Offprint from

SEMIOTICA

12: 2, 1974

edited by

THOMAS A. SEBEOK

MOUTON

THE HAGUE - PARIS

JOHN M. LIPSKI

A Topology of Semantic Dependence

I. INTRODUCTION

The history of linguistic investigation has been characterized by an increasingly wide scope of inquiry. Linguistics scored a notable achievement in the rigorous description of natural languages by moving from the study of the individual word to the complete sentence. Contemporary linguistics has broadened the scope even further, to include data concerning the structure of entire DISCOURSES. While the study of discourse-grammars or TEXT-GRAMMARS is still in its infancy, it has, thanks to the work of a number of investigators, progressed to the point where significant model-construction may be encouraged. In particular, it has been possible in many cases to characterize some of the formal properties inherent in such text-grammars, with the result that it is now feasible to pose a number of more specific questions concerning the mathematical structures contained in these grammars. Since mathematics is merely a formalization and axiomatization of the same human cognitive capacity manifested in natural language, nearly every branch of mathematics has been tentatively extended to the domain of linguistic theory. With a few noteworthy exceptions, however, most of these formal extensions have been quite rudimentary, sketching the structure of analogies and metaphors holding between mathematics and natural languages.[1] The way has, nonetheless, been paved for more comprehensive analyses of the

1 A typical example of the metaphorical approach is found in the paper of Schoeck (1968).

mathematical characteristics of human language, with the ultimate goal of achieving a complete synthesis between mathematics and linguistic science. The path towards this synthesis is of necessity taken by means of small steps, each serving as a launching point for additional research and consideration. The remainder of this paper is offered as just such a step, by entertaining the possibility of applying to the study of text grammars that branch of mathematics which deals most directly with spatial representations, namely topology. As indicated by the title, the primary subject matter concerns the notion of semantic dependence in a text or discourse, and the goal of this study will be a rudimentary attempt at extending a formal topological model to the theory of text-grammars. Rather than elaborating on more general characteristics of topological models, a task better relegated to a complete study of its own, it is the intention of this paper to propose a specific model, together with suggestions for its possible implementation. Since the foremost objective is the construction of a formal model, detailed discussion of specific examples must be postponed until such a time as the theoretical details of the model can be worked out. The remainder of this paper will therefore be divided into the following sections: Section II will offer a general discussion of semantic dependence, Section III will develop a basic topological model which will then be further considered in Section IV, while Section V will contain concluding remarks.

II. DEFINING SEMANTIC DEPENDENCE

2.1 It is clear that the very choice of terms used in qualifying an expanse of linguistic production as a unified text or coherent discourse implies a certain definable degree of coherency between the various logico-semantic structures comprising the text. In general, this means that the semantic interpretation of certain sequences is in some way dependent upon the semantic interpretation of other sequences occurring in the text. More specifically, Irena Bellert (1970: 335) has offered the following characterization: "A discourse is a sequence of utterances S_1, S_2, ..., S_n, such that the semantic interpretation of each utterance S_i (for $2 \leq i \leq n$) is dependent on the interpretation of the sequence S_1, ..., S_{i-1}. In other words, an adequate interpretation of an utterance occurring in a discourse requires the knowledge of the preceding context ... the semantic interpretation of any utterance is the set of consequences or conclusions

that can be drawn from that utterance." Bellert then goes on to discuss some specific cases, which will be returned to below. Of key interest at the present point is the fundamental role which semantic dependence relations play in the definition of textual coherence. Basic in this regard is the manner in which a text or discourse is generated by a text-grammar. As defined by Teun A. van Dijk (1973: 60), "a text is an ordered n-tuple of sentences ordered by the relation of preceding: *Prec* (S_{i-1}, S_i), defined over unanalyzed sentences". Thus, it is feasible, under these definitions of textual structure, to speak in terms of a formal relation of semantic dependence defined on successive pairs of sentences (S_i, S_j) occurring in a given text or discourse *D*. Let us therefore denote by $\rightarrow$ the relation of semantic dependence in the discourse, with $b \rightarrow a$ reading '*b* is semantically dependent on *a*'. It now remains to formulate a more precise definition of semantic dependence so that some empirical correlates can be added to this purely formal classification.

Intuitively, it seems clear that for a sentence S_j to be semantically dependent on another sentence S_i, the latter sentence must contain some information necessary for a complete, unequivocal interpretation of S_j. Bellert (1970: 336-37) believes that the answer lies in repetition of lexical items in an overlapping fashion across the expanse of the text: "The LS [logico-semantic] structure of each utterance S_i of a text S_1 ... S_n is such that at least one lexical item which occurs in it, or one proposition which can be inferred from it, can also be found within the utterances of the sequence S_1, ..., S_{i-1}, or within the propositions which can be inferred from these." Clearly, some sort of information repetition must take place in a coherent text, and the manner in which this repetition is effected may be taken to constitute a working definition of semantic dependence. The characterization offered above, while doubtless correct, is insufficiently detailed to permit a deeper inquiry into the matter of textual semantic interrelations; as a consequence, it will be necessary to consider in somewhat greater depth several specific aspects of semantic dependence. What will be offered is not an exhaustive classification, but merely an enumeration of the most obvious characteristics of semantic dependence, from which more subtle cases may be extrapolated. From the considerations to be given below, it will be possible to inductively infer a general definition of semantic dependence, which may be adopted as a working definition in lieu of a deductively generated definition.

2.2 The discussion may logically commence by considering the simplest case of semantic dependence, given as:

CASE 1: Given a text D and given two sentences S_i and S_j in D, $S_j \rightarrow S_i$ if S_j makes use of a linguistic index φ whose extralinguistic referent is expressed in, or may be inferred from, S_i.

Here the term 'linguistic index' is taken in the sense used by Bellert (1970: 346): "those devices of language whose function is to point to extralinguistic 'objects' or particulars that the hearer is supposed to identify in accordance with the instruction contained in the index and the situational or linguistic context." A linguistic index may refer either to another index, as in the case of certain pronouns or adverbs, or it may refer directly to an extralinguistic object or event, as in the case of nouns. When an index used in one sentence refers back to an index used in another sentence, the formerly-mentioned sentence is of necessity semantically dependent on the latter, since the latter sentence contains the extralinguistic reference missing in the first-mentioned sentence. Consider, as an example, the following two sentences:

(1) *David often visited New York*
(2) *He really didn't like it there*

In this example, sentence (2) is clearly dependent on (1), for (2) contains two indices whose extralinguistic referents are expressed in (1): *he*, corresponding to *David*, and *there*, corresponding to *New York*. Without the additional semantic information supplied in (1), sentence (2) may not properly occur as part of a coherent text, since it violates the requirement that lexical material from other sentences must be repeated. From a logical point of view, moreover, the utterance of sentence (2) in a text without a sentence like (1) occurring elsewhere in the text would constitute an inappropriate use of a linguistic index, since although the EXISTENCE of the referents of *he* and *there* may be taken for granted, there are no provisions for the UNIQUENESS of the referents. Hence, from a formal point of view, the indices cannot qualify as unit functions, and from a linguistic point of view, the sentence is semantically meaningless taken in isolation.[2] In terms of function, linguistic indices are referential and non-predicative, thus being comparable on an abstract level with indexing functions as used in mathematics.

