A bit later in chapter 6 of the Poetics Aristotle makes another remark that bears upon the character of the action that is imitated by tragedy. "Tragedy," he says, "is essentially an imitation not of persons but of action and life, of happiness and misery."\textsuperscript{192} Thus, tragedy imitates the sort of action whose progress is charted in relation to happiness and misery. Aristotle did not, however, conceive of happiness (\textit{eudaimonia} [\varepsilon\delta\alpha\imath\omicron\nu\omicron\ia]) and its opposite as states of mind. In the \textit{Nicomachean Ethics} he defines

\textsuperscript{192} Bywater translation, 6: 1450a16–7.
happiness as "activity in accordance with virtue." On the other hand, happiness and its opposite are not entirely separable from the presence or absence of pleasurable states either. Several lines after his definition of happiness as a form of activity, Aristotle suggests that it is an activity invariably accompanied by pleasure. Hence, happiness may be conceived of as an activity in accordance with virtue that is accompanied by pleasurable feeling, and unhappiness or misery may be conceived of as an activity not in accordance with virtue that is accompanied by pain. Tragedy, then, is to imitate an action that moves between the poles of two very different sorts of activities. Since the plot must represent that action, it also must be concerned with such activities, and towards the end of chapter 7, Aristotle says as much. The plot, he says, must contain whatever "is likely or necessary in an arising succession of incidents which happen to change from misfortune to good fortune or from good fortune to misfortune." The preceding quotation tends to emphasize the pleasure and pain that attend the activities or doings

193 Nicomachean Ethics, X: 7: 1177a12.
195 Telford translation, 7: 1451a13-5.
contained in the plot rather than the activities themselves. Nevertheless, it may be safely assumed that the various pleasures and pains are indeed connected with specific activities.  

The foregoing remarks have suggested that the plot of a tragedy is to imitate a certain kind of action, one which moves between the poles of happiness and misery and which encompasses both the internal and external aspects of such activities. In order to imitate such an action, the plot of a tragedy must, therefore, also concern itself with the activities of happiness and misery, and it must encompass all of the remaining five qualitative parts. Although the preceding conclusions hardly exhaust the manifold implications of Aristotle's pregnant observation that plot is the imitation of action, they will suffice for the moment, and attention may now be focused on the latter part of the remark that began the present discussion.

Having said that plot is the imitation of action, Aristotle explains that by "plot" he means the synthesis of the incidents or doings (pragma [πράγματα]). That part of the remark suggests two main points. First, it reinforces the

196 In book X, chapter 4 of the Nicomachean Ethics Aristotle states that "without activity pleasure does not arise" at all--X: 4: 1175a21.
idea that plot must encompass all the other qualitative parts. In order to have an incident—a doing—there must be concrete beings who, in order to achieve various ends, interact in visible and audible ways. As the synthesis of the incidents, plot must obviously encompass the various facets of those doings. Thus, it must encompass spectacle, melody, diction, thought, and character. The second point to be made is closely related to the first one. Although plot encompasses the other five qualitative parts and could not exist without them, it, nevertheless, represents something more than just the sum total of those parts. The plot of a tragedy constitutes the synthesis of the incidents, the ultimate unification of them into an overall pattern or structure; since the various incidents include elements of spectacle, melody, diction, thought, and character, the plot must also constitute the ultimate synthesis of the other five qualitative parts. Thus, it is the overall structure or organization of plot that distinguishes it from a simple sum of the other qualitative parts.

One further aspect of plot warrants discussion. In his book *Tragedy and the Theory of Drama*, Elder Olson claims that "events may be shown on the stage which form no part of the plot."197 Olson does not explicitly claim that the

---

preceding remark represents Aristotle's view of plot. However, Else makes something very like the foregoing claim during his discussion of a passage contained in chapter 17 of the Poetics. In that passage Aristotle offers the following advise to would-be playwrights:

> Also the argument of the play [logous—λόγος], whether one is taking it 'ready-made' or is composing it himself, should be drafted in general terms, then expanded with episodes in the spirit just indicated.

In his commentary, Else asserts that Aristotle is, during the course of chapter 17, concerned with the following question: "what is the proper relation between the plot and the 'episodes' or added material?" Now the very way in which the preceding question is stated presupposes that the episodes are not part of the plot; Else is, in effect, saying "assuming that the episodes are not part of the plot, what relation should be established between the two?" Later on, Else appears to soften his position a bit, claiming only that an episode is "by definition ... not an integral part of the plot." It is not, however, clear that even Else's weaker claim is true.

198 Else translation, 17: 1455a 34-1455b2.
199 Else, p. 504.
200 Else, p. 511.
In a footnote to his claim that episodes do not constitute integral parts of the plot, Else refers the reader to the following remark contained in chapter 9, which is presumably supposed to substantiate his claim:

Of simple plots and actions, the episodic are the poorest. By 'episodic' plot I mean one in which the episodes follow each other in no probable or necessary order.201

Else has obviously assumed that the preceding remark constitutes proof that Aristotle intended the word "episode" (επεισοδίον [ἐπεισοδίον]) to designate only those portions of a tragedy that are "episodic," those portions that do not fit into the overall sequence as probable or necessary. In other words, Else's remarks seem to imply that there is a crucial difference between incidents (pragmata) and episodes (επεισοδία): the former are not episodic and are, therefore, part of the plot, but the latter are episodic and are, therefore, not part of the plot, or at least not an integral part of it. However, the preceding passage from chapter 9 does not clearly indicate that all episodes are in fact episodic. Aristotle says that by "episodic plot" he means one in which the episodes follow each other in no probable or necessary order, i.e., one in which the episodes

201 Else translation, 9: 1451b34-6.
are related in an episodic way—which presumably means that episodes can also be related to each other in a non-episodic way. Moreover, a number of passages in the Poetics indicate that Aristotle conceived of the terms "episode" and "incident" not as names for radically different parts of a tragedy, but as synonyms for the very same parts. Even if the foregoing were not the case, however, the remark in question clearly indicates that the episodes must be part of the plot; they are, according to the remark, the primary constituents of an episodic plot—if they were not part of that kind of plot, there would not be a plot of that type at all.

Not only does chapter 17 contain no indication that Aristotle conceived of the episodes—or any other of the events enacted during the course of a tragedy—as lying outside the scope of a tragedy's plot, there are no indications of that anywhere in the text of the Poetics. Thus, whatever else may be said about Olson's claim that the plot of a tragedy may fail to include some of the events enacted during the course of the play, it does not appear to represent Aristotle's position on the matter. Telford

202 See, in particular, the definition, in chapter 12, of episodes as those parts which fall between the choral odes—in other words, the incidents—12: 1452b21-2. See also 4: 1449a28 and 24: 1459b30-3.
appears to come much closer to capturing Aristotle's position, "As the 'synthesis of the incidents,' plot is the concrete unity of every 'occurrence' in the play. . . . To give the plot would mean for Aristotle to produce the play in its entirety."203

SUMMARY

The preceding comments mark the end of the present study's consideration of Aristotle's six qualitative parts of tragedy on an individual basis. In order to gain a complete, well-substantiated picture of each qualitative part, it has been necessary to discuss a number of rather complicated arguments that have been advanced by previous commentators, especially those advanced by Else. In view of the length and complexity of some of those discussions, it may be useful to conclude the present chapter by characterizing each of the six qualitative parts in broad, summary terms.

Spectacle and melody represent the most sensual of the six parts. Together, they account for the fact that a tragedy must be seen and heard, whether in actual performance or in a reader's imagination. The term

203 Telford, p. 85.
"spectacle" can be used to refer (1) to everything that can be seen during the course of a tragedy—whether in performance or in the mind's eye, or (2) to the process of composing everything that can be seen. Similarly, "melody" can be used to refer (1) to everything that can be heard during the course of a tragedy—whether in performance or in a reader's imagination, or (2) to the process of composing everything that can be heard. "Diction" can be used to refer to either (1) the language contained in a tragedy, or (2) to the process of composing—of selecting and arranging—that language. "Thought" can be used to refer (1) to the dramatic personages' ability to express their thoughts and emotions, or (2) to those aspects of a tragedy's speeches that serve to manifest that ability, or (3) to the various processes involved in endowing the **dramatis personae** with that ability. Similarly, the term "character" can be used in three ways: it can be used (1) to refer to the dramatic personages' moral dispositions, or (2) to those aspects of a tragedy's speeches that manifest those dispositions, or (3) to the various processes involved in endowing the **dramatis personae** with such dispositions. Finally, "plot" refers to either (1) the synthesis of a tragedy's constituent incidents—including all of the various elements contained in
those incidents, or to (2) the processes involved in achieving such a synthesis.
Chapter VI
Aristotle's Conception of Causality and the Qualitative Parts of Tragedy

At the outset of the preceding chapter, two quite different views of the interrelation between the six qualitative parts of tragedy were described. According to one of those views, loosely referred to as the "atomistic" view, the qualitative parts represent the results of second-rate taxonomy, the haphazard division of tragedy into an aggregate of minimally related fragments. According to the other view, what was called the "Chicago School" view, the qualitative parts represent the results of a highly systematic analysis of tragedy into a hierarchy of synthetically integrated elements. The remainder of the

---

1 See above, pp. 350-53.

2 See above, pp. 353-55. As remarks contained in the last chapter indicated, Chicago School critics have only claimed that the top four qualitative parts, viz., plot, character, thought, and diction, exhibit the aforementioned systematic interrelations; see above, p. 353-54. Later in the present chapter, the bottom two parts, melody and spectacle, will be discussed; see below, pp. 530-42. Until that time, however, all references to the qualitative parts are to be construed as references to only the top four parts.

479
chapter was devoted to the task of ascertaining accurate characterizations of each of the parts, in preparation for a discussion of their relations to each other. This chapter begins with a brief discussion of the origins of the atomistic view. That is followed by discussions of Aristotle's conception of causality and its relation to the qualitative parts of tragedy, Aristotle's theory of imitation, and, finally, the plausibility of an extended version of the Chicago School view.

R. S. Crane and Bernard Weinberg have suggested that what is in the present study being called the "atomistic" interpretation developed during the years between the dawn of the Renaissance and the eighteenth-century. According to Weinberg, when the Poetics first became a major source of interest to Renaissance scholars during the sixteenth-century, Robortello and the other early commentators were

---

acquainted with virtually no previous interpretations of the work. With no prior interpretations to aid them, the sixteenth-century scholars fell back on a fifteenth-century tradition of textual exegesis wherein a writer would comment on a text line by line without ever attempting to state the overall meaning of lengthy passages—and without ever attempting to identify the function of a given statement as a first principle, intermediary step, or conclusion of an extended process of deductive reasoning. Naturally, the result of such an atomistic interpretive method was likely to be an atomistic interpretation. As Weinberg observes, when such a method is utilized, "The structure of a given document disappears; the form of the forest is lost and only isolated trees remain for contemplation." According to Weinberg, the use of the foregoing exegetical method, combined with several other factors—a tendency to read Aristotle in the light of a heavily rhetorical tradition, a tendency to "adapt" Aristotle to contemporary tastes, and a desire to read Aristotle as "a kind of Ur-Horace"—led to highly distorted readings of various points made during the

course of the Poetics. By the time commentators began to attempt to formulate more comprehensive interpretations, the distorted readings contained in the earlier, piecemeal commentaries had already become the standard views on the various issues and were, therefore, incorporated into the mainstream interpretive tradition.

Crane claims that one particular feature of the readings produced by the aforementioned interpretive trend was what is in the present study being called the "atomistic" view of the qualitative parts. According to him, the same factors discussed by Weinberg led commentators to ignore the causal structure that originally bound the qualitative parts together:

Of this causal scheme only the general outlines survived in the doctrines of subsequent critics in the "Aristotelian" line. The distinction of the four parts [i.e., plot, character, thought, and diction] was retained and, along with it, the substance of the rules Aristotle had formulated for their handling; what disappeared was precisely their rationale which in the Poetics had justified not only the rules, but the discrimination, definition, and ordering of the parts themselves.

Thus, if Crane and Weinberg are correct, the interpretation of the qualitative parts espoused by Hornby and Orsini has

8 Crane, p. 618.
its roots in the sixteenth-century. Moreover, if Crane's claims about the causal interrelations between the parts are correct, the atomistic interpretation advanced by Hornby and Orsini constitutes a hopelessly inaccurate reconstruction of Aristotle's original conception of the matter. Thus, the relative accuracy of the two major interpretations would appear to depend largely on the degree to which Aristotle's views on causality are genuinely relevant to interpretation of his remarks about the qualitative parts of tragedy. Naturally, consideration of the preceding matter will require, as a preliminary step, an examination of Aristotle's theory of causality.

ARISTOTLE'S THEORY OF CAUSALITY

Book II, chapter 3 of the Physics contains what is probably the most famous of Aristotle's explanations of causality. He begins the discussion by identifying four

There are, in addition to the remarks contained in Physics, II: 3, three other famous passages in which Aristotle engages in extended discussions of causality: (1) his definition of the term "cause" in the lexicon of philosophical terms in the Metaphysics, which constitutes a slightly abridged version of the remarks in Physics, II: 3, (2) a discussion of causation as it is related to demonstrative reasoning in Posterior Analytics, II: 1, and (3) a historical survey of what Aristotle's predecessors had to say about the subject in Metaphysics, I: 3-7. The aforementioned passages do not, of course, represent Aristotle's only comments on the topic--the entire Aristotelian corpus is peppered with remarks on the subject; they do, however, represent the most famous of his extended discussions of causality.
different senses, or uses, of the word "cause," (aitia [αίτια]). First, he identifies what has come to be known as the "material" cause:

In one sense, then, (1) that out of which a thing comes to be and which persists, is called 'cause,' e.g. the bronze of the statue, the silver of the bowl, and the genera of which the bronze and the silver are species.

Second, Aristotle identifies what has come to be known as the "formal" cause:

In another sense (2) the form or the archetype, i.e. the statement of the essence, and its genera, are called 'causes' (e.g. of the octave the relation of 2:1, and generally number), and the parts in the definition.

Third, he identifies the "efficient" cause:

Again (3) the primary source of the change or coming to rest; e.g. the man who gave advise is a cause, the father is cause of the child, and generally what makes of what is made and what causes change of what is changed.

---

10 Physics, trans. R. P. Hardie and R. K. Gaye, in The Basic Works of Aristotle, II: 3: 194b24-6. The Greek expression standardly translated as "material cause," "ex hou" [ἐξ οὗ], literally means "the from which."

11 Physics, II: 3: 194b26-8. The Greek word "eidos" [εἶδος] literally means "form."

12 Physics, II: 3: 194b29-31. The Greek expression "dia ho" [διὰ ὧν] literally means "the by means of which."
Last, Aristotle identifies the "final" cause, "Again (4) in the sense of end or 'that for the sake of which' a thing is done, e.g. health is the cause of walking about."\(^{13}\)

As the preceding passages suggest, Aristotle's use of the word "cause" was a good deal more comprehensive than is normal in contemporary usage. Indeed, Sir David Ross observes, that "of Aristotle's four causes only two, the efficient and the final, answer to the natural meaning of 'cause' in English."\(^{14}\) Ross goes on, moreover, to identify the underlying source of the difference in usage:

We think of matter and form not as relative to an event which they cause but as static elements which analysis discovers in a complex thing. This is because we think of cause as that which is both necessary and sufficient to produce a certain effect. But for Aristotle none of the four causes is sufficient to produce an event; and speaking generally we may say that in his view all four are necessary for the production of any effect.\(^{15}\)

Thus, the four causes would appear to constitute a set of conditions that are disjointly necessary but only conjointly

\(^{13}\) **Physics**, II: 3: 194b31. Either of two different Greek expressions appear when Aristotle is discussing the final cause: "h\(\nu\) h\(\nu\)\(\nu\)\(\kappa\)a" [\(\delta\upsilon\ \varepsilon\upsilon\varepsilon\upsilon\kappa\alpha\)], which literally means "the for the sake of which," or "telos" [\(\tau\epsilon\lambda\omega\)\(\)\(\upsilon\)], which literally means "end" or "purpose."


