

THE CONSTRUCTION AND USE OF A PROPOSITIONAL TEXT BASE\*

Althea Turner & Edith Greene  
Department of Psychology  
University of Colorado

Technical Report No. 63  
Institute for the Study of Intellectual Behavior  
University of Colorado  
Boulder, Colorado 80302  
April, 1977

\*This paper was supported by Grant MH15872 from the National Institute of Mental Health to Walter Kintsch and by Grant BNF-7202084 from the National Science Foundation to the Institute for the Study of Intellectual Behavior.

## TABLE OF CONTENTS

1	Introduction	1
	1.1 Theory	2
	1.1.1 Word Concepts	3
	1.1.2 Propositions	3
	1.1.3 Text Bases	4
	1.1.4 Macrostructure	5
	1.2 Conclusion	6
2	Relations and the Construction of Propositions	7
	2.1 Predication	9
	2.1.1 Propositions Which Use Case Relations	9
	2.1.2 Nominal Propositions	12
	2.1.3 Reference	13
	2.1.4 The Use of "Have" and "Is"	15
	2.2 Modification	16
	2.2.1 Qualifier Propositions	17
	Adjectival Propositions	18
	Adverbial Propositions	19
	Hedges	20
	2.2.2 Quantifiers	21
	2.2.3 Partitive Propositions	22
	2.2.4 Negative Propositions	23
	2.3 Connective Propositions	24
	2.3.1 Construction of Connective Propositions	27
	Conjunction	27
	Disjunction	29
	Causality	30
	Purpose	32
	Concession	33
	Contrast	34
	Condition	35
	Circumstance	36
	Time	36
	Location	38
	Manner	39
	2.3.2 Multiple Meanings: Ambiguous Connectives	40
3	Step-by-Step Construction of a Text Base	44
	3.1 The Ordering of Propositions in a Text Base	45
	3.2 A Night on the Town	47
	3.2.1 Text	47
	3.2.2 Procedure	47
	3.2.3 The Template Text Base	53
	3.2.4 The Base Structure	54
	3.3 Merino Sheep	55
	3.3.1 Text	55
	3.3.2 Procedure	55
	3.3.3 The Template Text Base	58
	3.4 <u>The Sun Also Rises</u> (excerpt)	59
	3.4.1 Text	60
	3.4.2 Procedure	60
	3.4.3 The Template Text Base	68
	3.4.4 A Note on Literary Passages	69

4	Scoring Recall Protocols	70
	4.1 Purposes	71
	4.2 Scoring Criterion--A Priori Considerations	72
	4.3 Examples of Scoring Techniques	73
	4.3.1 Scoring for Amount of Reproductive Recall	73
	Text--Merino Sheep	75
	Protocols--Merino Sheep	75
	Scoring Protocols	75
	4.3.2 Scoring for Form of Recall	76
	Analyzed Protocol with Propositions	78
	4.3.3 Scoring for Recognition	79
5	Conclusion	80
	References	81

---

Acknowledgments

We wish to give special thanks to Walter Kintsch, whose help and guidance have been invaluable to the development of this paper. In addition, we would like to thank the people who have commented on various drafts of this manual, in particular, Hans Brunner, Alinda Friedman, Eileen Kintsch, Ely Kozminsky, and Doug Vipond.

## 1 INTRODUCTION

To study complex processes using spoken and written texts, one requires some means to determine the conceptual meanings underlying the text. These meanings should be divorced from the surface words and sentences used to express this meaning in the text. Some nonarbitrary method is needed to describe the stimulus input (the text) in terms of its semantic content.

A basic prerequisite for experimental work using prose materials is that a method exist to represent the meaning of texts. Such a method ought to be useful for research purposes as well as sophisticated enough to be able to encode any conceptualization.

A number of theories have been proposed in recent years which represent knowledge in propositional form. Many researchers have been confused by the variety of representations. In addition, the proposals are often somewhat sketchy and do not serve as practical guides for the application of a particular representation.

One of the propositional theories has been suggested by Kintsch. This representation has proven to be a valuable tool in the investigation of memory for prose. The present work is an attempt to formalize this method of propositional analysis and to provide a guide for its use. It is hoped that this explication will aid in stimulating further research on text comprehension and memory. Little explanation of the theory is presented, since it has been presented in detail elsewhere (Kintsch 1970, 1972, 1974).