Often, of course, even the use of a noun as a linguistic index is not sufficient to ensure uniqueness of reference. For example, in sentence (1) it is assumed that the listener knows that unique individual which is

[2] Cf. the discussion in Bellert (1970: 347) concerning the use of unit functions in logic and in linguistics.

referred to by the index *David*; taken in isolation, however, the noun *David* is not unique. Thus, if sentence (1) is addressed to a listener whose personal knowledge does not permit him to uniquely identify the referent of *David*, this sentence must be accompanied in the discourse by another sentence which uniquely identifies the *David* being referred to. In such a case, sentence (1) is semantically dependent upon any sentence which serves to ensure the uniqueness of reference by the index *David*.

The presence of definite articles may also induce relations of semantic dependence, for the use of such an article presupposes that the listener is already aware of the unique object to which it refers. Just as a brief illustration, consider the following group of sentences:

(3) *Sam bought a new screwdriver*
(4) *The screwdriver was poorly made*
(5) *It broke the first time it was used*

Sentence (5) is clearly dependent on (4), by virtue of the pronominal replacement of *the screwdriver* by *it*. Notice, however, that (4) is also semantically dependent on (3), since (3) justifies the use of the definite article found in (4). Naturally, the mere use of a definite article in a sentence does not entail the presence of another sentence which justifies the use of the definite form, since the justification may be purely internal to the sentence, either explicitly or implicitly. For instance, (3) and (4) could be combined to yield

(6) *The screwdriver that Sam bought was poorly made*

where the justification for the use of the definite article is internally supplied by the clause *that Sam bought*. Implicit justification of the article occurs whenever it may be safely assumed that the listener can uniquely identify the object in question, for example in common phrases such as *the flag, the Pope, the Queen, the President, the pill, the bomb, the needle, the weather*, etc., or in the case of nouns taken in a generic sense, e.g. *the heart, the mother, the definite article*, etc., and where further qualification permits accurate identification, as in *the capitol of Spain, the King of the Blues*, and so forth. Closely tied to such examples are cases in which the justification for the use of the definite article may be INFERRED from another sentence by means of rules of deduction presupposed to be shared by both the speaker and the hearer. In the following pair of sentences:

(7) *Sam's screwdriver was poorly made*
(8) *The handle fell off the first time it was used*

the definite article construction *the handle* in (8) is a consequence of the presupposed knowledge that EVERY SCREWDRIVER HAS ONE (AND ONLY ONE) HANDLE. Thus, the index *handle* occurring in (8) may be assumed to specify a unique referent, due to the implications contained in (7).

In sentences (1) and (2), the required extralinguistic referents of the indices are explicitly stated in sentence (1), in this case by means of proper nouns. Sentences (7) and (8) have shown, however, that it is also possible for the extralinguistic referent of a linguistic index to be available only by rules of semantic projection including knowledge presupposed to be possessed by the listener. As a further example of this latter form of semantic dependence, consider the following pair of utterances:

(9) *It's hotter than you-know-what in this room*
(10) *That's too hot for me*

In sentence (10), *that* refers to the expression *hotter than you-know-what* in (3). However, anyone uttering (3) would presuppose on the part of his listener the knowledge that the semantically meaningless expression *you-know-what* is really a euphemism for *hell*, when found in such superficial discourse-frames as *hotter than*___, *harder than* ___, etc. Thus, from the standpoint of pragmatics, *you-know-what* in (3) is taken to 'refer' to the semantic concept of {hell}, with the result that the index *that* in (4) ultimately refers to the semantic representation of {hotter than hell}, via assumed rules of semantic projection. Hence, the semantic dependence in this case is determined not by direct reference, but by the intermediate step of a presupposed mutual knowledge shared by the speaker and the listener. Such mutual knowledge may also take the form of fixed grammatical rules; for example, the use of the masculine singular pronoun *he* in (2) to refer to *David* follows from the fact that *David* ordinarily receives the lexical specifications [animate], [masculine], [singular], etc. Similarly, the use of *there* in (2) to refer to *New York* is a consequence of the locative function of the slot in which the adverbs occur, as well as the locative usage of *New York* in (1). Such purely grammatical dependency relations are obviously not confined to pronouns but may occur in any situation including grammatical features. For example, relationships of verb tense, mood, and aspect, and of gender and number, as well as other overt grammatical markings which are found among the world's languages, may all be extended from sentence to sentence, in each case resulting in an instance of semantic dependence. Not only may such relations hold over successive pairs of utterances, but often features of the entire discourse are critical in determining inter-

sentential dependence; for instance, the tense specification of a narrative passage. In this paper, no account of such global textual features will be offered, although such an account must be included in a complete theory of text grammars.[3] Since the task of this study is to formulate a model based on the configurations of the individual utterances comprising a given text, intersentential dependence relations will be the only ones to be covered in detail.

It is also possible for semantic dependence of linguistic indices to involve logical rules of deduction or pragmatic rules of assumption, as noted above. As an example, the use of the adverb *there* instead of *here* in sentence (2) indicates that the utterer of (2) is not in New York at the time the sentence is uttered, by virtue of the pragmatic rules governing the use of these adverbs in English. Even if both (1) and (2) are uttered by the same speaker, however, the presupposition in (2) does not entail the semantic dependence of (2) on (1), except for the dependence of the pronominal referents. However, this situation changes if sentence (1) is replaced by

(11) *David often visited us here in New York*

Now, in addition to the semantic dependence induced by the use of the pronouns, sentence (2) is dependent on (11) since (11) implies that the speaker is in New York at the time the sentence is uttered, while (2) clearly implicates a speaker not in New York. Furthermore, given the information contained in (11), the index *there* in (2) may be interpreted as referring to either the semantic representation {in New York, where you are}, or to the representation of {in New York, where I was}, depending on whether both sentences were uttered by the same or by different speakers. Now let us consider the following pair of sentences:

(12) *Harold's mother has gray hair*
(13) *The woman is only forty years old*

Sentence (13) contains the index *the woman*, which refers to *Harold's mother* occurring in (12). In this case, the use of the index *woman* in (13) follows from the implicational formula

(14) $\forall x$ (x is human, x is a mother) $\supset$ (x is a woman)

thus illustrating another possibility for the semantic dependence of linguistic indices. This form of dependence relation, based on implica-

3 For a number of additional considerations, see Van Dijk (1972a: Chap. 3).

tional formulae, is somewhat different from a purely grammatically-conditioned dependence, of the type governing the use of *he* in sentence (2). In the latter case, the choice of the index is merely a function of the lexical specification of the referent *David*. On the other hand, in the case of sentence (13), the choice of the noun *woman* requires more than the simple lexical features of *mother*, which include such specifications as [human], [female], etc.; what is needed in fact is an implicational statement like (14) to permit the replacement of one noun by another fully specified noun, much in the same fashion as the implicational formula inherent in (7) permitted the insertion of the definite article and noun in (8). If, however, sentence (13) contained the pronoun *she* instead of the noun *woman*, the semantic dependence would revert to the purely grammatical type, in addition, of course, to the purely referential dependence required by the use of the pronoun.