\(^{15}\) Ross, p. 73.
sufficient for the coming-into-being of a particular entity or event.

Having briefly identified the four causes, Aristotle then proceeds to point out that they are often interrelated in rather complex ways. First, he discusses the interrelationship between the efficient and the final cause:

Some things cause each other reciprocally, e.g. hard work causes fitness and vice versa, but . . . not in the same way, but the one as end, the other as origin of change.\textsuperscript{16}

Thus, just as hard work is the efficient cause of fitness, fitness is the final cause of hard work. In fact, the relationship between the efficient and the final cause appears to be one of means to ends; the efficient cause represents the means whereby the final cause may be attained, and the final cause represents the end at which the efficient cause aims.\textsuperscript{17}

Several sentences later, Aristotle appears to suggest that the material and the formal causes also bear a reciprocal relation to each other:

\textsuperscript{16} As Ross observes, the interrelation between the efficient and final causes indicates that for Aristotle "mechanism and teleology are not mutually exclusive; where A mechanically necessitates B it may also be true that B teleologically necessitates A" (Ross, p. 72).

\textsuperscript{17} Physics, II: 3: 195a15-26.
The letters are the causes of syllables, the material of artificial products, fire, & etc., of bodies, the parts of the whole, and the premisses of the conclusion, in the sense of 'that from which.' Of these pairs the one set are causes in the sense of substratum, e.g. the parts, the other set in the sense of essence--the whole and the combination and the form.  

Thus, just as hard work is the efficient cause of fitness, which is, in turn, the final cause of hard work--the parts constitute the material cause of the whole, which, in turn, constitutes the formal cause of the parts. In fact, as Ross observes, "form" and "matter" are, for Aristotle, thoroughly correlative terms: a particular quantity of stuff constitutes matter only with reference to a particular form--a particular organizational structure--that is to be or has already been impressed upon it; and, likewise, a particular organizational structure counts as a form precisely because it is to be or has already been impressed on some particular quantity of stuff.  

Moreover, as Ross goes on to observe, the correlative character of matter and form implies that

---


19 Ross, pp. 71, 73.
the distinction of matter and form may be drawn at many different levels within the concrete thing. In the realm of art, [for instance,] iron, which is the finished product of the smelter, is matter for the founder.20

The paradigm cases where levels of matter and form can be distinguished are, of course, those which involve part/whole relations. Thus, to extend Aristotle's own example, while a syllable constitutes a form *vis a vis* the letters it contains, it constitutes matter *vis a vis* a word that contains it. It would appear, moreover, that in cases that involve part/whole relations, previous levels of structure or form will be preserved as subsequent levels are added; since the structural organizations or forms found on previous levels constitute the materials—the building blocks—out of which subsequent levels are made, they will persist as parts of those subsequent levels. Thus, for example, the internal structure of syllables is usually preserved as the syllables become parts of larger structures or forms like words, sentences, etc.—preserved because the syllabic structures constitute the very foundations upon which the larger structures are erected.

Towards the beginning of the preceding paragraph, it was observed that, like the efficient and final causes, the

20 Ross, p. 73.
material and formal causes are reciprocally related to each other, and it may be added that in the latter—like the former—case, the reciprocal relationship is of a means-ends character. Just as the efficient and the final causes are related as means to ends, so are the material and formal causes. In order to understand fully why that is so, it is necessary to realize that for Aristotle a thing's formal cause—the full realization of its potential form—may also constitute its purpose or final cause. The final cause or purpose of an acorn is, for instance, to become a full-grown, healthy tree—to realize completely its own potential form. Hence, as Aristotle says, "the form must [also] be the cause in the sense of 'that for the sake of which.'" Thus, the formal cause constitutes the end or goal towards the attainment of which the material cause is employed, and the material cause constitutes a means of attaining that end.

The remarks contained in the preceding paragraph have suggested that the final cause of a thing sometimes coincides with its formal cause—its organizational structure—e.g., the final cause or purpose of an acorn is

21 Physics, II: 8: 199a33. In fact, Aristotle maintains that in some cases the formal cause coincides with not only the final, but also with the efficient cause—see Physics, II: 7: 198a25-9.
to realize a particular organizational structure, viz., that of a full-grown oak tree. However, there is an additional way in which the formal and final causes of a thing sometimes coincide, a way which corresponds to an additional sense of formal cause. In what has been said thus far, the concept of form has been associated primarily with the notion of structural organization, but Aristotle also talks about form in terms of the various dispositions and capacities characteristic of a thing. For example, in On the Soul, Aristotle suggests that the form of an axe is cutting and the form of an eye is seeing; having the capacity to cut is the essence or form—is what it is to be—an axe; likewise, having the capacity to see is the form or essence—is what it is to be—an eye. However, cutting also constitutes the final cause of an axe—the purpose of an axe is to cut; and seeing constitutes the final cause of an eye—the purpose of an eye is to see. Thus, there would appear to be two senses of formal cause—formal cause qua structural organization and formal cause qua set of dispositions and capacities—and it would appear that the final cause of a thing might coincide with its formal cause in either or both of the preceding senses of formal cause.

There is, moreover, reason to think that when the final cause of a thing coincides with one of the two senses of formal cause, it will also coincide with the other. To see why that is so, it will be helpful to consider the examples of the acorn, axe, and eye. Earlier it was said that the purpose of an acorn is to realize a certain form *qua* organizational structure. That is not, however, its only goal or purpose; an acorn also strives to realize a certain form *qua* set of mature dispositions and capacities, e.g., reproduction. In fact, realization of the organizational structure constitutes a goal precisely because it constitutes a means of obtaining the further goal associated with the possession of such dispositions and capacities. Thus, form *qua* organizational structure and form *qua* set of dispositions and capacities are related as means to ends just as one goal or final cause constitutes the means of obtaining yet a further goal or final cause. The same would, moreover, appear to hold true in the case of the axe and eye; the form *qua* set of dispositions and capacities is, in each case, dependent on a particular organizational structure which makes that form possible. Thus, there are two different—but very closely related—senses of form, and when one way of describing the final cause of a thing coincides with the formal cause in one sense, there will be a second, more or less ultimate way of describing the final
cause that coincides with the other sense of formal cause.

During the latter portion of his discussion of causality in *Physics*, II: 3, Aristotle observes that the causes of a thing may be identified or described in a variety of more or less perspicuous ways. First of all, an individual may state the proximate cause of a thing, which, as Ross observes, will be commensurate with it, or he may state a more distant cause, some genus which includes the proximate cause, e.g., the cause of health may be said to be a doctor or, more generally, an expert.\(^{23}\) Second, an individual may pinpoint the specific cause of a thing, or he may state the cause in terms of a concomitant that is accidentally coextensive with the specific cause, e.g., "in one way 'Polyclitus,' in another 'sculptor' is the cause of a statue, because 'being Polyclitus' and 'sculptor' are incidentally conjoined."\(^{24}\) Moreover, Aristotle points out that a cause may be stated as a complex expression formed by combining the terms for both the specific and an accidental cause, e.g., the cause of a sculpture might be stated as "neither 'Polyclitus' nor 'sculptor' but 'Polyclitus, sculptor.'"\(^{25}\) Finally, Aristotle observes that all causes,

---

\(^{23}\) Ross, p. 72 and *Physics*, II: 3: 195a30-33.

\(^{24}\) *Physics*, II: 3: 195a33-36.

\(^{25}\) *Physics*, II: 3: 195b10-12.
whether they be proximate or generic, specific or accidental, may be "spoken of either as potential or actual; e.g. the cause of a house being built is either 'house-builder' or 'house-builder building.'"26 Aristotle then briefly summarizes what he has said:

> Cause means either what is particular or a genus, or an incidental attribute or a genus of that, and these either as a complex or each by itself; and all six either as actual or potential.27

Thus, each one of the four causes can be identified in a wide variety of ways.

While none of the aforementioned ways of describing a cause could be said, strictly speaking, to constitute an erroneous description, some ways of stating causes are, nevertheless, to be preferred in intellectual investigation:

> In investigating the cause of each thing it is always necessary to seek what is most precise (as also in other things): thus man builds because he is a builder, and a builder builds in virtue of his art of building. This last then is prior: and so generally.

Further, generic effects should be assigned to generic causes, particular effects to particular causes, e.g. statue to sculptor, this statue to this sculptor; and

---

26 Physics, II: 3: 195b4-6.

27 Physics, II: 3: 195b14-17.
powers are relative to possible effects, actually operating causes to things which are actually being effected. 28

Thus, the most proximate, most specific cause that is most appropriately related to its effect in terms of potency or actuality is to be preferred.

Now that a preliminary understanding of Aristotle's views about causality has been gained, the relevance of those views to his analysis of the qualitative parts of tragedy can be discussed. Commencement of that discussion will, however, be temporarily postponed to allow for some observations about the Aristotelian concept of mimesis (imitation). There has been so much disagreement about what Aristotle had in mind, and the concept figures so heavily in Aristotle's entire theory of tragedy, that a discussion of it would appear to constitute a necessary part of any consideration of the Poetics. The present seems an opportune time for such a discussion because, as the following remarks will argue, a grasp of Aristotle's views

28 Physics, II: 3: 195b22-28. When Aristotle observes that the art of building is "prior" to the other causes he presumably means prior in the order of nature; it may not necessarily be prior in the sense of being more readily apparent to an investigator. Aristotle makes the relevant distinction between "what is prior and better known in the order of being and what is prior and better known to man" in Posterior Analytics, trans. G. R. G. Mure, in The Basic Works of Aristotle, I: 2: 71b34-72a6.
about causality is essential to an adequate understanding of his conception of *mimesis*.

**ARISTOTLE'S THEORY OF IMITATION**

As Aristotle observes in the *Physics*, whereas natural objects contain their own source of genesis, artifactual objects do not; they come into being when an external source, an artisan, imposes on a particular quantity of matter a form that it would not naturally come to possess. A bronze statue comes into being, for example, not by nature, but because a sculptor imposes on a particular quantity of bronze a form or shape it would not otherwise have come to possess. Thus, one feature of imitative objects is that they are man-made rather than natural.

A second feature of imitative objects is that the object imitated constitutes the formal cause of the product. The artist strives to impress upon his raw material a form similar to the form possessed by the object he seeks to imitate; since that object or, more precisely, the form of that object determines the form of the imitative product, it constitutes the product's formal cause.

As the preceding remark suggests, there are several different ways in which a person might describe the object

---

29 *Physics*, II: 1: 192b8-32.
imitated--there are, in fact, as many as there are ways of stating the object's formal cause. For example, when identifying the object imitated by a certain statue, a person might pick out a particular woman; "This is a statue of that woman," the person might say. While not incorrect, the preceding way of describing the object imitated is not, however, the most perspicuous for it is not everything about the woman in question that constituted the object of imitation--only her visible form. Thus, a more perspicuous way of identifying the object of imitation would be to say, "This is a statue of the visible form of that woman." The difference between the two preceding ways of identifying the object of imitation can be precisely stated as follows: the latter statement identifies the specific formal cause of the statue--it is specifically the visible form of the woman that the artist set out to imitate; the former statement identifies an accidental formal cause--it is only because the various other attributes possessed by the woman happen to be conjoined with her visible form that she qua possessor of them all can be said to be the object the artist imitated.  

---

30 See above, p. 492. See also Physics, II: 3: 195a29-195b3.
Most commentators would agree with the remarks contained in the preceding paragraph. However, there are two central, closely related questions about Aristotle's theory of imitation that have sparked a good deal of disagreement: 1. must all artworks imitate actual objects or events, i.e., objects or events that exist outside the artist's imagination?, and 2. what exactly is the nature of the imitative relation, i.e., the relation between the imitative product and the imitated object?

Some writers, e.g., Nelson Goodman, believe that imitation theories are committed to what Goodman has called a "copy theory of representation"; to say that an artwork is an imitation automatically implies (1) that there exists some particular, actual object or event that the work imitates, and (2) that the work constitutes a "faithful copy" of that object's sensible form. Whether or not some

---


32 Goodman, pp. 9, 25. Goodman does not explicitly say that Aristotle in particular held the aforementioned view, but he presents the view as something to which all versions of imitation theory are committed and that presumably includes Aristotle; after all, Aristotle's name has always been the one most closely associated with imitation theory.
imitation theories are in fact committed to the foregoing claims, several of Aristotle's remarks make it quite clear that his theory is committed to neither claim. With regard to the first of the two aforementioned claims, Aristotle explicitly says, in chapter 9 of the Poetics, that most writers of comedies and some writers of tragedies had imitated incidents that never actually occurred, and he even appears to encourage poets to make-up the actions that they imitate.\footnote{Poetics, 9: 1451b12-33.} Thus, Aristotle did not believe that for every imitative product there must be an actual object or event imitated; there will, for example, be no actual event imitated when a poet makes-up the action portrayed in a tragedy.\footnote{In logical parlance, the preceding remark amounts to a statement that Aristotle's theory of imitation does not guarantee the success of the following existential generalization: if x is an imitation of y, then there exists an actual object or event y such that x is a copy it—Arthur Danto also makes the preceding point in The Transfiguration of the Commonplace (Cambridge: Harvard University Press, 1981), pp. 68-9.}

The second of the aforementioned claims appears to conflict with Aristotle's statement in the Politics that music provides the most vivid imitations of various passions, e.g., anger and gentleness, courage and
temperance. If it were the case that Aristotle thought artworks must faithfully copy the sensible form of objects or events, the preceding statement would have to mean, absurdly, that musical compositions sound exactly like people do when they are angry or courageous, etc. People in emotional states characteristically express their feelings in linguistic terms, but instrumental music is manifestly non-linguistic in character, so there can be no question of exact duplication of sensible form. In fact, it is hard to see how the relation between musical compositions and emotional states could be characterized in terms of similarity of sensible form at all—a point that is worth emphasizing because even those commentators who believe Aristotle could not possibly have thought of imitation in terms of slavish duplication often maintain that, according to Aristotle, it is the sensible form of objects or events that artworks imitate. Towards the end of the chapter in the Politics that contains the aforementioned passage, Aristotle remarks that

36 See, for example, Butcher, p. 127 and Telford, p. 65.
There seems to be in us a sort of affinity to musical modes and rhythms, which makes some philosophers say that the soul is a tuning, others, that it possesses tuning.\textsuperscript{37}

The remark seems to suggest that there is some sort of resemblance—an affinity—between musical compositions and states of the soul like emotional states, and the resemblance or similarity is presumably one of form; but it is clearly not a resemblance of one external, sensible form to another. Thus, while many imitative products may resemble the external, sensible form of the objects they imitate, not all imitative products need be so related to their objects.