The purpose of this work is to serve as a practical manual for the application of propositional analysis. Its emphasis, therefore, is on

the process of deriving a propositional text base for a text, and the subsequent utilization of that text base. Thus, this is not a theoretical handbook, but a manual for the use of a particular propositional system.

The system presented below is intended to be restricted in use to the semantic content of texts. Pragmatic considerations have not been addressed. Conversational speech acts are usually not appropriate material for a propositional analysis, as the pragmatic aspects of conversation would not be represented. Likewise, information which is solely syntactic is not represented in a propositional text base.

A distinction must be made between the method of propositional analysis and Artificial Intelligence parsing devices. The major difference is that propositional analysis is experimentally oriented. It serves as a tool for research into cognitive processes which use texts as stimuli or responses. There is no intention to formulate or build a computer parsing device. No interface system between input text sentences and the propositional output has been developed; no syntactic analyzer has been formalized. Indeed no computer is required to analyze a text into a propositional text base, giving one greater flexibility and the ability to deal with more complexity than would be possible in a computer parser.

### 1.1 THE THEORY

The method of propositional analysis to be presented here is based on Kintsch's propositional theory in The Representation of Meaning in Memory (1974). While some details of propositional construction may differ from details which were given previously, the theory remains as expressed in recent work by Kintsch.

In order to deal adequately with the study of texts in an

experimental situation, it is necessary to be able to formally represent their meaning. The meaning of a text is assumed to be represented by a list of connected propositions. This proposition list is called a text base. The propositions themselves are n-tuples of word concepts, one serving as a relation (referred to as predicator previously), the others as arguments of the proposition. Propositions are idea units, each one representing a single idea.

#### 1.1.1 WORD CONCEPTS

Propositions consist of abstract word concepts. These are the entries in the lexicon of semantic memory, and may be represented by one or more words in the surface structure. To distinguish between words and word concepts, the word concept is written in capital letters in the proposition. The actual word or words which are chosen to represent the concept are tokens for the abstract concept. It is important to understand that the abstract concept may be represented by a number of words in a given language. Since the word concepts do not always map directly on to a one-word token, at times the token of the concept in the proposition will consist of several words.

#### 1.1.2 PROPOSITIONS

The propositions themselves consist of two or more word concepts, such that the first is a relation, the others arguments. The relation is a connection between a set of arguments forming a single idea. Propositions have been classified into three distinct classes, based on the type of relation they contain: Predication, Modification, and Connection. Each class of propositions has certain constraints as to what kind of arguments may be taken by the relation in forming a proposition.

— — — —

Predicate propositions express ideas of actions or states. The relations are usually verbs. Included as predicate propositions are nominal propositions and references.

Modifier propositions change a concept by restricting it or limiting it by means of another concept. Negation is included as a modification of a concept also.

Connective propositions relate propositions or facts in the text to each other. These propositions may be left unexpressed in the text, and must then be inferred. They are important for providing a text with cohesion.

Whole propositions are often embedded as arguments in other propositions. In particular, the arguments of connective propositions tend to be other propositions.

### 1.1.3 TEXT BASES

The propositions in a text are connected into a cohesive, inter-related whole called a text base. The text base represents the meaning of the text. Three different kinds of text bases are distinguished here. They are base structure, template text base, and protocol text base.

The base structure is the set of propositions in the author's mind which originally gave rise to the text. Some percentage of these propositions are explicitly stated in the surface structure of the text. Others can be inferred from the text. Many receive no representation in the text whatsoever. To a large extent, the base structure of a text is unknowable.

The second form of text base is the template text base. This is an idealized representation of the meaning of a given text, including all the propositions which are directly expressed in a text as well as all



the inferences which are necessary in order to form a connected and coherent text base. The template text base can serve as a model, which can be compared with the reports from those who read or listened to the text.

The last form of text base is the protocol text base. This is the list of propositions which are represented in a subject's recall protocol. This is a more conservative estimate of meaning than is a template text base. Generally, a text is produced from a highly interconnected base structure. A subject's protocol, however, is not likely to be as well integrated as the text. Thus, propositions which are not given explicitly in the protocol are not included in the protocol text base.