2.3 The next general type of semantic dependence may be covered by the following characterization:

CASE 2: Given a text D and two sentences S_i and S_j in D, $S_j \rightarrow S_i$ if S_i contains information which serves to resolve an ambiguity present in S_j.

Here the notion of semantic dependence must be defined on the level of surface structures, for on the underlying level there will presumably be no ambiguity, at least from the point of view of the speaker. Without going deeply into the question of precisely what constitutes ambiguity, one may follow the generally accepted formal division into structural ambiguity and lexical ambiguity. Cases involving structural ambiguity are well known from discussions of transformational syntax; one need only consider sentences of the *flying planes can be dangerous* variety to visualize the manner in which semantic dependence may be realized through intersentential disambiguation based on ambiguities of this nature. Purely lexical ambiguity is also readily exemplified, and the manner in which semantic dependence may be realized through this process may be described as simply specifying a single, unique meaning out of a set of possible variant readings. For instance, the following sentence is lexically ambiguous:

(15) *Herbert brought the case to his father*

In this sentence, the noun *case* has a variety of possible readings, as may be verified by looking in a dictionary, and virtually all of these readings

are possible within the context of sentence (16) taken in isolation. This ambiguity would be resolved by the cooccurrence in the same text of a sentence like one of the following:

(16) *Herbert wanted his father, who was a carpenter, to fix his violin case*
(17) *Herbert wanted his father, who was a lawyer, to help him in a legal matter*
(18) *Herbert wanted his father to replace the glass in the trophy case*

By an appropriate choice of such sentences, all but one of the possible readings for *case* are removed from the text, although sentence (15) of course remains ambiguous taken by itself. The semantic dependence of (15) on a sentence like (16)-(18) creates a non-ambiguous textual macrostructure.

2.4 A further instance of semantic dependence may be described as follows:

CASE 3: Given two sentences S_i and S_j in a text D, $S_j \rightarrow S_i$ if S_j contains an elliptical construction which is fully specified in, or may be inferred from, S_i.

Elliptical constructions represent a very common means of inducing semantic dependence by permitting information to be repeated through inference rather than be explicit duplication. This may be illustrated by the following two sentences:

(19) *Dick likes detective novels*
(20) *I wonder why*

The semantic deep structure of (20) is roughly 'I wonder why Dick likes detective novels', with the final clause being deleted from the superficial representation by a transformation of ellipsis. In order for this transformation to operate, the clause to be deleted must be explicitly stated in the preceding context, thus highlighting the properties of semantic dependence inherent in the process of ellipsis. Taken independently of the surrounding context, a sentence like (20) is anomalous, since it presupposes previous knowledge of information which has obviously been deleted. No coherent text may contain an instance of such a sentence without also containing an instance of an accompanying sentence like (19), and consequently any expanse of text containing (30) but not (19) is necessarily incoherent, although the appearance of incoherency may be removed by considering a larger textual fragment. This matter will be returned to below.

2.5 Another major-type-form of semantic dependence relations to be considered here deals with the notion of implication or conclusion which may be derived from a sentence. The concept of deriving conclusions from utterances of natural languages is not identical to the process of deriving conclusions in a formalized logical language, but the analogy is a close one.[4] Bellert (1970) speaks of two types of rules of inference which may be used to extract conclusions from linguistic data: rules based on deductive reasoning, which may be properly considered as part of the logico-semantic structure of language, and rules based on inductive reasoning, involving the listener's internalized knowledge of extralinguistic reality. An example of the former category is provided by the following pair of sentences:

(21) *Benjamin died yesterday afternoon*
(22) *Benjamin was alive yesterday morning*

Sentence (22) follows from the logico-semantic specifications of the verb *die*, which must include an implicational formula of the form

(23) (x died at time t) $\supset$ (x was alive before time t)

and also from the characterization of *morning of day X* being prior in time to *afternoon of day X*. Bellert has referred to words whose lexico-semantic representations include such implicational formulae as IMPLICATIONAL WORDS. Such words logically induce semantic dependence between the sentences in which they occur and any sentences containing the conclusions which may be derived from them. It is important to note, however, that the concept of semantic dependence is in no way equivalent to the concept of logical implication. If, given the two sentences S_i and S_j, we have the implication $S_i \supset S_j$, we also clearly have $S_j \rightarrow S_i$, but the converse is not necessarily true, as the above examples illustrate. In fact, given the relation $S_i \supset S_j$, it is not usual to find both sentences cooccurrent in the same text, since the presence of S_j would be superfluous.

To visualize conclusions based on inductive reasoning derived from knowledge of the world, consider the following sentence:

(24) *Henry teaches at Yale*

For most speakers aware of the existence of this university, their personal knowledge would enable them to infer the following:

4 See, for example, Bellert (1970: 337-38).

(25) *Henry teaches in the United States*

Listeners with even more specific knowledge would be able to infer sentences like the following, which may be extended to any degree of detail, depending upon the specific knowledge possessed by theindividual listener:

(26) *Henry teaches in Connecticut*
(27) *Henry teaches in New Haven*
(28) *Henry teaches at an Ivy League university*

Here the conclusions which may be drawn are a function of each listener's INDIVIDUAL knowledge of the world, and may not be included as part of the lexical representation for any of the words in (24). Only the word *teach*, by virtue of its implicative properties, permits the direct inference of a sentence similar in overall form to (26)-(28), roughly

(29) *Yale is an institution of learning*

In order to fill in the details, however, inductive reasoning on the part of the listener is required, which accounts for the fact that the formal structure of the text itself may not be expected to yield all the conclusions that such a sentence might provide for a listener with the appropriate extralinguistic knowledge. The indefinitely large number of conclusions which may be inductively inferred from such a sentence is, as it were, independent of the text-grammar generating the sentence, and hence by extension independent of the utterer.[5] Consequently, while a sentence like (25) is in a sense semantically dependent on (24), this dependence is not the sort that can be characterized in terms of a formal text-grammar, unless the statements which permit the inductive reasoning are also present in the text; hence such a relation will not generally be included under the heading of 'semantic dependence' as used in this paper.