Since Aristotle explicitly denies that every tragedy must imitate some particular, actual object or event, it might seem that imitation should be characterized in terms of a relation between the imitative product and imagined objects or events, and several writers, e.g., Ross, come close to suggesting such an interpretation.\textsuperscript{38} Such a view appears to possess a certain amount of plausibility. Even in the case where, for example, a sculptor has a model

\textsuperscript{37} \textit{Politics}, VIII: 5: 1340b17-9.

\textsuperscript{38} Ross, p. 278.
sitting directly in front of him, it makes sense to say that he is imitating not the model \textit{per se} or even the model's visible form as such, but rather his conception of the model's visible form. There is, moreover, textual evidence that seems to support such an interpretation. In the \textit{Metaphysics}, for example, Aristotle observes that "from art proceed the things of which the form is in the soul."\textsuperscript{39} Perhaps, then, the most precise way of describing the formal cause—the object imitated—is to say that it is a concept or image in the artist's soul, which may or may not have been drawn from a single, particular object or event in real life. While it may be true that the foregoing constitutes the most precise way of describing the formal cause of an imitative product, and while Aristotle may have considered the relation between the form of an imitative product and the form of a corresponding image or concept in the artist's soul to be in a sense imitative, it does not seem likely that Aristotle thought imitation should be characterized \textit{primarily} in terms of that relation. As Else observes, "\textit{mimesis}" was traditionally associated with a relation between the product made and "pre-existant things," and while Aristotle was apparently trying to broaden the

\textsuperscript{39} \textit{Metaphysics}, VII: 7: 1032a35.
traditional concept of imitation, it is extremely unlikely that he intended to completely sever the connection between imitative product and things or events in the actual world; ancient Greeks would have regarded such a conception as having missed the essence of imitation entirely.  

The problems inherent in attempts to characterize imitation in terms of a relation between a particular product and a particular object or event, whether actual or imaginary, have led some commentators, e.g., Butcher, to characterize imitation in terms of a relation between a particular product and a universal, i.e., a type or kind of object or event found in real life. However, it is not clear that such a view succeeds in escaping the sorts of problems just discussed. The view implies that for each individual imitative product there is a type of object or event found in real life that it imitates. As it stands, the preceding claim is, however, too vague because the scope of the imitated types has not been specified; are they

40 Else, p. 322.

41 Butcher, pp. 153-54. Unfortunately, Butcher also insists that art, "in imitating the universal imitates the ideal" (p. 153), a view that assumes Aristotle saw the universal and the perfect as one and the same—which is extremely questionable. The present references to Butcher's views concern only his claim that art imitates the universal, not his additional claim that the universal is also the ideal.
to be extremely broad types, e.g., action in general, or narrower ones, e.g., the unbridled pursuit of revenge? The examples used by most of the commentators who have advanced the present view indicate that they have fairly narrow types in mind.\textsuperscript{42} It would appear, moreover, that the types must at least be narrower than, for example, action in general if the characterization of imitation in terms of a relation between a type of object or event and each particular work is to have any real point; if the type in question is so broad that all products of a given sort, e.g., all tragedies, can be said to imitate the same type of object or event, why identify one of the \textit{relata} in the imitative relation in particular rather than generic terms? Thus, it would appear that the aforementioned view is committed to a correspondence between particular products and fairly narrow types of objects or events found in real life. The problem is that it is not clearly the case that for each tragedy there is a particular type of action found in real life that it imitates. Consider, for example, the type of action imitated by Euripedes' \textit{Bacchae}. The action of that play is so dependent on the existence of supernatural personages and events that it would be difficult to describe it

\begin{footnote}
\textsuperscript{42} See, for example, Butcher, pp. 359-60, 363-66.
\end{footnote}
adequately as a type without some mention of its supernatur­
al elements, yet such actions do not exist in real life.
Thus, the view of imitation advanced by critics like Butcher
fails for much the same reason that the copy theory fails;
just as it is not the case that for every particular work of
art there exists in real life an individual object or event
that the work imitates, it is not clear that for every
particular work of art there exists in real life a specific
type of object or event that the work imitates.43

The preceding remarks suggest that the concept of
 imitation is caught in a seemingly insoluble puzzle.
Imitation apparently cannot be characterized in terms of a
relation between the imitative product and actual objects or
events; both (1) attempts to characterize it in terms of a
relation between a particular product and an individual
actual object or event and (2) attempts to characterize it
in terms of a relation between a particular product and a
type of actual object or event fail because for any
particular product there may be no individual object or
event and no specific type of object or event in real life
that corresponds to it. However, imitation cannot be

43 Subsequent remarks will identify an additional, even
more serious flaw in the account of imitation advanced by
writers like Butcher—see below, p. 506.
characterized primarily in terms of a relation between
imitative product and imaginary objects or events, whether
they be particulars or types, either because ancient Greeks
would not have considered such a relation to be a case of
imitation at all.

The way out of the preceding puzzle becomes apparent
only when the concept of imitation is seen in the light of
what Aristotle says about the various ways in which the
causes of a thing can be stated. Earlier remarks have cited
a passage in which Aristotle observes that each of the four
causes of a thing can be stated as actual or as potential,
in proximate or generic terms, and in specific or accidental
terms. Later in the same passage, Aristotle observes that
the effect of a cause can be stated in a similar variety of
ways, e.g., a person might identify the cause of "this
statue" or of "statue" or of "image" generally." However,
as earlier remarks indicated, Aristotle goes on to say that
"generic effects should be assigned to generic causes,
particular effects to particular causes, e.g. statue to
sculptor, this statue to this sculptor." The preceding
example concerns the relation between efficient cause and

44 See above, pp. 493.
45 Physics, II: 3: 195b8.
46 See above, p. 493.
effect, but much the same would appear to hold true in the case of formal cause and effect, e.g., the formal cause of statue is shape—the cause of this statue is this shape. Likewise, the formal cause of tragedy is action—the cause of this tragedy is this action.

It is easy to see how previous accounts of Aristotle's concept of imitation can be classified on the basis of the preceding distinction. Attempts to characterize imitation in terms of a relation between a particular product and an individual object or event, whether actual or imaginary, constitute attempts to characterize imitation on the individual rather than the generic level—and they do not work. Attempts to characterize imitation as a relation between a particular product and a type of object or event not only fail, they fly in the face of Aristotle's assertion that generic effects should be assigned to generic causes, particular effects to particular causes; they attempt to characterize imitation in terms of a relation between specific effects, individual imitative products, and generic causes, types of objects or events found in real life. It is, needless to say, very unlikely that Aristotle would have characterized imitation in terms that he explicitly rejects as methodologically illegitimate. There is, however, one possibility left. Perhaps both sides of the imitative relation ought to be characterized in generic terms. In
other words, perhaps imitation ought to be characterized in terms of the relation between types of product and types of actually existent objects or events. Such a view would entail that for every type of imitative product there exists a type of object or event found in real life that it imitates, which would, in turn, entail that if \( x \) is an imitation, then \( x \) belongs to a type of product that imitates a type of object or event found in real life. However, the generic correspondence between product and actually existent object does not guarantee that there will be a similar correspondence on the individual level. Thus, such a view is not committed to the claim that for every individual imitative product there is a particular object or event in real life that it imitates; nor would it be committed to the claim that every imitative product directly imitates some specific type of object or event in real life. The present account insures a connection between every imitative product and a type of object or event in real life, but the connection occurs indirectly rather than directly; each individual imitative product can be said to imitate indirectly a type of object or event in real life because it belongs to a type of product, and that type directly imitates a type of object or event in real life. So, for example, the statement that a particular tragedy is an imitation of an action necessarily implies only that the work in question belongs to the
artistic kind known as tragedy and that there is some kind of object or event found in real life, viz., action, that tragedy qua kind of product imitates. Of course, it might also be the case that there is a less general, type of action—or even a specific, individual action—found in real life that the tragedy resembles, but that need not be the case. Thus, the present account insures the required connection between imitative products and actual objects or events without entailing the problematic commitments that accompany previous accounts.

If the preceding account is correct, then there must be some sort of correspondence between tragedy qua kind of artistic product and action qua kind of occurrence found in real life. What exactly might that mean? First of all, it may be assumed that the correspondence in question is one of form; to say that tragedy imitates action is to say both that there is a particular form that all actions possess and that all tragedies exhibit a similar form. The next step would be to ascertain the form of action as it exists in real life, a task which has already been accomplished during the course of a discussion in the preceding chapter. As remarks contained in that discussion indicated, "praxis" refers to a series of activities that exhibit the following formal features: the activities are characterizable as happy or the reverse, and they are both internal
and external in form. Thus, the form of action \textit{qua} type of occurrence found in real life can be characterized in terms of the pattern or organization impressed upon the series of activities that make-up its matter; it is a pattern that encompasses both the internal and external and that moves between the poles of happiness and misery.

To say that tragedy imitates action is, then, to say that tragedy \textit{qua} kind of product exhibits a form similar to that just described—and, in fact, it does. First of all, since tragedy encompasses all six qualitative parts, it, like action in real life, will encompass both the internal and external. Moreover, since inclusion of all six parts necessitates the interaction of dramatic personages who possess hopes and desires and who experience emotional states as they try to attain the objects of their hopes and desires, the fact that tragedy encompasses all six qualitative parts also guarantees that, like action in real life, the activities it portrays will move between the poles of happiness and misery. Thus, the correspondence between the form of tragedy \textit{qua} kind of imitative product and the

\footnote{47 See above, pp. 467-71. To be precise, it is not really all action, but only a type of action that tragedy imitates. As the remarks contained on the aforementioned pages indicated, besides the form of action being described there is a form of action that is entirely internal. For the sake of simplicity, the distinction between the two types of action is being temporarily ignored.}
form of action qua kind of occurrence found in real life is insured by the fact that tragedy qua kind of product includes all six of the qualitative parts. \(^{48}\) It is, incidentally, worth noting, in light of what was said earlier about the relation between music and its objects of imitation, that the aforementioned similarity between the form of tragedy and the form of action could not be characterized exclusively in terms of external, sensible resemblance.

Before discussion of the relation between Aristotle's views about causality and the interrelation between the six qualitative parts begins, it may be useful to review briefly the conclusions reached during the preceding discussion, while at the same time pointing out some implications of those conclusions that have not yet been made explicit. "Mimesis" refers primarily to a relation between types of product and types of actual objects or events, a relation that consists in a resemblance of form. There may also be derivative or secondary senses of "mimesis." For example, Aristotle probably regarded the

---

\(^{48}\) The preceding remark is only intended to imply that inclusion of all six parts is sufficient to insure that a kind of product constitutes an imitation of action; given the other kinds of product that Aristotle says imitate action, e.g., the epic, it can be assumed that he did not regard the inclusion of all six as a necessary condition for the imitation of action.
correspondence between a particular product and an individual object or event, or a type of object or event, found in real life—when there is one—as imitative, and he may also have regarded the relation between a particular product and an image or concept of it possessed by the artist as imitative. In all of the preceding, secondary senses of "mimesis," the relation involved is the same type of relation that serves to characterize the primary sense of the term—viz., similarity of form. It is, however, the similarity between the form of a type of artistic product and the form of a type of object or event in real life that is primarily responsible for the imitative quality of certain artifactual objects.

THE CAUSAL INTERRELATIONS AMONG THE QUALITATIVE PARTS

An opponent of the Chicago School interpretation might well ask what evidence can be provided in support of the claim that Aristotle's views about causality are genuinely relevant to his discussions of the qualitative parts of tragedy. The following remarks will attempt to substantiate the Chicago School claim. The remarks begin by arguing for the broader claim that it is—in general—prudent to interpret the Poetics in light of the content of Aristotle's other works. The discussion then argues that of the various views expressed in his other works, Aristotle's views about
causality are—in particular—relevant to interpretation of substantial portions of the Poetics. Finally, attention is focused on the specific claim that Aristotle's discussions of the qualitative parts of tragedy are among those portions of the Poetics that ought to be interpreted in light of his views about causality.

As Grene points out, Aristotle's works are filled with cross-references and are often related to each other in strikingly systematic ways. Given the highly integrated nature of the Aristotelian corpus, it would appear that interpretation of the views contained in one work in light of the views contained in his other works is, in general, a sound methodological strategy. The foregoing would, moreover, appear to be especially true in the case of the Poetics. The Poetics is, as Else puts it, "an extraordinarily crabbed, difficult piece of writing." Indeed, it is so crabbed and difficult that most scholars assume it represents not a finished work intended for general circulation, but a set of Aristotle's lecture notes or even a set of notes recorded by a student who attended his lectures. In view of the fact that a fair number of

49 Grene, p. 16.
50 Else, p. vii.
51 Cooper, pp. 4-5.
the work's passages are virtually uninterpretable on their own, interpretation of the Poetics in light of the views developed in Aristotle's other works would appear to be unavoidable. As Butcher says, "Aristotle's Poetics cannot be read apart from his other writings."  

The remarks contained in the preceding paragraph merely make the general point that the Poetics should be interpreted in light of the views developed in Aristotle's other works—without indicating which of those views might be especially relevant. Why, it might be asked, should his views about causality be considered particularly relevant? As earlier remarks suggested, the four causes constitute a set of conditions that are disjointly necessary and conjointly sufficient for the coming-into-being of a particular entity or event. As such, the four causes represent the primary avenues to be followed when attempting to explain why a particular entity or event is the way it is:

The number [of the causes] is the same as that of the things comprehended under the question 'why.' The 'why' is referred ultimately either (1) in things which do not involve motion [i.e., to the formal cause], e.g. in mathematics to the 'what' (to the

52 Butcher. p. 114.

53 See above, pp. 485–86.
definition of 'straight line' or 'commensurable,' &c.), or (2) to what initiated a motion [i.e., to the efficient cause], e.g. 'why did they go to war?--because there had been a raid'; or (3) we are inquiring 'for the sake of what?'--'that they may rule' [i.e., we are inquiring about the final cause]; or (4), in the case of things that come into being, we are looking for the matter [i.e., we are looking for the material cause].

Furthermore, since the four causes constitute a sufficient condition for the coming-into-being of an entity or event only when conjoined, the explanation of why an entity or event is the way it is will not be complete until the "why-question" has been answered in terms of all four causes. Aristotle makes it clear, moreover, that a complete causal analysis is no less necessary where explanation of an artificial, i.e., man-made, object is concerned than it is where explanation of a natural being is at stake. In fact, the four major, extended discussions of causality contain as many--if not more--examples drawn from the province of art as from the various other realms of inquiry distinguished by Aristotle; many of the examples are even drawn from what are today distinguished as the fine arts.

54 Physics, II: 7: 198a15-20.
55 Physics, II: 7: 198a22-5.
56 Physics, II: 2: 194a21-194b8.
It appears, therefore, extremely unlikely that so methodical a philosopher as Aristotle, having in his other works stressed the importance of causal analysis in the explanation of artistic products and having provided numerous examples of such analysis, would have failed to utilize systematically the procedure of causal explanation in what may have been his only full treatise concerning a productive art; the Poetics is, at any rate, Aristotle's only extant work devoted primarily to the study of a productive art.

Another reason for thinking that Aristotle's views about causality are generally relevant to the content of the Poetics is that the work as a whole comes out looking like a more coherent, tightly reasoned and internally consistent document under a heavily causal interpretation than it does under minimally or non-causally oriented interpretations. Limitations of the present study prohibit an enumeration of the many ways in which the former sort of interpretation increases the apparent intelligibility of the Poetics as a whole, but inclusion of such a list is actually unnecessary.