#### 1.1.4 MACROSTRUCTURE

The present work attempts to present the essentials involved in a propositional analysis of a text and the experimental use of the resultant text base. The text base that is constructed consists of "microstructure" propositions, the detailed propositions of a text. Yet a text cannot be adequately characterized by this sequence of microstructure propositions alone. A text also must have a more global aspect called macrostructure. Macrostructure corresponds to a formal concept of the "gist" of a text. Since this overall structure contains a higher level of meaning in the text, the macrostructure is also a sequence of propositions. A certain set of operations is assumed to govern the formation of the macrostructure from microstructure propositions. These macropropositions could be propositions which are expressed directly as part of the microstructure of the text, or they could be constructed from the microstructure propositions through the application of macro-operations which control the formation of the macrostructure. A text may have one or more levels of macrostructure,

each more global than the previous one.

No explication of the principles or derivation of macrostructure is offered here. At the present time, work is progressing on the theory of macrostructure and its relationship to microstructure. Those interested in macrostructure are referred to Kintsch and van Dijk(1975).

## 1.2 CONCLUSION

No justification is given for this system other than the fact that it appears to work well as a tool in experimental work and has some psychological validity. No comparisons will be made between Kintsch's method of propositional representation of meaning and other systems. Our purpose is to make the system more available to others by explicitly stating guidelines for the construction of propositions, text bases, and the utilization of the text bases in experimental settings.

This work is organized into three major sections. Relations and the Construction of Propositions deals with the construction of individual propositions. This is subdivided into three parts, according the three types of relations. Each of the classes of relations is discussed in depth in this section.

The second division, The Step-by-Step Construction of a Text Base, details several examples of text base construction from a given text. This section concentrates on the development of the template text base.

Section three, Scoring Recall Protocols, examines several potential approaches to the utilization of template text bases in recall or recognition experiments. No attempt is made to deal with this subject exhaustively. Instead, a few approaches are outlined and short examples of scoring techniques are given.

## 2 RELATIONS AND THE CONSTRUCTION OF PROPOSITIONS

Propositions are representations of conceptual units. Each one consists of a relation and a number of arguments. The relation and its arguments are abstract word concepts, not to be confused with the actual words used in a given text. The concepts may be represented by one or more words.

Several conventions have been adopted for writing propositions. The relation and arguments are written in capitalized letters to indicate that they are word concepts, not individual words. The relation is written first in the proposition, followed by the arguments in the order specified by the constraints of the given relation. These are separated from each other by commas. The entire proposition is enclosed in parentheses.

The relation is the pivotal concept in the proposition. It connects its arguments in such a way that a single idea results. The idea may be compound or complex. That is, other ideas may be embedded in the proposition as arguments. Relations fall into three major classes: Predication, Modification, and Connection. It is possible to classify propositions on the basis of the type of relation they are expressing.

The arguments of a proposition are constrained by the relation. Certain semantic requirements must be met before a word is acceptable for a particular argument of a relation. Some of these constraints are dictatts are an individual's world knowledge. Other constraints are caused by the nature of the relationship. Many of these constraints are listed below in the appropriate sections.

Predicate relations express an action or a state. They are frequently verbs. When a relation is a verb, its arguments fill particular "cases." These cases are the slots in a verb frame, each of which holds a particular semantic relationship to the action or state specified by

the verb relation. For example, the verb "to hit" would have cases for the person or thing who performs the action of hitting (the agent), the person or thing which receives the action of hitting (the object), and the means by which the action was accomplished (the instrument). Only verbs take cases. Other predicate relations are the copula "ISA," which relates its arguments in terms of set membership, and "REFERENCE," which identifies its arguments as having the identical referent.

Relations of modification express various forms of restrictions or limitations of one concept by another. Four different types of modifiers are discussed: Qualifiers, Quantifiers, Partitives, and Negations. The suggested relations are standardized to indicate explicitly the specific type of modification that is involved.

The third category is Connection. Eight categories of connectives are specified; Conjunction, Disjunction, Causality, Purpose, Concession, Contrast, Condition, and Circumstance. Connectives relate facts or propositions in the text to each other. The purpose of connective propositions is to co-ordinate the text and provide it with coherence. Often connective propositions are left unstated, and have to be inferred from the rest of the text.

## 2.1 PREDICATION

In logical terminology, predication refers to a statement or assertion about the subject of a proposition. A predicate proposition is one that expresses a relationship of action (sentences 1-3 below) or being (sentences 4-6). Sentence 5 is a nominal proposition, example 6 is a reference. These will be detailed below.