There is, in addition to the previously-mentioned cases of inference, another way in which inferential relationships may be used to induce semantic dependence, which may be roughly stated as follows:

CASE 4: Given two sentences S_i and S_j in a text D, $S_j \rightarrow S_i$ if S_j is expected to yield a conclusion which is NOT the result of the presence of an implicative word, and which may NOT normally be assumed to be directly derivable from the listener's personal knowledge, and if this conclusion is expressed in or inferrable from S_i.

5 Bellert (1970: 342-43).

To give an example of this characterization, consider the semantic structure of the following sentence, taken in isolation:

(30) *Since the New York State Police couldn't arrest him there, Rocky felt safe as soon as he crossed into Jersey City*

For most speakers residing in the northeastern United States, sentence (30) appears quite natural, and yields a number of conclusions. For example, a speaker familiar with the metropolitan New York City area will know that Rocky was travelling west from Manhattan, that he crossed the Hudson River via the Lincoln Tunnel and reached the immediately adjacent New Jersey coast; i.e., Jersey City. Moreover, a typical American listener, knowing that under ordinary circumstances state police only have authority in their own state, could conclude, without further geographical knowledge, that Jersey City is not in New York, but rather immediately adjacent to some part of New York. Such a listener would therefore find (30) acceptable, although the geographical subtleties would perhaps be lost on him. If, however, sentence (30) was presented to someone from another country, ignorant both of American geography and of American police systems, it would be unacceptable, and in fact mystifying, for the use of the qualifier *since* presupposes that the listener UNDERSTANDS BEFOREHAND the reasons why one might feel safe from state police upon reaching a given city. Thus, in order for (30) to occur in an acceptable fashion in a text meant for such listeners, it would have to be accompanied in the text with one or more additional sentences which would impart the knowledge necessary for (30) to yield the intended conclusions. As an even more striking instance of this variation of semantic dependence, consider the effects of replacing (30) by

(31) *As soon as he reached Jersey City, Rocky felt safe from the state police*

Anyone uttering (31) would most probably intend his listener to grasp the full significance of this sentence, by understanding that Rocky was escaping from the New York State Police, that Jersey City, while in New Jersey, is immediately adjacent to New York City, that New York State Police have no jurisdiction in New Jersey, etc. However, for a large number of individuals, this information may not be gathered by (31) alone, but requires additional textual material supplying the necessary background material. A sentence like (31) may of course occur in a text without any supplementary sentences which apprise the listener of the situation, but only if the listener may be assumed to already be in possession of the necessary information will the text be coherent on this account; for example, (31) might safely occur in a New York City newspaper. In any event, if such a sentence cooccurs in a text with another sentence which supplies information of the sort referred to above, there will exist by definition a relation of semantic dependence between the two sentences, regardless of whether or not the additional information is actually required by any given listener.

2.6 It is, as noted previously, impossible to exhaust all the possible variations of the property of semantic dependence within the confines of a short review; the cases presented up to this point represent broad, general categories which cover many of the most common cases in which one may speak of a relation of semantic dependence. Each of these categories could be broken down into finer subdivisions, to yield a more detailed and comprehensive characterization of the relation under consideration.[6] Such a task will, however, not be attempted in this paper, for it is only the general properties of semantic dependence which are crucial to the ensuing discussion. Consequently, rather than increasing the detail with which individual cases of semantic dependence are tabulated, this portion of the discussion will conclude with a final classification, somewhat of a catch-all category, which partially overlaps some of the preceding categories, but which also provides for additional categorization:

CASE 5: Given two sentences S_i and S_j in a text D, $S_j \rightarrow S_i$ if S_j contains a presupposition P which is stated in, or may be inferred from, S_i.

I do not intend to participate in the debate over a rigorous definition of presupposition; rather, 'presupposition' as used above may be simply regarded as a condition which must obtain in order for a given utterance to be well-formed. Every well-formed linguistic utterance contains a set of presuppositions, most of them quite trivial, which may or may not occur explicitly in the discourse accompanying the utterance in question. To take a trivial example, given a sentence S containing a linguistic index φ, S contains the presupposition that 'φ exists in some manner', or at least that 'φ may be referred to'.[7] Therefore, while S does not have

[6] An excellent discussion of intersentential relations is found in Van Dijk (1972a: esp. Chap. 2).

[7] Naturally, in order to be referred to by a linguistic index, an object does not have to 'exist' in the real world; we need only consider a sentence such as *flying dragons*

to be accompanied in a text by another sentence which states, in effect, that 'φ exists', if such a sentence, say R, does occur, then $S \rightarrow R$ by definition. To turn to a somewhat less trivial example, note that sentence (32) below contains, among others, the presuppositions which are explicitly expressed in (33)-(34):

(32) *Steve got married yesterday*
(33) *Immediately prior to yesterday, Steve was not married*
(34) *Steve is at least the legal age for marriage*

It would not be at all unusual for sentences like (33) and (34) to cooccur in a text with (32), by supplying information or background, even though (33) and (34) are implicitly embedded in (32) as presuppositions. As noted by Bellert, a critical requirement for the coherence of texts is the successive repetition of information, and one way in which this condition may be effected is by explicitly stating the presuppositions of certain sentences. Therefore, we shall consider this process of repeating presuppositions in explicit form to constitute yet another form of semantic dependence, capable of contributing to the overall semantic unity of a text.

2.7 In the preceding paragraphs an attempt was made to inductively approach a definition of semantic dependence, which could then be used in the construction of a mathematical model. Therefore, let us assume, on the basis of the foregoing discussion, that there is indeed a formally defined RELATION of semantic dependence. Such a methodological assumption is of course not strictly true, but it is unlikely that any new properties of semantic dependence will be discovered which are substantially different from those discussed above. One may therefore consider the property of semantic dependence, denoted by $\rightarrow$, as a formal binary relation, and study its relational properties accordingly. First of all, it may be stated that semantic dependence is essentially a REFLEXIVE relation; that is, relation; that is, one may trivially claim that for every utterance S, $S \rightarrow S$. Given the properties of semantic dependence discussed or inferred above, it also appears that in certain cases, this relation may be considered TRANSITIVE; that is, given three utterances S_i, S_j, and S_k,

can be dangerous. Strawson (1952) notes that "from the fact that something can be referred to no philosophical conclusions follow about its nature". Clearly, however, in order to serve as a linguistic index, a word must be capable of referring to some referent, perhaps only by arbitrary convention, for if no general consensus can be reached regarding the referent of a given word, this word may not properly be regarded as a linguistic index.