57 See fn. # 9 on p. 483 for a list of the major discussions of causality. Being defined as "a state of capacity to make, involving a true course of reasoning," Aristotle's use of the term "art" (techne [téχνη]) includes both the useful arts, e.g., medicine, carpentry, etc., and the fine arts—Nicomachean Ethics, VI: 4: 1040a9-11.
Even a cursory comparison of, for example, Telford's reading of the *Poetics* and the interpretations provided by Hornby and Orsini will provide more than enough evidence that the preceding statement is in fact true. Telford's commentary provides a particularly good example because it is by far the most heavily causal reading to date; Telford explicitly says that he intends to identify not only "the causes of tragedy, but the causes of those causes, and so on" until "at last the argument is grounded in an ultimate cause or principle which is self-caused or reflexive."58

Although the preceding paragraphs have indicated that there is fairly strong evidence that Aristotle's views about causality are generally relevant to the content of the *Poetics*, an opponent might still argue that he has not yet seen compelling evidence. After all, he might say, it is possible that Aristotle did not always follow his own methodological advice. Nowhere in the *Poetics* does Aristotle explicitly say that his views constitute the results of causal analysis. Moreover, Aristotle's customary word for cause, "aitia," occurs only eight times

58 Telford, pp. 66-7. Incidentally, Aristotle himself claims that the causes of causes can and should be traced back until an ultimate cause is identified (*Posterior Analytics*, I: 24: 85b27-86a3).
during the entire course of the work. A causal interpretation is therefore justified, he might insist, only in the case of those eight passages. Attempts to interpret additional passages in causally oriented terms represent cases of reading more into a text than is strictly warranted.

Luckily, the text of the Poetics contains evidence that Aristotle's use of causal analysis cannot possibly have been limited to the eight passages that contain occurrences of the term "aitia." Aristotle's closing summary of what he has accomplished during the course of the Poetics runs as follows:

So much for Tragedy and Epic poetry—for these two arts in general and their species; the number and nature of their constituent parts; the causes [aitia] of successes and failures in them; the Objections of the critics, and the Solutions in answer to them.

Aristotle's statement that he has identified "the causes of successes and failure" in tragedy and epic poetry clearly suggests the use of causal analysis, but the identification of those causes cannot be argued to coincide with the seven other passages in which the term "aitia" occurs; of those

---

59 The term occurs in lines 1448b5, b12, b19, 1450a1, 1451b16, 1452a8, 1455b7, & 1462b18.

60 Bywater translation, 26: 1462b16-20.
seven passages, only one concerns a cause of artistic excellence or lack thereof in poetic works. Thus, although Aristotle's other discussions of the causes of artistic success do not explicitly contain the term "aitia," it is clear that they nevertheless constitute the result of a causal analysis.

The remarks contained in the preceding paragraph have furnished evidence that Aristotle did not always bother to point out the causal underpinnings of the views he expresses in the Poetics. When combined with the fact that he says in other works that causal analysis should be utilized in all fields of inquiry and the fact that a heavily-causal interpretation increases the intelligibility of the Poetics as a whole, those remarks furnish compelling evidence that Aristotle's views about causality are in fact relevant to substantial portions of the Poetics. However, none of what has been said thus far furnishes evidence that Aristotle's discussions of the qualitative parts should—particularly—be interpreted in causally oriented terms. An opponent of the Chicago School view of the qualitative parts could admit that Aristotle's views about causality are

---

61 The passage in question, located at line 1452a8, suggests that the degree to which the incidents in a tragedy are causally related to each other constitutes a cause of the work's success or failure.
relevant to interpretation of much of what is contained in
the Poetics, e.g., he could admit that the four main clauses
in the definition of tragedy serve to identify the four
causes of tragedy's nature, and still insist that there is
no connection between the aforementioned views and those
passages in which the qualitative parts are discussed. The
assumption that the parts constitute a causal hierarchy is,
an opponent might add, particularly unwarranted; the
preceding paragraphs have presented no evidence that
suggests Aristotle even possessed the notion of hierarchical
organization.

Although the remarks contained in the preceding
paragraphs have not cited evidence that indicates Aristotle
possessed the notion of hierarchical organization, earlier
remarks have cited such evidence. Towards the beginning
of the present chapter, during the discussion of
Aristotle's theory of causality, it was observed that the
distinction between matter and form can be drawn at several
different levels within the same entity. Moreover, that
discussion went on to suggest that some of Aristotle's own
examples—the relation between parts and wholes and the
relation between letters and syllables—could, if developed
a bit further, be turned into examples of hierarchical

organization; and, as a matter of fact, during the course of his logical works Aristotle himself develops the implications of the latter example.\(^63\) Just as the aforementioned passage from the Physics draws attention to the incorporation of letters into syllables, the opening chapters of On Interpretation trace the incorporation of syllables into the smallest independently meaningful linguistic units, e.g., nouns and verbs, and the subsequent incorporation of those into propositions.\(^64\) The opening chapter of Prior Analytics carries the analysis one step further: propositions are, in turn, combined to form syllogisms.\(^65\) Thus, the units located on each of the foregoing levels appear to be parts of a causal hierarchy: letters, syllables, words, propositions, and the syllogism constitute an ascending hierarchy of more inclusive forms in which the syllogism represents the ultimate organization or form and letters represent the ultimate matter, or that which is ultimately organized.\(^66\)

\(^{63}\) See above, p. 488.


\(^{65}\) Prior Analytics, I: 1: 24a10-20.

\(^{66}\) Although some of the foregoing levels coincide with the levels identified by modern linguists, Aristotle is not attempting to identify the constituent elements of language proper; he is, instead, attempting to identify the basic
The remarks contained in the preceding paragraph have cited evidence that strongly suggests that Aristotle did in fact possess the notion of hierarchical organization. By itself, that does not, however, prove that Aristotle conceived of the qualitative parts of tragedy as an instance of hierarchical organization. Additional evidence is needed to substantiate the latter claim. One source of additional support, the one commonly adduced by Chicago School critics, lies in the fact that the top four qualitative parts fit neatly together when seen as part of a causal hierarchy. Telford's commentary contains what is probably the most complete formulation of the preceding line of defense to date. He begins by delimiting the scope of diction, "diction is not the organization of speech into propositions"; that follows, says Telford, from the fact that in chapter 21 of the Poetics the highest form of diction is found to be metaphor, which is simply the "transfer of a word belonging to something else." To make such a transference is not to predicate, and it is predication which is the essential condition of propositions.67

The function of synthesizing diction in such a way that elements utilized in the process of logical reasoning. Thus, syllables and words are being considered not qua constituents of natural language but qua constituents of logical arguments.

67 Telford, p. 84.
propositions result is, however, expressly affirmed of thought. Hence, thought must represent the immediate organization or synthesis of diction, "for words put together to form unities are thoughts." From the foregoing Telford concludes, "Thought therefore is the end to which diction is the means, and thought stands to diction as form to matter." He next turns his attention to the relation between thought and character. After quoting Aristotle's statement that "speeches do not possess character which do not in a general way include what the speaker chooses or avoids," Telford goes on to say:

For speeches to include "in a general way" what is chosen or avoided must mean that the thought contained in the choice must be related to a general tendency and this establishes of what sort the agent is, i.e. his character. Character is thus the organization of thought, for it is character that

68 Telford, p. 84.

69 Telford, p. 84. The last two quotations from Telford appear to involve two different uses of the term "thought"—as the name of a form-matter compound and as the name of the form or organization inherent in that compound. Just as the term "square" can be used to denote both a particular form and also a concrete embodiment of that form, so "thought" can be used to denote both the organizational structure that is imposed on diction and the combination of that form plus diction. It is, of course, thought in the former sense that is related to diction as form to matter; thought in the latter sense is related to diction as hylomorphic compound to the matter out of which it is made. The names of the other qualitative parts exhibit the same sort of systematic ambiguity, and subsequent remarks should be read accordingly.
relates thoughts to a unity. Therefore, just as thought is the form of diction, so character is the form of otherwise separate thoughts, and in this latter relation thought is the material rather than the form.  

Finally, Telford discusses the relation between character and plot:

characters would remain separated and out of relation except for their interrelation through action. And since plot is the imitation of action and "the synthesis of incidents," it must stand to characters as character does to thought and as thought does to diction. . . . Plot is therefore the form of all the other formal parts and is the ultimate end to which they are the means, for it is the requirements of plot that determine what kinds of characters, thoughts, and diction ought to be included.

Thus, Telford manages to show that the Chicago School interpretation represents a tidy, coherent view of the top four qualitative parts.

Plausible as Telford's account is, it does not rule out the possibility that the Chicago School interpretation represents a case of reading more into the relevant passages than Aristotle intended. An opponent might admit that the interpretation constitutes a consistent and interesting reading, but he might still argue that he has not seen

70 Telford, pp. 84-5.
71 Telford, p. 85.
evidence that it is, beyond a reasonable doubt, what Aristotle actually had in mind. Fortunately, the Poetics contains several passages that constitute rather compelling evidence that Aristotle did indeed have something like the Chicago School interpretation in mind.

After he has finished introducing the six qualitative parts, Aristotle presents four arguments designed to prove that plot is the most important of the parts. In the first of those arguments, he maintains that the agents do not act in order to portray the Characters; they include the Characters for the sake of the action. So that it is the action in it, i.e. its Fable or Plot, that is the end and purpose of the tragedy; and the end is everywhere the chief thing.\(^7\)

The passage provides clear evidence that Aristotle thought of a tragedy's plot as the final cause of its characters. The plot is, as he says, that for the sake of which the characters are included; he flatly states that plot is the end of tragedy and even utilizes one of the two terms he customarily uses when discussing final causation, "telos." Moreover, as the end of tragedy, plot must constitute the final cause of everything else contained in a tragedy; it must, in other words, constitute the final cause of not only

\(^7\) Bywater translation, 6: 1450a20-3.
character, but of the other qualitative parts as well.

As earlier remarks have indicated, the final and formal causes of a thing sometimes coincide, and the text of the Poetics contains evidence that such is the case where the relation between plot and the other qualitative parts is concerned. For example, Aristotle's definition of plot as the synthesis of the incidents indicates that plot is the formal cause of the other qualitative parts; as the synthesis—the arrangement or pattern—of the incidents, plot constitutes the form qua structural organization of the other parts. Moreover, after he has finished presenting all four arguments designed to prove that plot is the most important of the qualitative parts, Aristotle concludes that plot "is the principle and, as it were, the soul of tragedy, while characters are second." In On the Soul, Aristotle asserts that the soul constitutes both the formal and final cause of the body, and the point of the preceding passage is apparently that plot stands in a similar relation to everything else contained in a tragedy; just as the soul constitutes both the final and formal cause of the body, so plot constitutes the final and formal cause

73 See above, pp. 489–92.

74 Telford translation, 6: 1450a37–8.
of everything else contained in a tragedy, viz., the other qualitative parts.\textsuperscript{75}

The remarks contained in the last two paragraphs have cited specific passages from the text of the \textit{Poetics} that show that Aristotle's ranking of plot as the most important of the qualitative parts was not based purely on the whims of personal preference, as many commentators have supposed; it was based on the results of a systematic, causal analysis. Aristotle deemed plot the most important of the parts because his analysis showed it to be the formal and final cause of the other parts, and, for him, the final cause— the end to be attained— always possesses the greatest amount of importance. What has been said about the ranking of plot applies, moreover, to Aristotle's subsequent ranking of the other qualitative parts. The evidence cited during the course of the foregoing discussion appears at the beginning of a section of the \textit{Poetics} in which Aristotle ranks not only plot, but each of the other five qualitative parts as well. At the outset of the section, Aristotle has said that plot is the principle and the soul of tragedy, while characters are second.\textsuperscript{76} Since plot and character are ranked in the very same construction, it seems safe to

\textsuperscript{75} \textit{On the Soul}, II: 4: 415b8-11.

\textsuperscript{76} \textit{Poetics} 6: 1450a37-8.
assume that they are being ranked according to the same criteria, namely, their place in a chain of formal-final causes. Following hard upon the heels of the preceding remark come Aristotle's assertions that "Third is thought," and, "Fourth is the diction of the speeches." Since Aristotle does not say otherwise, it must be assumed that the criteria used in determining the importance of each part has remained constant—one part is deemed more important than others because it constitutes the formal and final cause of the others.

In addition to the preceding evidence, there is also the fact that Aristotle repeatedly refers to the qualitative parts as forms, "eide." Occurrence of that term does not—by itself—necessarily prove that Aristotle was thinking of the parts as forms in the literal sense of organizational structures for he sometimes used the word as a means of referring to species or kinds. However, the fact that he defines several of the qualitative parts as syntheses indicates that his reference to the qualitative parts as "forms" was not simply a means of distinguishing them as distinct types or kinds, but was, in addition, intended to convey the idea that they are quite literally

77 Telford translation, 6: 1450b3 and b13 respectively.

78 See, for example, Poetics, 6: 1450a12-3 and 12: 1452b14-5.
organizational structures. Moreover, at the beginning of chapter 12, Aristotle's reference to the qualitative parts as "eide" stands in stark contrast to his reference to the quantitative parts as "divisions" (\textit{diaireitai}).\textsuperscript{79} If Aristotle's use of the term "eide" when referring to the qualitative parts was only meant to convey the sense of "kinds" or "types," then the quantitative parts should count as "eide" just as much as the qualitative parts; they are distinct kinds or types too. Aristotle seems to suggest, however, that the quantitative parts are not "eide"; whereas it is appropriate to refer to the qualitative parts as "forms" because they are literally forms in the sense of organizational structures, the quantitative parts are not organizational structures, and should not, therefore, be referred to as "forms."

Not only does the \textit{Poetics} contain evidence that Aristotle conceived of the top four qualitative parts as forming a chain of successive formal-final causes, it also contains evidence that Aristotle conceived of the lower ranked parts as material causes of the parts ranked above them.\textsuperscript{80} First of all, if Aristotle's statement that plot is

\textsuperscript{79} \textit{Poetics}, 12: 1452b14-5.

\textsuperscript{80} To say that one part constitutes the material cause of the parts above it is, of course, to say that it constitutes the cause of those parts \textit{qua} hylomorphic compounds—\textit{not qua} organizational structures.
the soul of tragedy represents a claim that plot stands to the other qualitative parts as the soul stands to the body, then not only is it the case that plot constitutes the final-formal cause of the other parts, it is also the case that the other qualitative parts constitute the material cause of plot. Second, Aristotle’s discussion of the parts of diction in chapter 20, when combined with his remark that thought involves demonstration, bears out Telford’s claim that diction is the material cause of thought. Thus, specific passages from the text of the Poetics can be used to show that Aristotle saw diction as the material cause of thought and that he saw not only diction and thought, but also character as material causes of plot. The only link that cannot be supported on the basis of specific textual evidence is that between thought and character, and it would seem that the force of the overall pattern is sufficient to secure that link.

The preceding remarks bring to a close discussion of whether or not the Chicago School interpretation can be defended against charges that it represents the results of reading too much into the Poetics. As that discussion has shown, the Chicago School reading can be defended not only on the grounds that it suggests an interesting, consistent
view of the qualitative parts, and on the grounds that it is in keeping with statements made in Aristotle's other works, it can also be defended on the basis of specific passages drawn from the text of the Poetics itself. The preceding discussion has, however, focused on the relations contracted among only the top four qualitative parts. The next order of business is, therefore, to ascertain whether or not the causal chain that links the top four parts extends down to the bottom two parts, melody and spectacle.