it may be that if $S_k \rightarrow S_j$ and $S_j \rightarrow S_i$, then $S_k \rightarrow S_i$. It is in fact clear that a necessary and sufficient condition for the existence of the property of transitivity is that each instance of semantic dependence be based on an IDENTICAL type of semantic linkage. The mere fact of information repetition is not sufficient to ensure semantic dependence of the extreme members of any chain of utterances that might be constructed. The necessary condition for a chain of semantically related sentences to exhibit the property of transitivity is that the information transfer be related to the same basic proposition in each case, so that the elements of the same set of information-units are utilized in forming the dependence relation between each pair of sentences. The precise conditions which must obtain in order to engender the property of transitivity are not clear, and are surely very complex. Since the model to be discussed below makes no use of the property of transitivity in dealing with strings of related utterances, no attempt will be made to state these conditions. Just as an illustration, however, consider the following groups of semantically related sentences, the first of which exhibits transitivity, and the second of which does not:

(35) (i) *Professor Jones is my father*
(ii) *He teaches at Harvard*
(iii) *He enjoys the job very much*

(36) (i) *My cousin wrote me a card*
(ii) *He visited Maine last summer*
(iii) *The water there is very cold*

It is also apparent that the relation of semantic dependence is not symmetric; that is, if $S_j \rightarrow S_i$, then it is not necessarily the case that $S_i \rightarrow S_j$, and *vice versa*, even in cases involving repetition of the same basic information. Thus, even when transitivity obtains, the relation of semantic dependence is not an equivalence relation. Also, since, as noted above, semantic dependence is not to be equated with logical inference, it is not possible to consider any form of 'biconditional equivalence' between two sentences; i.e., if $S_i \rightarrow S_j$ and $S_j \rightarrow S_i$, there are no general grounds for claiming that S_i and S_j are 'semantically equivalent', since the dependence may be based on different parameters in each direction. In such a case, one may only say that S_i and S_j are 'semantically interdependent' with any further conclusions only being possible by making reference to the particular sentences. On the other hand, the relation of semantic dependence is not strictly asymmetrical; that is, if $S_i \rightarrow S_j$ and $S_j \rightarrow S_i$, it is not necessarily the case that $S_i = S_j$. Therefore, the relation of semantic dependence may not be said to partially order a text, but it may be considered as a PREORDER on a text, and as such may legitimately define chains of related sentences.

Semantic dependence as defined in the preceding paragraphs is clearly a SURFACE relation between sentences in a written or spoken text. It is, in other words, a relation defined from the standpoint of the listener or reader exposed to a sequence of linguistic utterances, and as such takes into account only those properties which may be apprehended through an examination of the superficial representations of these utterances. For this reason, there is no necessity for requiring semantic dependence to hinge on questions of relative ordering; for example, to require that if $S_j \rightarrow S_i$, then S_i precedes S_j in the text. In view of the above discussion, it may be seen that a given sentence S_j may be semantically dependent upon another sentence S_i which occurs further along in the same text and which serves to disambiguate or more fully specify S_i.[8] The use of such semantic configurations is a common literary device used to gradually develop a plot, and is particularly successful in building up a condition of suspense. It is at present an unsolved question whether at the underlying semantic level semantic dependence is a strict ordering relation. If one allows for the consideration of semantic structures of an arbitary degree of abstractness, this question becomes almost a tautology, for as one approaches the deepest levels of cognition the so-called ordering

[8] Van Dijk (1973: 61) has put the matter thus: "The determination of semantic content does not depend merely on the set of transition rules ... but is globally programmed by the abstract underlying macro-structure. At the surface level this implies that $S_{i+1}, S_{i+2}, ..., S_n$ may retro-determine the semantic structure of the propositions $S_1, S_2, ..., S_i$, which were at first either ambiguous or unspecified. The semantic relations holding in the ordered n-tuple, then, seem therefore essentially symmetric: if sentence S_i is dependent on S_j, then S_j is dependent on S_i. Of course, this is true only with respect to the general property of dependency, not of specific semantic relations between elements of the respective sentences: if R (x, y) then ~R (y, x)." It is imperative to note that, while talking about essentially the same sort of relation, Van Dijk is not using the term 'dependence' in the same sense as is to be found in the present paper. As used by Van Dijk, dependence merely signals the semantic interrelatedness of two sentences, with further more specific properties being given by additional formal relations. This is more clearly indicated by the following statement, from Van Dijk (1971: 29): "Many modern texts ... do not respect the rule of linear coherence. In such cases there is no identifiable semantic relation between two or more subsequent sentences. This violation of one of the basic conditions of text coherence is often compensated on the level of the whole text: the semantic deep structure of mutually incoherent sentences can reveal a coherent textual deep structure ... ." In the present study, the term 'semantic dependence' is taken to denote an inherently asymmetrical or at least non-symmetrical relation, exhibiting clear properties of directionality with respect to the semantic structures of two related sentences.

relation for semantic dependence appears to constitute a necessary condition for rational thought and expression. On the other hand, if one deals only with semantic deep structures of the types which have appeared in recent literature, there seems to be no need to posit such an ordering relation, as long as certain global or textual criteria are satisfied. The ultimate resolution of this question, however, hinges upon the construction of an adequate model of textual macro-structures, a task which has not yet been accomplished in sufficient detail to permit a solution.[9]

III. A TOPOLOGICAL MODEL

3.1 In a certain sense, the relation of semantic dependence defined above is an idealized model, restricted to idealized texts. By this it is meant that, in addition to the cases defined above, semantic dependence requires, in an indirect fashion to be sure, certain global conditions of textual unity. For example, given any sentence S_i in a discourse D, one presupposition on the part of the speaker is roughly 'I assume that you heard and understood all the preceding utterances $S_1, ..., S_{i-1}$ in this discourse D'. Presuppositions of this nature are difficult if not impossible to incorporate into a formalized theory, and yet they form the most fundamental basis for any sort of meaningful discourse. If semantic dependence were to be defined in terms of such presuppositions, every sentence in a text would be considered dependent on all preceding sentences, and any finer classification would be impossible. While such relations do in fact account for much of the inherent unity of coherent texts, the present study is aimed at more specific relations of semantic dependence, which may be studied pairwise as well as across the entire expanse of a text. The notion of dependence developed above is based on explicitly defined parameters which may be used to investigate in detail the semantic structure of complete texts. While not foolproof, the characterization offered above seems to concur with most intuitively derived conceptions of semantic dependence, and hence it may be provisionally adopted as a working definition. With this definition in hand, it will now be possible to formulate a mathematical model which will formally illustrate the intuitive notions of textual coherency. What is needed, in effect, is a model which will explicate, by means of a non-arbitrary formal characterization, the properties of semantic dependence to be found in coherent and non-

[9] For significant work along these lines, see the works of Van Dijk referred to above.

coherent texts. Such a model should yield formal results which are compatible with more traditional analyses, so that the model may subsequently be extended to texts of a length and complexity not generally encountered in previous textual analyses. Most importantly, this model should yield further insight into the mathematical structure of natural languages by showing yet another way in which a formal model may account for linguistic data. In the present case, it appears fruitful to approach the study of semantic dependence from the standpoint of topology, the mathematical study of fundamental spatial properties. Since the distribution of semantic material throughout a text is essentially a form of (abstract) spatial patterning, this application of a topological model to study textual dependence can potentially yield meaningful results.

3.2 Topology, representing an outgrowth of axiomatic set theory, provides a means of formally decomposing a given domain or SPACE into a network of constituent parts, utilizing the Boolean operations of union and intersection. Informally speaking, a topological structure or TOPOLOGY defined on a given domain is a collection of subsets closed under the operations of union and intersection. More rigorously, one may utilize the following dfienition:

DEFINITION 1: Given a domain D and a collection $\mathscr{I}$ of subsets of D, $\mathscr{I}$ is a TOPOLOGY over D if the following conditions are satisfied:

(1) $D \in \mathscr{I}$ and φ [the null set] $\in \mathscr{I}$.

(2) Given any collection of members of $\mathscr{I}$ $\{x\alpha : \alpha \in \mathscr{A}\}$ where $\mathscr{A}$ is any indexing set, the union: $\bigcup_{\alpha \in \mathscr{A}} x\alpha \in \mathscr{I}$.

(3) Given any collection of members of $\mathscr{I}$ $\{x_1 \ldots x_n\}$, where n is a FINITE number, the intersection: $\bigcap_{i=1}^{n} x_i \in \mathscr{I}$.

Criterion (3), calling for finite intersections, is crucial for the purely mathematical definition of a topology, in order to avoid certain contradictory characterizations, particularly in the case of the Euclidean spaces.[10] However, for the purposes of textual analysis it is assumed that

[10] As a counterexample demonstrating the need for restricting the definition to finite intersections, let D be the set of real numbers, and let $\mathscr{I}$ be the set of al l open intervals (a, b), where a and b are real numbers and $a < b$; i.e. the standard Euclidean

only a finite number of elements is involved; consequently, (2) and (3) may be effectively paraphrased by requiring that all unions and all intersections of members of $\mathscr{I}$ also be members of $\mathscr{I}$. Since the only formal prerequisite for applying topological analysis to a given domain is the ability to define sets, there is in principle no obstacle in the way of establishing a topological model of textual structure. There have, in fact, been a number of partial but interesting attempts at defining topological structures for natural languages, generally based on the point-set topological notions of neighborhood and distance.[11] In order to study the relation of semantic dependence, however, the machinery of metric topology is of no avail, for the object of study is not individual neighborhoods, but arbitrarily long strings of semantically related utterances. We must therefore direct our efforts to the construction of a topology which will make fundamental use of the basic defining properties of semantic dependence. Clearly, in any given domain D, there is no single unique topology which is assigned to this space; in fact, given any collection K of subsets of D such that all elements of D are contained in the members of K (i.e. K is a COVERING of D), one may construct a topology $\mathscr{I}$ consisting of all unions of members of K, and all finite intersections of these unions. It is, therefore, the choice of this set K, called the SUBBASIS for the topology $\mathscr{I}$, which determines the structure of the resulting topology. The task which remains for the present study is then the construction of an appropriate subbasis for a topology of semantic dependence, which will formally reflect the structure of this relation. No algorithm exists for the construction of such a subbasis; it is up to the individual investigator to incorporate those features necessary to his analysis. In the present case, the topology to be constructed follows from the general remarks on semantic dependence given above, as well as from more nearly universal mathematical considerations. No claims are made as to the superiority of this model over alternative models which might be proposed, since at this juncture any such discussion would be hopelessly hypothetical. Instead, the feasibility of the model will be

topology. Define an infinite series of sets x_i as follows: $x_i = (1 - \frac{1}{i}, 2 + \frac{1}{i})$. The infinite intersection $\bigcap_{i=1}^{\infty} x_i$ is the closed interval [1, 2], which is not a member of the Euclidean topology $\mathscr{I}$. This example of course holds for any open interval (a, b), and may be analogously extended to Euclidean spaces of any dimension.

[11] See, for example, Edmundson (1967), Marcus (1971), Kálmár (1967), Lipski (1974a, b) for some proposals concerning the application of topology to natural languages.

demonstrated by showing the compatibility of the results yielded by this model with the notions of textual coherence referred to earlier.

3.3 Crucial to the construction of a topology of semantic dependence is the ability of this relation to form linear strings of related sentences, by virtue of the fact that this relation is a preorder on the set of utterances of any given discourse. Let us therefore take as the basic defining sets for the proposed model all MAXIMAL strings of semantically dependent utterances. Given a text D consisting of the sentences $S_1, \dots, S_n$, a maximal string will be defined as any string of sentences $S_i, \dots, S_j$ such that $S_j \to, \dots \to S_i$, and such that there exists no sentence S_k such that $S_i \to S_k$. Let K be the collection of all maximal strings in a text D. Such a collection is clearly a covering of D, and hence forms a subbasis for the topology $\mathscr{I}$ consisting of all unions of members of K and all intersections of these unions. It is trivial to verify that this topology satisfies the above definition, and hence one may proceed directly to the investigation of the properties exhibited by this particular topology. It is obvious, even on first inspection, that the sets comprising the topology $\mathscr{I}$ are directly dependent on the properties of semantic dependence shown by the text under consideration, via the action of the two Boolean operations. The intersection of two maximal strings is itself either a maximal string or an empty intersection; this is the case even when a given sentence S_i is semantically dependent on each of two sentences, say S_j and S_k, neither of which is in turn dependent upon another sentence in the text. In such a case we take the combined pair $\left\{ \begin{matrix} S_j \\ S_k \end{matrix} \right\}$ as the starting point for those maximal strings whose second member is S_i. In case two maximal strings $S_i \dots S_m$ and $S_j \dots S_m$ both converge on the same sentence S_m, then any string including a sentence S_k such that $S_k \to S_m$ will have to simultaneously include both the string beginning with S_i and the string beginning with S_j; this may be denoted as:

$$\begin{matrix} (S_i \dots) \searrow \\ (S_j \dots) \nearrow \end{matrix} S_m, S_k$$

The union of two maximal strings is itself either a maximal string or two maximal strings with a common intersection; if the intersection of these two strings is empty, then the union will simply be two disjoint maximal strings. Consequently, the set of maximal strings in a text D is already closed under the operation of intersection, and we need specify $\mathscr{I}$ to only consist of all maximal strings and all unions of maximal strings,

with the null set φ added for completeness. Thus, the set of maximal strings forms not only a subbasis but also a BASIS for the topology $\mathscr{I}$, thereby demonstrating more clearly the manner in which the topology is created from the class of defining sets.