According to Telford, the causal chain does not include melody and spectacle:

Melody and spectacle, however, are not related to each other or to the other four parts in quite the same way. These two parts Aristotle calls the ornaments or pleasing accessories of tragedy. As ornaments they will be determined ultimately by the requirements of plot, and in this sense they are materials on which the higher forms impose an organization. That is to say, what melodies are needed to ornament diction will be determined by the interpretation desired, the thought expressed, who expresses it, and under what circumstances. So also spectacle, which is simply the visual aspect of the drama, the scenery, costumes, gestures, etc., will be determined by the requirements of the action presented. But diction obviously cannot be regarded as the synthesis of melody in the sense in which thought is the synthesis of diction. Nor can melody in any sense be the synthesis of spectacle—that is, its unity or form. The relation of melody and spectacle to the other formal parts is therefore of an entirely different order, and this relation is expressed by Aristotle in the term ornament or pleasing accessory. Melody
is the ornament of diction or tragedy's ultimate poetic material, while spectacle is the ornament of plot or tragedy's ultimate poetic form.81

Thus, Telford claims that the causal chain includes only the top four qualitative parts.

While some of what Telford has to say seems correct, some of it seems questionable. For example, while it is true that towards the beginning of chapter 6 Aristotle utilizes the term "kosmos" [κόσμος] or "ornament" when referring to spectacle ("one part of tragedy will be the ornament of spectacle"), the other term that Telford alludes to, "hedysmata" [ἡδυσμάτα], literally means "sweetenings."82 Telford's translation of the term as "pleasing accessories" gives the impression that spectacle and melody are, in some sense, unnecessary. Since the remarks in which the two terms are associated with spectacle and melody are both contained in the chapter during which Aristotle sets out to prove that all six parts are absolutely necessary ingredients of tragedy, the notion that the two terms are meant to convey the idea that the two parts in question are somehow unnecessary seems a rather implausible one. There is,

81 Telford, p. 86
82 Poetics, 6: 1449b31-4 and 1450b16-7 respectively.
moreover, a much less problematic way of construing the import of the two terms. Since both terms were primarily associated with items that appeal to the senses, "edysmaton" being used with reference to sauces and spices and "kosmos" being used with reference to apparel, jewelry, etc., it seems probable that Aristotle's use of the terms when characterizing spectacle and melody was intended to suggest, not that they are in some sense unnecessary, but that they are the most sensual of the qualitative parts. Aristotle's recognition that spectacle and melody are by far the most sensual of the parts has, however, absolutely no bearing on the question of whether or not they form part of the overall hierarchy whose top level is plot.

Telford's assertion that diction cannot be regarded as the synthesis of melody in the same sense that thought is the synthesis of diction appears to be based on the traditional assumption that "melopoia" refers only to the strictly musical aspects of tragedy. Given that assumption, what he says makes perfect sense. The thought being expressed, the dramatic circumstances, etc., could certainly be expected to influence the choice of music used for the sung portions of a tragedy; here music that expresses sorrow is needed, there music that expresses joy is required. However, the music could not be regarded as the material cause of diction, nor could the diction be regarded as the
literal unification of the music. However, all of the preceding claims are based on an interpretation of "melopoia" that remarks contained in the preceding chapter have shown to be untenable. Those remarks showed that the meaning of "melopoia" is not restricted to aspects of tragedy that are musical in the strict sense; the term can be used to refer to the sound made when a text is delivered--whether the text is being sung or spoken. As such, melody does represent the material cause of diction; words are literally made out of sounds. Moreover, diction represents the formal cause of melody; diction represents the synthesis of melody in the same way that thought represents the synthesis of diction--just as a thought represents the synthesis of diverse words into a unity, a word represents the synthesis of diverse sounds into a unity. Of course, some aspects of melody, e.g., segments of instrumental music, non-linguistic vocalizations, sound-effects, etc., do not become incorporated into words and do not, therefore, attain their unity on the level of diction. Some of those aspects, e.g., non-linguistics vocalizations, may become unified on the level of thought; some, e.g., segments of instrumental music, may not become unified until the level of plot is reached. The preceding phenomenon is, however, to be found

---

on other levels as well. For example, Aristotle's remark that "speeches do not possess character that do not in a general way include what the speaker chooses or avoids" clearly indicates that some of the thought contained in a tragedy does not become unified on the level of character; it is at the level of plot that such thought becomes unified. Moreover, Telford himself claims that certain remarks in the Poetics indicate that "just as there may be speeches (expressive of thought) which do not exhibit character, so there may be diction which does not exhibit thought." Thus, the fact that some aspects of melody do not become unified on the level of diction does not constitute a legitimate reason for denying melody a proper place in the causal chain whose highest level is plot.

84 Telford translation, 6: 1450b9-11.
86 Incidentally, the phenomenon discussed above marks one of the most obvious differences between the hierarchical organization exhibited by the six qualitative parts of tragedy and the sort of hierarchical structure exhibited by natural languages. As remarks contained in the first chapter of the present work indicated, Pike identified a phenomenon in linguistic hierarchies whereby a unit on one level also constitutes a single unit on a higher level before it is incorporated into a larger unit on a still higher level—see above, p. 60. However, the phenomenon described above is characterized by the fact that units on one level completely skip one or more levels before becoming incorporated into a larger unit on a higher level, a phenomenon that Pike himself emphatically distinguishes from the so-called "portmanteau" phenomenon (see Pike, p. 440).
The preceding remarks have shown that the causal chain includes melody, but what about spectacle? Telford claims that spectacle and melody do not exhibit anything even remotely similar to the reciprocal relations exhibited by adjacent qualitative parts nearer the top of Aristotle's list, e.g., diction and thought. The following discussion will, however, show that, while spectacle and melody do not exhibit exactly the same relations as do other pairs of adjacent parts, there are some very important similarities, important enough to warrant inclusion of spectacle in the causal chain. The discussion will begin with an examination of whether spectacle can be considered the material cause of melody.

Telford's remarks indicate that he does not believe that spectacle can be regarded as—in any sense—the material cause of melody. However, several passages contained in book II, chapters 7 and 8 of *On the Soul* suggest that Telford's apparently plausible claim may be mistaken. Having, in chapter 7, discussed sight and its object, the visible, Aristotle turns his attention, in chapter 8, to hearing and its object, the audible. 87 "What is required for the production of sound," he says," is an

impact of two solids against one another and against the air."\textsuperscript{88} Being solid, the entities needed for the aforementioned sort of collision will, of course, be visible—at least they will be potentially so.\textsuperscript{89} Thus, sounds are produced by the collision of visible objects. Now the collisions apparently constitute the efficient cause of sound—they set the air in motion; and it is the movement of the air that, according to Aristotle, activates the appropriate sense organs, viz., the ears; and the objects themselves apparently constitute the material cause of sound—the objects involved in the collisions constitute the material cause of sound just as quantities of bronze and marble might constitute the material cause of a statue.\textsuperscript{90}

The remarks contained in the preceding paragraph suggest that there may be a sense in which spectacle can be

\textsuperscript{88} On the Soul, II: 8: 419b20-1.

\textsuperscript{89} It is, of course, possible that the two colliding bodies responsible for producing a sound might be encased in a larger object that obscures them from actual view. For example, in chapter 8 of On the Soul (420b28), Aristotle analyzes the production of speech along the lines sketched above, the windpipe being a major factor in the collisions responsible for setting the air in motion. While the inside of the windpipe is, of course, normally obstructed from view by the layers of tissue and muscle that surround it, it is, nevertheless potentially visible in the sense that it could be seen were it laid open to view.

\textsuperscript{90} On the Soul, 7: 419a25-32.
considered the material cause of melody. Since visible objects constitute the material cause of sounds, and since the visible aspects of the objects contained in a tragedy fall under the rubric of spectacle and the audible aspects fall under the rubric of melody, there would seem to be a sense in which the former constitutes the material cause of the latter. While the preceding remark indicates that the relation between spectacle and melody bears some similarity to the relations contracted between other pairs of adjacent qualitative parts, the extent of that similarity is limited. Whereas a thought is, for example, literally made out of diction, i.e., words, a sound is not literally made out of sights. A sound is made out of colliding objects—which happen to be visible. It is not, in other words, the sights themselves, but objects to which the sights are tied that constitute the material cause of melody. Thus, in Aristotelian terminology, spectacle is only an incidental or concomitant cause of melody, whereas diction constitutes the specific material cause of thought.  

Were the question of whether spectacle should be considered part of the causal chain to be answered solely on the basis of what has been said so far, the answer might

---

91 The distinction between specific and concomitant causes is discussed above; see p. 492.
have to be negative; on the one hand, spectacle constitutes the material cause of melody, but, on the other hand, it does not constitute the specific material cause of melody—as is the case with other adjacent pairs of qualitative parts. It would, however, be premature to try to answer the preceding question solely on the basis of what has been said so far because only one side of the reciprocal causal relation has been discussed; the possibility that melody constitutes the final-formal cause of spectacle has not yet been considered. Is there any sense in which melody can be said to be the final-formal cause of spectacle? The answer would seem to be, at least in some cases, affirmative.

Consider, for example, a case in which a poet creates a dramatic situation that requires one of the *dramatis personae* to applaud another one. Production of a particular sound clearly represents the final cause, the goal to be achieved, and in order to attain that goal, the spectacle will have to be formed in a certain way: the spectacle will have to include the appearance of the dramatic personage repeatedly joining and separating his hands if the series of collisions required in order to make the sound are to be brought about. Thus, in the preceding example, melody constitutes the final cause of spectacle, and it also constitutes the formal cause in the sense that it determines the overall form or organization that the spectacle must be
given. There is, in addition, a sense in which melody constitutes the synthesis or unity of spectacle: it is the sound produced that unites the otherwise disconnected sights, allowing a perceiver to see them as parts of a single act. The relationship between spectacle and melody illustrated by the preceding example is, moreover, one frequently found in a play; whenever a dramatic personage delivers dialogue, the sounds he makes constitute the final and formal cause of the visible movements of his mouth, etc.

There is, however, an additional sense in which thought can be said to represent the formal cause of diction that does not apply in the preceding sorts of examples. Since diction is the specific material cause of thought, a given thought can be said to constitute the literal, concrete unification of the various words it contains. However, the same cannot be said in the case of melody and spectacle. In the preceding example, for instance, the sound of applause represents the factor that gives unity to a sequence of sights, but it cannot be said that the sound literally is the concrete unification of those sights.

The remarks contained in the preceding paragraph have indicated that melody does not constitute the final-formal cause of spectacle in precisely the same way that each of the other qualitative parts constitutes the final-formal cause of the part immediately below it. Those, and earlier,
remarks, have, however, shown that the relation between the former two parts does resemble the relations contracted between the latter pairs of parts in a number of important respects. Of particular importance is that fact that the requirements of melody often determine the form that must be given to spectacle. When combined with the various other similarities that have been identified, the preceding fact provides fairly strong support for the claim that Aristotle considered spectacle to be part of the causal chain.

During the preceding pages, several arguments designed to show that the causal chain includes both melody and spectacle have been presented. An opponent might, however, object that none of those arguments are based on specific passages located in the text of the Poetics itself. The argument designed to show that melody is part of the causal chain claimed that once the true meaning of "melopoia" is understood, its place in the causal chain will become obvious. The argument designed to show that spectacle is part of the causal chain relied heavily on passages located in On the Soul, not on ones located in the Poetics. Is there, an opponent might ask, specific evidence located in the text of the Poetics itself that supports the interpretation in question? The answer is yes. Earlier remarks have shown that Aristotle's ranking of the qualitative parts in chapter 6 is
based on the place each one occupies in a causal chain.\textsuperscript{92} However, the top four parts are not the only ones ranked during the passage in question. Aristotle goes on to say that "Of the remaining parts the making of melody is the most important," which, in turn implies that spectacle is the least important.\textsuperscript{93} Since Aristotle does not indicate that the criteria upon which he is basing his assessment of importance has changed, it seems safe to assume that the assessments of melody and spectacle are based on the same causal criteria that he has used in ranking the top four parts. Thus, it would appear that melody and spectacle must also be part of the causal chain.

The preceding remarks bring to a close the present discussion of whether the Chicago School view of the qualitative parts represents a historically plausible reconstruction of Aristotle's conception of the matter. That discussion has not only shown that the Chicago School view of the top four qualitative parts can be supported in a number of ways, ways which include the citation of specific, corroborating passages from the text of the \textit{Poetics}, the discussion has also shown that similar sources of evidence

\textsuperscript{92} See above, pp. 526-27.

\textsuperscript{93} Telford translation, 6: 1450b16.
can be adduced in support of an extension of the Chicago School view that allows for inclusion of spectacle and melody as parts of the overall causal hierarchy.
CHAPTER VII

Conclusion

This work has advanced the Structuralist study of literature in a number of ways. First, it has shown how the problematic features of several key linguistic concepts utilized by Structuralists can be resolved, thereby rendering their literary application more plausible. In Chapter Three, for example, problematic features of the distinction between langue, parole, performance, and competence as well as the distinction between signifier and signified were isolated and solutions were suggested.

Second, discussions in the first half of the study have spelled out the details of the literature-language analogy to a greater extent than have the works of previous authors. So doing has not only helped to clarify the nature of the Structuralist approach, it has also increased the plausibility of various claims that are basic to the approach. For example, the second chapter has shown that it is possible to formulate philosophically respectable characterizations of Structures in both epistemological and ontological terms,
that both linguistic units and literary works fit those characterizations, and that that fact can be used to strengthen the case for the literature-language analogy. Likewise, Chapter Four has shown how the distinction between syntax, semantics, and pragmatics can be applied within the domain of literature, thereby providing a more detailed conception of the analogy between the generative approach to language and a generative approach to literature; and it has shown that the latter approach possesses more plausibility than most previous writers have thought.

The greatest potential contribution of the present study lies, however, in its demonstration that the qualitative parts of tragedy constitute a series of hierarchically organized levels of dramatic structure. As remarks contained in Chapters Three and Four indicated, one of the greatest stumbling blocks facing Structuralist criticism has been its failure to identify a comprehensive hierarchy of structural levels in the domain of literature analogous to those identified by linguists. The remarks contained in Chapter Six suggest that Aristotle's qualitative parts constitute just such a hierarchy.

Naturally, that suggestion opens a whole new area of future research. The next step in that research would consist of efforts to isolate the primary types of discrete,
quantitative units located on each of the six qualitative levels—the dramatic analogues of phonemes, morphemes, etc.—which could then be classified into paradigmatic sets of various sorts. Next, it would be necessary to investigate the integrative relations whereby units on lower levels combine to form units on higher levels. Perhaps it will even be possible to characterize those relations in generative terms.

The present study has thus not only articulated the nature of the Structuralist approach to dramatic literature and strengthened the arguments in favor of that approach, it has also suggested a solution to one of the major problems facing the Structuralist enterprise; and, in so doing, it has opened new avenues of future research.
APPENDIX A

Before the second reason why it does not seem necessary to provide an exhaustive, point-by-point discussion of deconstructionist attacks on Structuralist concepts and techniques can be clearly stated, it will be necessary to consider some basic characteristics of deconstructionist criticism. The following discussion will, therefore, begin with some further remarks about the nature of deconstructionism itself.