It is also necessary at this point to discuss the manner in which closed loops of semantically dependent utterances are to be represented within the proposed model. For the first case, let us consider a simple closed loop containing no branching structures; i.e. a sequence of sentences $S_i \dots S_j$ such that $S_j \to \dots S_i \to S_j$. Clearly, it is impossible to speak of this loop as having a beginning or an end; we may, however, consider the entire closed loop to represent an instance of a maximal string, represented as $\{S_j \dots S_i, S_j\}$. This convention is consistent with the preceding definitions, since the intersection of such a maximal loop with any other maximal string will of necessity be void, unless such a maximal string is itself contained in the loop, in which case the intersection will simply be the maximal string in question. The union of a closed loop and another maximal string will also satisfy the definition, since it will either be the closed loop, or the disjoint union of the closed loop and another maximal string.

As the next case, one may consider a maximal string ending in a closed loop; i.e. a string of the form $S_k \to \overset{\nearrow S_k}{S_j} \to \dots \to S_i$. In such a situation, one may represent the maximal string by simply 'reading around' the loop at the end: $\{S_j, S_k, S_j, \dots, S_i\}$; the procedure is identical for terminal loops with more than two members. In a similar fashion, this mode of representation may be extended to strings which begin with a closed loop, strings of the form $S_j \to \dots \to S_{i+1} \to S_i \to S_{i+1}$. A string of this nature may be represented as the set $\{S_j, \dots, S_{i+1}, S_i, S_{i+1}\}$. Closely tied to the preceding two cases are strings containing a closed loop in an intermediate position, exhibiting a configuration of the form $S_k \to \dots \to S_{j+1} \to S_j \xrightarrow{\nearrow S_{j+1}} S_{j-1} \to \dots \to S_i$. In this case, one again reads around the internal loop before proceeding to the remaining sentences: $\{S_k, \dots S_{j+1}, S_j, S_{j-1}, \dots, S_i\}$. By representing closed loops in this fashion, it is possible to ensure closure of the operation of intersection defined on the class of maximal strings, thus meeting the requirements for a topological model and remaining within the confines of the proposed model. These conventions of definition may of course be extended to configurations of any degree of complexity, involving combinations of loops and open chains. The simple algorithm of representation given above will yield the correct results in all cases, and consequently any expanse of text may be reduced to a collection of well-defined sets by applying the topological model which has been described.

IV. APPLYING THE MODEL

4.1 It is, as noted earlier, impractical to discuss at this point specific examples drawn from actually-occurring texts. Limitations of space preclude the necessary computations and representations, and in any case the theoretical cause would be little advanced by such a procedure. It is interesting, however, to consider in schematic form the manner in which the formal model just presented may be used to provide a rigorous definition of textual coherence. Intuitively speaking, a text is coherent if it contains no breaks, 'holes', or is not composed of mutually disjoint sections for which no common connecting theme may be found. In a similar vein, it is possible to consider the form in which a given topological space is represented as the union of its constituent sets; i.e. whether the pieces overlap or fit smoothly, or whether any 'holes' are introduced into the union. Formally, this topic is covered by the property of CONNECTEDNESS. A space is connected if it cannot be non-arbitrarily decomposed into mutually disjoint subsets; formally:

DEFINITION 2: Given a space D with a topology $\mathscr{I}$, D is CONNECTED if there do not exist two NONEMPTY sets x_i and x_j, both members of $\mathscr{I}$, such that $x_i \cap x_j = \varphi$ and $x_i \cup x_j = D$.

It is claimed that the property of connectedness as applied to the model under consideration is mathematically equivalent to the notion of textual coherence, both globally and locally. In order to visualize the relation between the global and the local properties of connectedness, it is necessary to consider the following definition, giving the method of restricting a topology to a particular subset:

DEFINITION 3: Given a space D with topology $\mathscr{I}$, and given a subset $A \cup D$, the topology $\mathscr{I}$ induces a RELATIVE TOPOLOGY $\mathscr{I}_A$ on A consisting of all intersections between A and members of $\mathscr{I}$.

It is trivial to verify that $\mathscr{I}_A$ defined in this fashion is in itself a topology on the set A. From this definition, it is possible to define connectedness

with respect to a particular subset A, by requiring that A not be representable as the disjoint union of two nonempty members of $\mathscr{I}_A$. It is important to realize that, while a given subset $A \subset D$ may be disconnected with respect to $\mathscr{I}_A$, the entire space D may be connected with respect to $\mathscr{I}$; similarly, a disconnected space may contain many subsets which are connected with respect to their relative topologies. The relative independence of the global and the local properties of connectedness is of the utmost importance to the topological model of semantic dependence, for it closely parallels the characteristics of linguistic texts with respect to the property of coherence.

4.2 The simplest example of semantic dependence is offered by an idealized text $D = S_1, \ldots, S_n$ in which every utterance is dependent upon the preceding ones. In such a text, the sets forming the basis for the topology proceed in an overlapping fashion: $\{S_1\}$, $\{S_1, S_2\}$, ... $\{S_1, \ldots, S_n\}$. Any union of members of the basis will itself be a member of the basis, so in this case, the basis is also the topology. A text of this nature cannot be represented as the disjoint union of nonempty members of the topology $\mathscr{I}$, since the sets comprising the topology form a nested collection; hence, any such text is by definition connected, and of course also coherent. It also follows from the formal specifications of the topological model that, in more general terms, ANY coherent text is connected. This assertion can be easily proven by contradiction: assume that a given coherent text D is disconnected, and may be decomposed into the two components x_i and x_j. Now x_i and x_j, by virtue of being members of the topology $\mathscr{I}$, each consists of either a maximal string or a union of maximal strings of semantically dependent utterances of D. Since $x_i \cap x_j = \varphi$, no sentence occurs in both x_i and x_j; hence there are no semantic dependence relations between x_i and x_j. However, this mutual independence of textual macro-structures defines an INCOHERENT text, and thus it is impossible for a coherent text to be disconnected.

The converse of the above assertion, namely that every connected text is coherent, is even easier to demonstrate. Given a connected text $D = S_1, \ldots, S_n$, the only decomposition into disjoint members of $\mathscr{I}$ is the trivial union $D \cup \varphi$. This means that the text D itself is either a single maximal string $S_1, \ldots, S_n$ of semantically dependent utterances, or the union of a set of maximal strings $K_1, \ldots, K_m$, with a nonempty common intersection; i.e. such that $K_1 \cap K_2 \cap \ldots \cap K_m \neq \varphi$. If D consists of a single maximal string, then it is coherent by definition. If, on the other hand, D is the union of a non-disjoint collection of maximal strings, then in particular, for each $i \leq n$ it is possible to choose from D a collection of i semantically related sentences. Thus, given any two utterances S_i and S_j in D, either $S_i \rightarrow S_j$ or $S_j \rightarrow S_i$; i.e., the text is coherent.

Although the inverses of the above statements, namely that textual disconnectedness is equivalent to textual incoherency, follow logically from the above statements by contraposition, it is useful to explicitly schematize the relationships in question. Let us therefore consider a greatly simplified but nonetheless typical schematic representation of an incoherent text, a text $S_1, \ldots S_n$ in which, as in the above example, each sentence is semantically dependent on the preceding sentence, except for a sequence of sentences $S_j, \ldots, S_k$ which is independent of the remainder of the text, and which forms a semantically-related string in its own right: $S_k \rightarrow \ldots \rightarrow S_j$. The main text therefore is representable as $S_n \rightarrow \ldots \rightarrow S_{k+1} \rightarrow S_{j-1} \rightarrow \ldots \rightarrow S_1$. In this case, the sequence $A = \{S_j, \ldots, S_k\}$ is a member of the topology $\mathscr{I}$, since it is not itself dependent on any other sentences. Also a member of the topology $\mathscr{I}$ is the longest maximal string which may be defined on the remainder of the text, namely $B = \{S_1, \ldots, S_{j-1}, S_{j+1}, \ldots, S_n\}$. Now $A \cup B = D$, and $A \cap B = \varphi$; hence, the space is not connected. This example may be extended to any case in which there exists a string or loop $S_i, \ldots, S_j$ of semantically related utterances which is independent semantically from the remainder of the text. In each such case, the 'island' $\{S_i, \ldots, S_j\}$ will be a member of the topology $\mathscr{I}$, as will be the longest maximal string contained in the remaining text, thus providing a partition of the text into mutually disjoint subsets.

4.3 The possibility for extension to cases of arbitrary complexity demonstrates that any incoherent text is disconnected with respect to the topology $\mathscr{I}$, for any incoherent text by definition contains one or more sequences of utterances which bear no demonstrable semantic relationship to the rest of the text. The converse of this statement, that every disconnected text is incoherent, follows directly from the definition of connectedness: if a given text may be exhaustively decomposed into two or more disjoint maximal strings of semantically dependent utterances, then each such sequence is semantically independent from every other such sequence; i.e., the text is incoherent.

The equivalence of connectedness and coherence on a local level follows from the definition of relative topology, and it may thus be seen how locally incoherent passages may be embedded in globally coherent expanses of text.

V. CONCLUSION

The demonstrated equivalence between a mathematical model and a precise although non-mathematical characteristic of linguistic texts promises to be of significance to the study of text-grammars, and to mathematical linguistics in general. By regarding a text as a topological space, it is possible to pass naturally from highly localized descriptions of textual structure to global characteristics spanning the entire text. Any investigation of text-grammars which aims for a complete characterization of textual macro-structures must bring to bear analytic machinery which is capable of providing information about entire texts, or about extensive fragments of texts. The domain of pragmatics promises to provide one such analytical tool, and it is also likely that certain logico-semantic structures may be discovered which incorporate material from entire texts, for example in the form of global 'transderivational constraints' of the form currently being investigated within the field of generative grammar.[12] The topological model presented in this paper also represents an analytic procedure for obtaining specific data from large expanses of text, and as such may potentially be incorporated into a model of text-grammar. A more precise characterization of the mathematical properties of natural languages will facilitate the construction of more adequate text-grammatical models, and may also contribute to the field of machine translation. In particular, given the formal characteristics of linguistic texts, it has been apparent for some time that an investigation into the topological properties of such texts would probably be fruitful. The model presented in the preceding paragraphs exemplifies only one of many topological possibilities exhibited by natural languages, and is not to be considered in any sense exhaustive. The sole aim of this study has been to demonstrate the feasibility of searching for formal mathematical characteristics of textual macro-structures. It has been seen that by suitably constructing a topology based on a rigorous notion of semantic dependence, it is possible to formally characterize textual coherence in terms of topological theory. Needless to say, the model and examples presented above were of an extremely rudimentary nature, and do not accurately reflect the complexity to be found in typical specimens of linguistic discourse. It is hoped, however, that this practical idealization does not undermine the theoretical issues under consideration, for the model is complete in every detail, being schematically presented only

12 Some further considerations along these lines may be found in Van Dijk (1972b) and elsewhere in works by the same author.

for ease of exposition. The reported results are of course still tentative, for an exhaustive characterization of semantic dependence remains to be offered. Nonetheless, it appears at this point that the topological model under discussion is essentially correct in all major aspects. By carefully working out the necessary details, and by combining the results obtained from this model with other areas of textual investigation, it may be possible to achieve a degree of convergence which will place the topological model on a firm basis within textual theory.

REFERENCES

Bellert, Irena
1970 "On a Condition of the Coherence of Texts", *Semiotica* 2, 335-63.

Edmundson, H. P.
1967 "Mathematical Models in Linguistics and Language Processing", *Automated Language Processing*, ed. by H. Borko (New York: Wiley), 33-96.

Kálmár, László
1967 "Le langage comme structure algébrique", *Cahiers de Linguistique Théoretique et Appliquée* 4, 73-82.

Lipski, John M.
1974a "Towards a Topology of Natural Languages", *Poetics*, in press.
1974b "Connectedness in Poetry: Towards a Topological Analysis of E. E. Cummings", *Language and Style*, in press.

Marcus, Solomon
1971 *Poetica matematică* (Bucharest: Editura Academici Republicii Socialiste România).

Schoeck, Richard J.
1968 "Mathematics and the Languages of Literary Criticism", *Journal of Aesthetics and Art Criticism* 26, 367-76.

Strawson, P. F.
1952 *Introduction to Logical Theory* (London: Methuen).

Van Dijk, Teun A.
1971 "Some Problems of Generative Poetics", *Poetics* 2, 5-35.
1972a *Some Aspects of Text Grammars* (The Hague: Mouton).
1972b "A Note on Linguistic Macro-structures", *Linguistische Perspektiven, Referate des VII Linguistischen Kolloquiums* (Tübingen: Niemeyer), 75-87.
1973 "Models for Text Grammars", *Linguistics* 105, 35-68.

John M. Lipski (b. 1950) is an instructor of foreign languages at Newark State College, Union, New Jersey. He describes his research interests as the mathematical theory of texts, the behavioral correlates of language, and foundations of phonological theory. Recent and forthcoming publications: "The Survival of a 'Marked' Segment in Portuguese" (1973); "Features, Markedness, and Shorthand (to appear); "Towards a Production Model of Spanish Morphology" (to appear); and "Orthographic Variation and Linguistic Nationalism" (to appear).