Of the various assumptions of traditional Western thought that deconstructionists have called into question, two warrant special consideration—not only because virtually all deconstructionists have challenged them, but also because virtually all previous deconstructionist arguments either attempt to prove that the assumptions are indeed questionable or depend on that claim as a premise. First, deconstructionists have denied the traditional assumption that at least some utterances have a more or less determinate meaning. In "The Critic as Host," for example, Miller suggests that "the notion of determinable meaning" represents "a demonstrable impossibility"; and in "Limited
inc abc . . .," Derrida maintains that there can be no "adequation between meaning and saying"—that

[we are left with] no choice but to mean (to say) something that is (already, always, also) other than what we mean (to say), to say something other than what we say and would have wanted to say, to understand something other than . . . etc.¹

The traditional assumption that at least some statements are true has constituted a second target of deconstructionist criticism. Thus, in Of Grammatology, Derrida claims that the experience of the effacement of the signifier [the visible or audible aspect of language] in the voice is not merely one illusion of

¹ Miller, "Host," pp. 439-40 and Derrida, "Limited Inc," p. 200, respectively. The ellipses appear in the original Derrida quotation; nothing has been deleted from the passage in question. In fairness to Derrida, it should be pointed out that he does not deny the phenomenon of meaning altogether. In Writing and Difference, for example, he suggests the possibility that "the meaning of meaning (in the most general sense of meaning and not of indication) is infinite implication. . . . the unchecked referral from signifier to signifier. . . . [that] its force is a certain pure and infinite equivocalness, which gives signified meaning no respite, no rest, but engages it within its own economy to go on signifying and to differ/differ" (Writing and Difference, trans. Alan Bass [Chicago: University of Chicago Press, 1978], p. 25). As the preceding quotation indicates, however, no utterance could ever possess a determinate meaning on such a view; each utterance would be "ininitely equivocal." The important point for the purposes of present discussion is Derrida's claim that while utterances may be meaningful in some sense, they do not possess determinate meanings—in any sense.
many—since it is the condition of the very idea of truth. . . . This illusion is the history of truth . . . .

Similarly, in "Tradition and Difference," Miller approvingly quotes Nietzsche's assertion that "truths are illusions whose illusory nature we have forgotten . . . ."\(^3\)

The deconstructionists' efforts to call traditional assumptions about meaning and truth into question have, however, invariably produced paradoxical results, and there appear to be reasons for thinking that no argument could possibly prove that the assumptions in question are, in fact, false.\(^4\) Consider, for example, what will occur when


\(^3\) J. Hillis Miller, "Tradition and Difference," *Diacritics*, 2 (Winter 1972), p. 11. It is important to realize that the preceding claims are not directed solely at so-called "correspondence" theories of truth. As Culler points out, the deconstructionist attack is directed just as much at so-called "pragmatist" theories of truth as it is at correspondence theories (Culler, *Deconstruction*, pp. 152–55). In fact, the deconstructionist claim would appear to be that no utterances are true—regardless of what theory is used to explicate the notion of truth—be it a correspondence, coherence, semantic, redundancy, pragmatic—or any other—theory.

\(^4\) The claim that no argument could possibly prove false the assumptions about meaning and truth in question naturally raises the question of whether such assumptions constitute necessary truths, and, if they do, whether they constitute analytic truths or synthetic a priori truths. Unfortunately, such questions could not be answered without first defining the terms "necessary truth", "analytic truth," and "synthetic a priori truth," a task that would
a deconstructionist attempts to state the conclusion of an argument intended to prove that the aforementioned traditional assumption about meaning is false. The conclusion of

require a series of long and complicated discussions. No attempt will, therefore, be made to provide answers to the foregoing questions.

It is, incidentally, worth noting that the deconstructionists have not been the only twentieth-century writers to deny that utterances have a determinate meaning. In the second chapter of his book *Word and Object*, Willard Van Orman Quine has made similar claims—*Word and Object*, Willard Van Orman Quine (Cambridge: The M.I.T. Press, 1960), pp. 26-79. Quine focuses on cases of so-called "radical translation," i.e., attempts to translate totally unfamiliar languages, but repeatedly says that his arguments also apply to the utterances of a person's own native language—see especially pp. 26-27, 59, 78-79. However, there appear to be a number of *prima facie* differences between the claims made by Quine and those made by the deconstructionists. For example, it is often said that, as Gilbert Harman puts it, Quine only objects to "philosophical (as opposed to ordinary) talk about meaning, meanings, propositions, and propositional attitudes" (Gilbert Harman, "An Introduction to 'Translation and Meaning': Chapter Two of Word and Object", in *Words and Objections: Essays on the Work of W. V. Quine*, ed. Donald Davidson and Jaakko Hintikka [Dordrecht, Holland: D. Reidel Publishing Company, 1969], p. 15). Derrida, on the other hand, insists that "'everyday language' is not innocent or neutral. It is the language of Western metaphysics, and it carries with it not only a considerable number of presuppositions of all types, but also presuppositions inseparable from metaphysics, which, although little attended to, are knotted into a system" (Derrida, *Positions*, p. 19). Thus, it is unlikely that Derrida would consider "ordinary" talk of meaning immune to his arguments. However, Derrida's view raises the question of why Quine should consider ordinary talk of meaning and meanings acceptable. If all efforts to explicate meaning are doomed to failure, as Quine seems to think, then ordinary talk of meaning must presumably be based on illusory notions just as its philosophical counterpart is. Of course, Quine may say that ordinary talk is acceptable in the sense that the illusory nature of such talk does not lead to problems the way the
the argument is presumably to be either of the following two, logically equivalent statements: "it is false that at least some utterances have a determinate meaning," or "no utterance has a determinate meaning." Regardless of the premises used in such an argument, as soon as a deconstructionist attempts to state the conclusion, the question of whether the conclusion itself has a determinate meaning will arise. If a deconstructionist assumes that the conclusion does not itself have a more or less determinate meaning, the argument will be rendered incapable of proving its own point—for a statement without any determinate meaning can prove nothing whatsoever. If, on the other hand, a deconstructionist assumes that the conclusion does have a more or less determinate meaning, he will be in the position of having to assume that the conclusion is false at the very moment he claims it to be true; in order to assert that no philosopher's talk of meaning allegedly does. In other words, Quine might want to say that although no utterances possess a determinate meaning, it is alright—for the purposes of everyday conversation—to pretend as if they did. In that case, however, the prima facie difference between Quine and the deconstructionists would not really amount to much after all; much the same might hold true in the case of the other prima facie dissimilarities as well. If such were to be the case, Quine's arguments for the indeterminacy of meaning would also be subject to the following criticisms. Although the foregoing possibility is intriguing, it will not be explored in this study.
utterance has a determinate meaning, he must assume that at
least one utterance, viz., the utterance he is in the pro-
cess of asserting, does have a more or less determinate
meaning. Thus, either the conclusion of the proof does not
itself have a determinate meaning and is, therefore, inca-
ble of proving anything whatsoever; or the conclusion turns
out to be self-contradictory and, therefore, false. Either
way, the argument will fail to prove what it set out to
prove.

Not only do deconstructionists have to assume that at
least some utterances have a determinate meaning in order to
deny that very assumption, but, as Christopher Norris
observes, they typically require others to make the same
assumption:

[Deconstructionists typically] demand that
their texts be properly understood—or at
least intelligently read—while ostensibly

5 Of course, the conclusion is not self-contradictory
in the explicit way that "p and not p" is. The contradic-
tion lies in the relation between what must be the case in
order to assert the statement and what the statement asserts
to be the case—between what must be the case if the con-
clusion is to possess what philosophers like Austin and
Searle have called "illocutionary force" and what the
conclusion would assert if it did have any illocutionary
force (Austin, p. 100 and Searle, Speech Acts, p. 30).
Thus, the statement involves what might be called an "illo-
cutionary contradiction."
denying the power of language to encompass any such end.

In fact, it is not unusual for a deconstructionist to claim, within the space of one and the same essay, both that an utterance could never possess a determinate meaning and that opponents have misunderstood the meaning of his own previous utterances. For example, Derrida's declaration that a person can never mean what he says, say what he means, understand what someone else has said, understand what someone else meant to say, etc., occurs in the midst of what amounts to an eighty-nine page catalogue of charges that Searle has misunderstood various remarks Derrida made in an earlier essay. The following constitutes but one of the seemingly endless items in that catalogue:

Had Sari [Derrida's name for Searle and several associates] been sufficiently present to what it was writing or rewriting, the passage in question might have cleared up the misunderstanding: in it, what is discussed, with an insistence that should have prevented all haste and confusion, concerns not permanence, but remainders or remains, non-present remains.

Thus, Derrida apparently thinks that "Sari" misunderstood

6 Norris, p. 127.
the meaning—whatever that could be—of his earlier remarks—and after Derrida had so insistently tried to prevent such confusion.⁸

To further complicate matters, deconstructionists often use various devices like putting quotation marks around the term "meaning" in their criticisms of opponents, thereby raising questions about whether they "really" "mean" what they are saying about their opponents having misunderstood the "meaning" of their own previous remarks. So doing has the distinct advantage of allowing deconstructionists to imply that they cannot be held responsible for their criticisms of opponents even as they assert them—after all, how can an opponent justifiably hold them responsible for what they have said if he can never be sure that they said what they "meant," "meant" what they said, etc.? In fact, the deconstructionists' habit of criticizing or bracketing the very

---

⁸ It is perhaps possible that Derrida wishes to accuse "Sari" of misunderstanding his previous remarks without implying that "Sari" misunderstood the "meaning" of those remarks. Derrida might, in other words, think that it is possible to provide an account of linguistic understanding that does not invoke the notion of "meaning." In view of the fact that he offers no such alternative account of linguistic understanding, and in view of his repeated assertion that "nothing is conceivable" without the notion of a "transcendental signified" (his term for "determinate meaning"), the foregoing seems, however, a rather remote possibility—see, for example, Grammatology, pp. 13-14 and Writing and Difference, pp. 280-81.
concepts and principles they are using has—not surprisingly—made most commentators rather wary about attributing firm beliefs of any kind to them. For example, Christopher Butler feels obliged to qualify his entire discussion of Derrida in advance:

> the relatively firm positions which I (and others) will attribute to him in what follows, are in fact alien to his own method, which is 'slippery' and in a perpetual movement of self-qualification.  

In view of the hazards that obviously accompany any attempt to interpret the meaning of utterances made by authors who claim that no utterance ever has a determinate meaning, it may be prudent to hereby issue a similarly qualificatory caveat *vis a vis* the present study's attempts to interpret the meaning of deconstructionist texts.

Deconstructionist attempts to challenge the traditional assumption that at least some utterances are true exhibit inconsistencies precisely parallel to those mentioned above. Consider, for example, what will occur when a deconstructionist attempts to state the conclusion of an argument advanced as a proof that the aforementioned traditional assumption about truth is false. The conclusion of the argument is

---

9 Butler, p. 60.
presumably to be either of the following two, logically equivalent statements: "it is false that at least some statements are true," or "no statements are true."

Much as in the aforementioned case of arguments concerning the determinacy of meaning, regardless of the premises used in such an argument, as soon as a deconstructionist attempts to state the conclusion, the question of whether the conclusion itself is true will arise. If the conclusion is not assumed to be true (either because it is assumed to be false or because it is assumed to have no determinate truth value), the argument will be rendered incapable of proving its point—for a statement that is not true cannot succeed in proving anything at all. If, on the other hand, a deconstructionist assumes that the conclusion is true, he will be in the position of having to assume that the conclusion is false at the very moment he claims it to be true; in order to assert that no statements are true, he must assume—on the contrary—that at least one statement is true, viz., the statement he is in the process of asserting. Thus, either the conclusion of the proof is not itself true and is, therefore, incapable of proving anything whatsoever; or the conclusion will be caught in an illocutionary contradiction. Either way, the argument will fail to prove what it set out to prove.
There is, moreover, a further parallel between the inconsistencies that arise in the deconstructionist treatments of meaning and truth for it is not unusual for a deconstructionist to claim, within the confines of one and the same essay, both that truth and falsity are illusions, and that the statements of an opponent are false—while at the same time attempting to relieve himself of any responsibility for the latter claim by placing the word "false" in quotation marks. For example, within the space of several pages, Derrida asks the following presumably rhetorical question: "if Sarle's [previously mentioned] statement is difficult to accept, is it not because in each of its words it is 'obviously false'?"—and he criticizes "Sar1" for assuming that his own statements can be evaluated as either true or false:

Instead of precipitously [assuming that the statements can be evaluated as true or false] . . . , a theoretician of speech acts who was even moderately consistent with his own theory ought to have spent some time patiently considering questions of this type: Does the principle purpose of Sec [Derrida's earlier essay] consist in being true? In appearing true? In stating the truth? And what if Sec were doing something else?

What? . . . offering the performance of a text which, by raising in passing the question of truth . . . does not simply succumb to its jurisdiction and remains, at this point . . . irreducible to "verdictive" . . .
sentences of the type: this is true, this is false, "completely mistaken" or "obviously false." Thus, "Sari's" statements can be criticized as "obviously false"—although the quotation marks warn that Derrida might not "mean" that they are "really false"—truth and falsity being, after all, illusions. Derrida's statements appear, however, to fall conveniently outside the scope of such illusory notions as truth and falsity.

There are several ways in which deconstructionists might attempt to respond to the foregoing discussion of their efforts to challenge traditional assumptions about meaning and truth. Earlier remarks have drawn attention to the fact that deconstructionists sometimes place quotation marks around words, one of several devices which, as M. H. Abrams observes, deconstructionists employ in order to indicate that they are using a given word—and accepting the various assumptions that go along with its use—only "provisionally," or sous rature ("under erasure"). In logic,

10 Derrida, "Limited Inc," pp. 181 & 179 respectively.

11 Abrams, "How to Do Things," p. 573. In Grammatology, for example, Derrida says at one point, "Like all the notions I am using here, [the concept of experience] belongs to the history of metaphysics and we can only use it under erasure [sous rature]" (Grammatology, p. 60; the bracketed French terms appear in the original translation).
the "reductio ad absurdum" or "indirect proof" enables a person to provisionally accept a claim as true at the beginning of an argument in order to later prove that the claim is false, and it might be suggested that deconstructionist arguments exhibit the preceding pattern of argumentation and thereby manage to avoid all the problems discussed above. However, a closer look at the pattern exhibited by a reductio proof will show that not even it could be used to prove that the traditional assumptions about meaning and truth are false.

The overall pattern of a reductio argument is roughly as follows: at the beginning of such a proof, a person provisionally assumes the truth of the statement he wishes to prove false, say, for example, the statement "not p." He then attempts to show that "not p" logically implies a contradictory statement. Now all contradictory statements are necessarily false, and a false statement cannot be validly inferred from all true premises. If he succeeds in


13 See Kahane, pp. 80-81 and Salmon, pp. 30-32 for descriptions of the pattern of reductio proofs much like the following one.
validly deducing a necessarily false statement from his original premise, "not p," he will, therefore, have succeeded in proving that it is also false and can go on to state the conclusion of his argument: "it is false that not p," or the logically equivalent statement, "p." Having shown that the argument's provisional premise logically implies a contradiction, he can, in other words, discharge that provisional assumption before going on to state his conclusion--but that is precisely what the arguments concerning meaning and truth will not allow a deconstructionist to do. A deconstructionist cannot provisionally accept the traditional assumption about meaning or the assumption about truth as a premise that he can discharge before he goes on to state the conclusion of his argument--for that assumption is required by the conclusion itself: if he does not continue to assume his original premise when stating the conclusion, the conclusion will be incapable of proving anything whatsoever--but if he does continue to assume the original premise, he will thereby contradict the conclusion, and the entire argument will automatically reduce itself to absurdity. Thus, any attempt to utilize the pattern of a reductio proof in hopes of avoiding the sorts of problems discussed earlier will fail.

Perhaps it might be argued that the preceding paragraphs have not, however, taken into account the full range
of implications associated with the technique of employing terms "under erasure." During the present century, it has become standard practice among logicians to draw a distinction between the "object language" of an investigation and the "metalanguage" in which the investigation is carried out. The distinction can also be applied to the domain of natural languages, and it will perhaps be easier to explain the distinction by utilizing examples from that domain. Consider the example of a linguist who writes a paper in English about the German language. In that investigation, German would constitute the object language—the object being investigated—and English would constitute the metalanguage—the language used to carry out the investigation. Of course, if a linguist were to write a paper in German about the English language, precisely the opposite would be the case; English would constitute the object language and German would constitute the metalanguage. Closely associated with the foregoing distinction is a distinction between "using" a term and "mentioning" a term. Logicians would say that in the English article about the German language, the English words are being "used" while the German words

14 See, for example, the discussion of the distinction contained in Richmond H. Thomason, Symbolic Logic (London: The Macmillan Company, 1970), pp. 53-54.

15 See, for example, Thomason, pp. 56-57.
are only being "mentioned." Just the reverse would be true in the case of the German article about the English language.

One of the reasons that the aforementioned distinctions have become important in the area of logic is that logicians often formulate proofs concerning the properties of one logical system or language within the framework of a separate logical system or language; one system constitutes the object language while the other constitutes the metalanguage. Moreover, a proof carried out in the metalanguage often "mentions" terms that belong to the object language being investigated.  

The relevance of the foregoing distinctions to the present discussion is that someone might wish to argue that when a deconstructionist employs the terms "meaning" and "truth" in his proofs, readers are to assume that the proof is being carried out in a metalanguage, that the terms in question belong instead to an object language, and that they are being mentioned rather than used. That, it might be argued, is what is meant by saying that the terms are being used "under erasure."

Utilizing the foregoing strategy, a deconstructionist might reformulate his arguments against the traditional

16 Thomason, pp. 56-60.
assumptions concerning meaning and truth as follows. Suppose that a hierarchy of languages \( L_1, L_2, \ldots, L_n \) can be constructed, and suppose that each language contains the predicate "has a determinate meaning" and the predicate "is true." Thus, \( L_1 \) will contain the predicate "has a determinate meaning" and "is true," \( L_2 \) will contain "has a determinate meaning" and "is true," etc. Now consider a version of the \textit{reductio} argument concerning meaning that is formulated in the language of \( L_2 \), and assume that every time the predicate "has a determinate meaning" occurs during the course of the proof it is "has a determinate meaning" rather than "has a determinate meaning" that is involved—assume, in other words, that the expression constitutes a term from an object language that is being mentioned rather than a term from the metalanguage that is being used. Such a strategy will enable a deconstructionist to claim that the conclusion of his proof does not fall prey to the vitiating dilemma discussed earlier. He can say that the conclusion of his proof \textit{does} have a determinate meaning in the sense of "has a determinate meaning 2"; the conclusion only denies that any utterance has a determinate meaning in the sense of "has a determinate meaning 1." The conclusion is not, therefore, involved in any kind of self-contradiction. The same strategy will, of course, also be applicable to the argument concerning truth; the conclusion will be true in the sense
of "is true_2," although it denies that any statement is true in the sense of "is true_1." Thus, it might be claimed that within any language L_x, the arguments in question will only claim that no utterances have a determinate meaning or that no utterances are true in the sense of "has a determinate meaning_x-1" or "is true_x-1."

While the preceding reformulations of the deconstructionist arguments successfully avoid the dilemmas discussed earlier, they appear to fall prey to the following equally serious dilemma. Either the hierarchy of constructed languages will have a top level, or it will go on infinitely. If a highest level is reached, call it L_n, the question will arise whether the conclusions of the proofs formulated in that language have a determinate meaning and are true in the sense of "has a determinate meaning_n" and "is true_n." If a deconstructionist admits that they do have a determinate meaning and that they are true, then it will turn out that at least some utterances have a determinate meaning and that at least some utterances are true after all. If he denies that they have a determinate meaning and are true in the specified senses, then they will be incapable of proving that no utterance has a determinate meaning and that no utterances are true in the sense of "has a determinate meaning_n-1" and "is true_n-1." That will, in turn, mean that the conclusions of the proofs formulated in the language L_n-1
are incapable of proving their claims *vis a vis* the utterances of $L_{n-2}$, and so on all the way down to the bottom language.

If, on the other hand, there are an infinite number of ascending languages, it will always be possible to show that the proofs formulated on any given level are not really capable of proving what they set out to prove. Given any language $L_n$, it will always be possible to show, by ascending to the next higher language $L_{n+1}$, that the conclusions of the proofs formulated in $L_n$ only *appeared* to have a determinate meaning and only *appeared* to be true in the sense of "has a determinate meaning$_n$" and "is true$_n$." Since they turn out not to have a determinate meaning and not to be true *vis a vis* $L_n$ after all, they turn out not to have been capable of proving their claims *vis a vis* the utterances of $L_{n-1}$ either. Thus, regardless of whether the hierarchy of constructed languages has a top level or goes on infinitely, there will, once again, be no way of proving that no utterance has a determinate meaning or that no utterances are true.

Not only is it the case that the preceding sort of strategy inevitably fails to prove that the traditional claims about meaning and truth are false, a number of Derrida's remarks indicate that he does not himself have such a strategy in mind. For example, in "Limited Inc," he
repeatedly attacks the very distinction upon which the whole preceding strategy is based— the distinction between use and mention. Moreover, in Writing and Difference, he asserts that

we can pronounce not a single destructive proposition which has not already had to slip into the form, the logic, and the implicit postulations of precisely what it seeks to contest.

Thus, Derrida seems to think that deconstructionists will have to assume—not concepts that are located on a level one step above the level of the concepts they wish to deny— but, instead, the very same concepts they wish to deny. The preceding pages have shown that even if he were to reformulate his arguments along the aforementioned lines, however, he would still not be able to prove his point.

Insurmountable problems will thus arise every time a premise or the conclusion of a deconstructionist argument attempts to challenge the traditional assumptions about meaning and truth. The preceding remarks have already shown what happens when such a claim is to be the conclusion of an argument, and similar problems will arise when such a claim

18 Derrida, Difference, pp. 280-81.
is to be used as a premise. Since, under every possible option, the premise will ultimately fail to be true, the argument in question will turn out to have at least one premise that is not true—which means that even if the argument is valid, it will not succeed in proving that its conclusion is true. Thus, every deconstructionist argument that contains such a claim as either its conclusion or as a premise will be incapable of proving that its conclusion is true. Since virtually all deconstructionist arguments fall into one or the other of those two categories, they will virtually all be incapable of proving what they set out to prove.

Throughout the preceding remarks, it has been assumed that the basic principles of logic are reliable—that, for example, all contradictory statements are necessarily false. However, the traditional claim that at least some principles of logic are correct or reliable has itself been the target of deconstructionist attacks. For example, Derrida responds to statements in which "Sar1" assumes the notion of logical priority as follows:

Logic, the logical, the logos of logic cannot be the decisive instance here: rather, it constitutes the object of the debate, the phenomenon that must first be explained before it can be accepted as the deciding instance. The matter we are discussing here concerns the value, possibility, and system of what is called logic in general. The law
and the effects with which we have been dealing, those of iterability for example, govern the possibility of every logical proposition, whether considered as a speech act or not. No constituted logic nor any rule of a logical order can, therefore, provide a decision or impose its norms upon these prelogical possibilities of logic.  

Of course, any attempt to prove that the aforementioned traditional assumption about logic is false will encounter much the same problems already discussed in connection with the arguments about meaning and truth. When a deconstructionist attempts to state the conclusion of his argument, "it is false that at least some rules of logic are reliable" or "no rules of logic are reliable," questions concerning the reliability of the rules employed in the very proof he is advancing will arise. If he assumes that the rules he has employed are not reliable, there will be no basis for believing that the conclusion is true, and the conclusion will, therefore, be rendered incapable of proving what it is supposed to prove. If, on the other hand, he assumes that the rules he has employed are reliable, then he will be in the position of denying the conclusion at the very moment he asserts it. If he attempts to construct an ascending series of logical systems, differentiates between the metalanguage

in which any given proof is being carried out and the object language being investigated, and claims that the conclusion of an argument only denies the legitimacy of the logical principles located in the object language; his arguments will still encounter the problems discussed several pages ago. Once again, every possible strategy will ultimately fail.

A deconstructionist might, however, attempt to claim that it is inappropriate to use rules of logic in assessing deconstructionist arguments about meaning, truth, and logic—not because he can prove that the rules of logic are unreliable—but because the so-called "arguments" in question were not really intended to constitute arguments at all. In a remark quoted earlier, Derrida suggests that he is "offering the performance of a text [that raises] the question of truth . . . [without] succumb[ing] to its jurisdiction," and a deconstructionist might wish to respond to the preceding discussion of deconstructionism in much the same terms: deconstructionist texts do not attempt to say anything true or prove that anything is the case (or attempt to use language to convey determinate meanings?); deconstructionist texts

20 See above, pp. 563-64.
merely attempt to raise certain sorts of questions.  

The foregoing sort of response might well lead a critic to wonder, however, what the point of such deconstructionist

21 See above, p. 556. As remarks contained in Chapter One of this study indicated, deconstructionism appears to constitute a form of skepticism (see above, pp. 5-6, 9, 12); and it is worth noting that the difference between what might be called the "assertive" version of deconstructionism considered during the course of the foregoing pages and the version mentioned above, what might be called the "interrogative" version of deconstructionism, bears interesting similarities to the difference between two distinct versions of classical skepticism, "academic skepticism" and "Pyrrhonian skepticism." Philosophers have advanced a variety of slightly different characterizations of the two brands of skepticism, but the difference can be put—in very rough terms—as follows: "academic skepticism" is based on the claim that it is impossible to attain knowledge; "Pyrrhonian skepticism" urges man to suspend judgment concerning the answer to any question whatsoever, including the question of whether it is possible to attain knowledge (see, for example, Richard H. Popkin, The History of Scepticism from Erasmus to Descartes, revised edition [1964; rpt. New York: Harper & Row, Publishers, 1968], pp. ix-xi, and Craig B. Brush, Montaigne and Bayle: Variations on the Theme of Skepticism [The Hague: Martinus Nijhoff, 1966], pp. 4-13). Thus, whereas an academic skeptic might claim that he knows that he knows nothing and can prove it, the Pyrrhonian refrains from making any such claims, and although he may present arguments, he typically refrains from saying that he knows them to be sound (see Popkin, pp. xi-ix). The difference between the two possible readings of deconstructionist texts depends, likewise, on whether deconstructionists are construed as claiming that they can prove that no utterance has a determinate meaning, no utterance is true, no logical principles reliable, and hence that no knowledge is possible—or—whether their texts are, on the other hand, given a thoroughly non-assertive interpretation. Whereas the classical versions of skepticism were mostly based on "arguments" concerning the unreliable character of sense perception (with the exception of Agrippa's attempts to show that reason itself is unreliable—see Brush, pp. 10-13), the deconstructionist "arguments" are all
questions might be. Deconstructionists obviously cannot say that their questions could lead to insights into the nature of their own or other theories, or that they could lead to insights into the nature of meaning, or truth, or logic—even that they could lead to the insight that traditional assumptions about meaning, truth, and logic are illusory—because all such talk about gaining insights into various phenomena presupposes the very sorts of claims that they reject; a person can, for example, gain the insight that traditional assumptions about meaning, truth, and logic are illusory only if it is true that such assumptions are illusory. In other words, a deconstructionist cannot claim that his "interrogative" texts are valuable and, therefore, merit consideration because they might lead to any kind of knowledge, insights, truths, etc.—but what other reasons could there be for thinking that a text merits

ultimately based on their views about language. Thus, there appear to be some interesting historical differences as well as similarities. Deconstructionism constitutes a form of skepticism whose roots are firmly anchored in twentieth-century philosophy's preoccupation with the nature of language—what is sometimes referred to as "the linguistic turn" (see, for example, Blackburn, pp. 5-6). Deconstructionism might even be the first form of skepticism based primarily on linguistic concerns. Of course, the foregoing remark represents no more that a tentative hypothesis whose confirmation would require an exhaustive study of the entire history of skepticism. Nevertheless, it appears to be a hypothesis worth investigating.
consideration? What other reasons could a deconstructionist give when asked why other critics should concern themselves with such interrogative texts? Perhaps he might say that the texts are valuable and, therefore, deserve attention because they are capable of producing pleasurable experiences when read; or he might perhaps simply say that other critics should concern themselves with the texts because he would personally prefer that they do so. If, however, the question of whether deconstructionist texts merit consideration is to be determined solely on the basis of whether or not they produce pleasurable experiences in a given reader, or on the basis of personal preference, then a rival critic will have every right to completely ignore those texts solely because they happen not to give him pleasure, the

22 The preceding sorts of questions possess an additional relevance where dramatic works are concerned because the merits of dramatic theories have often been assessed partly on the basis of whether they imply interpretive strategies that can be profitably used in transferring a play from the written page to the stage. When asked how the deconstructionist approach will help a director determine the interpretation or the set of interpretations that will provide the most appropriate basis for the performance of a given play, the deconstructionist can apparently answer in only one of two ways. A thoroughly "interrogative" deconstructionist can say only that he does not know, and an "assertive" deconstructionist can say only that any interpretive strategy that is based on an attempt to figure out the meaning of the written script will constitute an exercise in futility—since the script has no determinate meaning. Needless to say, a director is not likely to find either of the preceding answers particularly helpful.
pleasantness of an experience lying as it does in the eye of the beholder, or solely because that is what he prefers to do. Either way, the question of whether a particular set of texts merit consideration will ultimately boil down to a matter of personal preference—much like the question of which flavor of ice cream a person prefers.

During his introduction to *Margins of Philosophy*, Derrida himself appears to raise the very issues that have just been discussed:

Certainly one will never prove philosophically that one has to transform a given situation and proceed to an effective deconstruction in order to leave irreversible marks... With what is one to authorize oneself, in the last analysis, if not once more with philosophy, in order to disqualify naivete, incompetence, or misconstrual, in order to be concerned with passivity or to limit pleasure?23

Thus, Derrida apparently realizes that references to his own personal preferences are not likely to persuade others that they should pursue the course of deconstructionism or even that deconstructionist texts merit consideration. The only means of doing that would require him to embrace the very assumptions of Western philosophical thought—assumptions

about meaning, truth, and logic—that he wishes to reject.

It is now possible to state the second reason why it does not seem necessary to consider what deconstructionists have had to say about each of the Structuralist concepts and principles that are discussed in the present study. If the deconstructionist discourses concerning a particular concept or principle are viewed as an attempt to present arguments, it will be possible to show that those arguments cannot possibly succeed in proving what they set out to prove. Since they are destined to fail, there would appear to be little reason for considering them. If, on the other hand, the discourses in question are taken to be interrogative or imperative in character, deconstructionists themselves will be unable to provide any reasons, beyond those having to do with personal preferences, why those discourses merit any consideration. Either way, there would appear to be little to be gained by considering what deconstructionists have had to say about the various topics discussed during the course of the present work, and that is why deconstructionist attitudes toward a given concept or principle are discussed only infrequently.
APPENDIX B

One of the additional issues that arises in connection with Saussure's views concerning the units and principles of langue concerns the relation between those views and two recent linguistic doctrines which Lyons refers to as "the doctrine of linguistic relativism" and "the doctrine of linguistic universalism." According to Lyons, the former doctrine, which is widely associated with the writings of Benjamin Lee Whorf, asserts that "the actualization of particular phonological, grammatical and semantic distinctions in different language-systems is completely arbitrary."

---

1 Lyons, p. 245. Unfortunately, the extent to which the foregoing statement accurately characterizes Whorf's own views is unclear. It is certainly true that throughout his writings Whorf repeatedly emphasized the existence of significant phonological, grammatical, and semantic differences between languages, and he clearly thought that the existence of such phenomena constitutes good evidence for thinking that the distinctions established within any given language are at least somewhat arbitrary—see, for example, the collection of his writings entitled Language, Thought, and Reality, ed. John B. Carroll (Cambridge: MIT Press, 1956), pp. 134-59, 246-70. However, it is not at all clear that he thought that all such distinctions are completely arbitrary.

574
Although Lyons distinguishes several different versions of linguistic universalism, only the weakest version will be relevant to the following remarks; it simply asserts the denial of the foregoing thesis—it is not the case that the actualization of all particular phonological, grammatical and semantic distinctions in different language-systems is completely arbitrary.²

Lyons points out that a number of recent studies have challenged the plausibility of the thesis of linguistic relativism—and have done so with respect to the sorts of examples that were previously thought to constitute the best evidence that the thesis is true—variation in the color concepts found within different languages.³ For example, after surveying the color terms employed within a wide variety of languages, Brent Berlin and Paul Kay concluded that whenever two or more languages employ the same number of basic color terms, those terms will all focus on the same part of the color spectrum: all languages with only two basic color terms have words whose focal points are in the areas of black and white; all languages with only three basic color terms have words for black, white, and red; all

languages with only four color terms have words for black, white, red and either green or yellow, etc. Thus, while the basic color concepts employed within any given language may be said to be arbitrary to the extent that some languages employ only two such concepts, some three, some four, etc., the choice of such concepts does not appear to be completely arbitrary, as the thesis of linguistic relativism maintains it is.

While Lyons is careful to point out that there are formulations of the Structuralist approach that are "compatible with various kinds of universalism," versions which do not "necessarily imply an acceptance of the doctrine of linguistic relativism," he repeatedly suggests that the foregoing does not hold true in the case of Saussure. Culler also seems to suggest that Saussure was committed to the view that the linguistic division of sound and thought is completely arbitrary. While the contents of the Course certainly lay heavy stress on the phenomena of acoustic and conceptual variation across languages, Saussure never says that all acoustic and conceptual distinctions are completely

---


5 Lyons, Semantics, pp. 245, 247.

6 Culler, Saussure, pp. 24-29.
arbitrary. Given the fact that Saussure never explicitly commits himself to such a strong view, and in light of the highly implausible nature of that view, it does not seem entirely fair to attribute it to him. The present study will, therefore, assume that Saussure's remarks are not incompatible with a version of universalism based on the claim that there are certain uniformities both in the way human beings are psychologically structured and in the sorts of environments to which humans are normally exposed that together place at least some restrictions on the sorts of phonological, grammatical, and semantic distinctions that are actualized in natural languages.  

The doctrine of linguistic relativism is not the only thesis attributed to Whorf that is relevant to the present discussion of Saussure. Whorf is also associated with what John R. Anderson calls "the thesis of linguistic determinism"—"the claim that language determines or strongly influences the way one thinks or perceives the world." As

7 A view of the above sort is discussed in some detail in Lyons, *Semantics*, pp. 246-49.

8 John R. Anderson, *Cognitive Psychology and Its Implications* (San Francisco: W. H. Freeman & Co., 1980), p. 384. Unlike the doctrine of linguistic relativism, it is clear that Whorf did, in fact, maintain at least a weak version of the thesis of linguistic determinism; while some of his remarks seem to assert the strong claim that a person's language completely determines the way he thinks and
Anderson points out, recent studies have also cast doubt on its plausibility. For example, Eleanor Rosch performed a series of experiments designed to determine whether speakers of Dani, a language that has only two basic color terms, are significantly less adept than English speakers at picking out the "focal colors" associated with the eleven basic color words of English. In one experiment, she tested the respective speakers' ability to learn nonsense names for focal versus nonfocal colors; and in another, she tested their ability to select a focal versus a nonfocal color chip they had been briefly shown from a set of 160 color chips. Contrary to what the thesis of linguistic determinism would lead a person to expect, speakers of Dani, like speakers of English, found it significantly easier to reidentify and learn nonsense words for focal colors than nonfocal colors—in spite of the fact that the Dani language contains no basic terms for those focal colors.

perceives the world (see, for example, Whorf, pp. 252, 256-57), in other remarks he seems to retreat to the weaker claim that a person's language strongly influences the way he thinks and perceives the world (see, for example, p. 138).

9 Summarized in Anderson, pp. 385-86.
10 Summarized in Anderson, pp. 385-86.
The question of whether Saussure's remarks commit him to the thesis of linguistic determinism would appear to be largely dependent on whether a person's concepts completely determine the way he thinks and perceives the world. As earlier remarks have shown, Saussure insisted that the existence of a particular concept or signified is dependent on its association with a distinct signifier; prior to the association of thought and sound in language, there are no distinct concepts. That claim, which will subsequently be referred to as the "thesis of conceptual dependency," appears to imply that a monolingual speaker's language determines the set of concepts he can be said to possess. However, the thesis of conceptual dependency would entail the thesis of linguistic determinism only if a person's concepts, in turn, completely determine the way he thinks and perceives the world. It is not, however, clear that such is the case.

In his discussion of a view which he calls "psychological nominalism," Wilfred Sellars makes a claim that is quite similar to—if not the same as—the thesis of conceptual dependency: "all awareness of sorts, resemblances, 

11 See above, p. 146-47.
facts, etc... is a linguistic affair." As Richard Rorty notes in a discussion of Sellars' view, there appear to be obvious counter-examples to the preceding claim, viz., the existence of certain perceptual states, e.g., "pains, whatever feelings babies have when looking at colored objects, etc." In order to avoid such counter-examples, says Rorty, Sellars invokes a distinction between "awareness-as-discriminative-behavior" and other more sophisticated forms of awareness.

Although Saussure does not himself explicitly draw a distinction between conceptual awareness and non-linguistic, discriminatory awareness, he does not say anything that would rule out such a distinction. In the interests of charity, the present study will assume that Saussure would have found such a distinction agreeable had he thought about it. So doing will increase the plausibility of Saussure's views in two respects: 1. it will allow him to avoid being committed to the highly implausible thesis of linguistic

14 Rorty, p. 182.
determinism, and 2, it will provide a defense against potential counter-examples to his thesis of conceptual dependency.

The next set of issues that require discussion concern the mentalistic character of Saussure's approach. Following the advent of behaviorism in the early decades of the twentieth-century, mentalistic approaches became the target of widespread attacks, attacks which Jerry A. Fodor suggests were largely motivated by a belief that mentalism entails a commitment to some form of mind-body dualism. During the last several decades, a number of writers have, however, argued that the difference between mentalistic and anti-mentalistic approaches lies not in any difference in ontological commitments, but lies instead in a tendency to employ levels of analysis that differ in their degree of abstractness.

Computers are commonly used to illustrate the possibility of utilizing levels of analysis that differ with respect to abstractness. When attempting to account for the behavior of a computer, a person could either try to formulate

---


16 See, for example, Chomsky, Aspects, pp. 193-94 and Rules, p. 226.
hypotheses about the kind of program it is running, or he could try to formulate hypotheses about the actual physical processes that are occurring in the hardware of the machine as it is running that program. The former enterprise would involve an abstract level of analysis; the latter would involve a more concrete level of analysis.

As Chomsky observes, the possibility of utilizing more and less abstract levels of analysis is, however, by no means limited to the case of computers:

In the study of any organism or machine, we may distinguish between the abstract investigation of the principles by which it operates and the study of the physical realization of the processes and components postulated in the abstract investigation. 17

The primary difference between mentalist and anti-mentalist approaches, suggests Chomsky, is simply that the former employ an abstract level of analysis in their study of intelligent behavior, whereas the latter prefer a more concrete level of analysis. 18

As Anderson observes, difference in the level of analysis employed also constitutes one of the primary differences between cognitive psychology and physiological

17 Chomsky, Rules, p. 226.
18 Chomsky, Aspects, pp. 193-94.
psychology: "Cognitive psychology is to physiological psychology much as computer science is to electrical engineering." To say that cognitive psychology employs a highly abstract level of analysis is another way of saying that it adopts a thoroughly mentalistic approach.

As Chomsky points out, the fact that mentalistic approaches employ an abstract level of analysis does not, however, entail any sort of commitment to dualism:

We may speak of the abstract study of human intellectual capacities and their functioning as the study of mind, without thereby implying that there is a res cogitans as a "second substance" apart from the body.

Just as a computer programmer's talk about abstract programs does not commit him to the idea that there must be some sort of immaterial substance which constitutes the ontological locus of those programs, so a mentalist's talk about concepts, ideas, etc. in no way commits him to some sort of mind-body dualism.

Saussure was, moreover, quite explicit about the issue in question: "Linguistic signs, though basically psychological . . . are realities that have their seat in the

---

19 Anderson, pp. 11-12.

20 Chomsky, Rules, p. 227.
brain." The preceding is just one of many passages in which Saussure insists that the two components of the linguistic sign are somehow realized in the neuro-chemical structure of the brain. Given what was known about neurophysiology at the time, Saussure did not really have much of a choice about whether to employ an abstract or a concrete level of analysis.

Mentalistic theories of language have, however, been criticized on grounds other than the ones just discussed. For example, as Lyons observes, the traditional view of language dating from the Middle Ages, which is mentioned in Chapter Three, tends to define the meaning of a linguistic expression in terms of the mental entity—be it a concept, or image, or idea, etc.—that is typically associated with the expression in the minds of competent speaker-hearers. Over the course of the last fifty years or so, that aspect of traditional mentalistic theories has been widely challenged—and with a good deal of

21 Saussure, p. 15.
22 See also Saussure, pp. 14, 19, 66.
success. As earlier remarks have indicated, Saussure's analysis resembles the traditional account in drawing a sharp distinction between the mental entity associated with a word-form and the referent of that form. Does his analysis also follow the traditional view in maintaining that the signified component of a linguistic sign encapsulates the meaning of that sign?

Due to a shift in Saussure's terminology, the answer to the foregoing question is not entirely clear. At one point near the beginning of the Course, Saussure does refer to the two components of the sign as sound-images and meanings ["sens"]. However, that remark may well have occurred prior to what the translator indicates was a shift in Saussure's terminology, and throughout most of the Course, Saussure employs the term "concept" ["concept"]. At no


25 See above, p. 142.

26 Saussure, p. 15.

27 Saussure, p. 75.
point does Saussure explicitly equate signifieds or concepts with meaning. Moreover, it appears that Saussure's analysis of signs was not motivated by a desire to answer the question "What is meaning?" He was, instead, trying to answer the question "What underlies a given community's regular use of linguistic forms?" His answer was that the members of a community have implicit knowledge of acoustic images that correspond to the various linguistic forms and that those images are, in turn, associated with concepts of some sort. In view of his primary interests, the present study will assume that Saussure was not committed to the claim that the meaning of a sign is identical with its signified component.

As Lyons observes, the traditional view also maintains that the concept or idea associated with a word-form determines its referent or extension—the set of things to which the word-form can be correctly applied. However, that assumption has also been challenged by recent writers. For example, in his paper "The Meaning of 'Meaning,'" Hilary Putnam imagines a science-fiction situation in which there is somewhere in the galaxy a planet called "Twin Earth" that is exactly like Earth except that what speakers there refer

to as "water" is not composed of hydrogen and oxygen, but has a molecular structure which Putnam abbreviates as XYZ. In all other respects, XYZ is to be indiscernable from $H_2O$: it has the same appearance and taste, it quenches thirst, it fills the lakes and oceans, the rain is made of XYZ, etc.

Consider, says Putnam, a typical Earthian English speaker Oscar$_1$ and a typical Twin Earthian English speaker Oscar$_2$, both of whom lived prior to the advent of chemical theory and who were exact duplicates in appearance, feelings, thoughts, interior monologues, etc. By hypothesis whatever psychological state Oscar$_1$ was in at any given time he thought about or referred to water, Oscar$_2$ was in precisely the same state at that time. Nevertheless, says Putnam, their tokenings of the word did not have the same extension. Putnam therefore concludes that "the extension of the term 'water' . . . is not a function of the psychological state of the speaker by itself"--which means that the extension of a word-form is not determined exclusively by the concept or idea that is typically associated with it in the minds of competent individuals, as the traditional

---

mentalistic theory of language assumed it is.\textsuperscript{30}

Saussure says very little about the phenomenon of reference in the Course so it is difficult to say whether he thought that the signified by itself determines the extension of its corresponding signifier. There is, however, a larger question raised by Putnam's example. Remarks contained in the penultimate paragraph suggested that Saussure was primarily interested in trying to answer the question "What underlies a community's regular use of linguistic forms?" Briefly put, his answer was that the members of the community possess certain sorts of implicit knowledge. However, Putnam's remarks seem to suggest that such knowledge by itself is not sufficient to account for all aspects of linguistic behavior. The verbal behavior manifested by the members of a linguistic community derives not only from the fact that they possess certain sorts of implicit knowledge, but also from the way they are situated in the world—various facts about their environment, facts

\textsuperscript{30} Putnam, p. 224. In opposition to Putnam, it might be argued that the extension of "water" depends on the observable characteristics rather than the underlying microstructural properties of the liquid in question. It seems probable that less objectionable examples could, however, be found.
about the ways in which they are causally related to that environment, etc. Saussure never denies that the foregoing is the case, but he does not discuss it either, and that is certainly a major omission. In view of the preceding remarks, Saussure should perhaps be construed not as trying to account for all of the factors that underlie a community's regular use of linguistic forms—only those factors that have to do with the implicit knowledge possessed by its members.

31 Putnam, pp. 229-35. The notion that the referent of many types of linguistic expressions is largely determined by the way its user is causally situated in the world has recently led to the development of the so-called "causal theory of reference." See, for example, Michael Devitt's book Designation (New York: Columbia University Press, 1981).
BIBLIOGRAPHY


---------. "How to Do Things with Texts." Partisan Review, 46 (1979), 569-75.


--------.  "Thinking about Non-Being."  Inquiry, 16, no. 3 (1973), 290-312.


---------, "From Morpheme to Utterance." *Language*, 22 (1946), 161-83.