1. Betty bought a balloon.  
(BUY, A:BETTY, O:BALLOON)
2. Fred went to Boulder.  
(GO, A:FRED, G:BOULDER)
3. Jane gave a present to Joey.  
(GIVE, A:JANE, O:PRESENT, G:JOEY)
4. Joe feels sick.  
(FEEL, E:JOE, I:SICK)
5. A collie is a dog.  
(ISA, COLLIE, DOG)
6. President Kennedy is JFK.  
(REFERENCE, JFK, PRESIDENT KENNEDY)

### 2.1.1 PROPOSITIONS WHICH USE CASE RELATIONS

The basic structure of a sentence consists of a predicate and one or more noun phrases. Each of these noun phrases assumes a different semantic relation (case) to the predicate. Each argument of a proposition stands in a certain case relationship to its predicator. Thus the general form of such predicate propositions is

$$(\text{RELATION}, C_1, C_2, \dots, C_n),$$

where  $C_1, C_2, \dots, C_n$  are arguments which hold a particular semantic case relationship with the relation. Obviously, the same argument may play different semantic roles in different propositions.

In general, the case system outlined below follows Fillmore (1968,

1969, 1971). The case grammar of Fillmore is more highly developed than others which have been proposed. Therefore it has been used in Kintsch (1972a, 1974) to classify the arguments of propositions according to their relationship to the predicate relation, when that relation is a verb. Other case systems exist which include some cases which will not be discussed here. This is not to be taken as a denial of the possible merits of these other systems or of the existence of other cases.

Fillmore's theory of case grammar identifies the following categories of cases:

1. Agent (A) - typically animate instigator of the state or action identified by the verb
2. Experiencer (E) - experiencer of psychological event
3. Instrument (I) - typically inanimate stimulus of an experience, a force or object causally involved in the state or action identified by the verb
4. Object (O) - object of an action which undergoes change or movement
5. Source (S) - source or state or action identified by the verb
6. Goal (G) - result or goal of state or action identified by the verb

Case relations are represented by category symbols (e.g. A: or O:) which will be used throughout the remainder of this work. Each symbol (and hence each case relationship) can occur only once in a simple sentence, although compound instances of a single case may be permitted through conjunction. In the sentence:

7. Beth hit the ball with a bat.  
(HIT, A:BETH, I:BAT, O:BALL)

each argument assumes a different semantic role, as designated by its case marker. Beth is the agent, bat is the instrument, and ball is the

object. Further examples:

8. The umpire angered Andy.  
(ANGER, A:UMPIRE, E:ANDY)
9. Pat practiced from noon until dusk.  
(PRACTICE, A:PAT, S:NOON, G:DUSK)
10. Ralph hit a home run.  
(HIT, A:RALPH, G:HOME RUN)

The presence of certain case relations may be either obligatory or optional in a given proposition. Consider these examples, as provided by Fillmore (1969):

11. The door opened.  
(OPEN, O:DOOR)
12. John opened the door.  
(OPEN, A:JOHN, O:DOOR)
13. The wind opened the door.  
(OPEN, I:WIND, O:DOOR)
14. John opened the door with a chisel.  
(OPEN, A:JOHN, I:CHISEL, O:DOOR)

The verb "open" requires an object, while an agent or an instrument are optional. This can be written

(OPEN, [A,] O, [I]),

where brackets indicate optional cases. This listing of the specific cases that can follow from a given predicator is called a "verb frame."

If a verb requires the presence of a certain case, but it is absent in the surface structure of the sentence, then an empty case marker (\$) must fill the case position in the proposition. This is often the situation when a noun phrase which would normally fill that case can be inferred from contextual information. For example,

15. The message was received.  
(RECEIVE, A:\$, O:MESSAGE)

The verb "receive" requires the presence of an agent, someone to receive the object. Since the case is empty here, an empty case marker is used to fill the agent slot.

It is possible to characterize verbs of action through the use of verb frames. Since verbs with the same frame can be grouped together, these listings provide a means of classifying predicate propositions. Verbs can be characterized not only by specification of the frames into which they fit, but also by analysis of their transformational properties. The choice of a particular noun phrase to become the surface structure subject, and the choice of a certain preposition to accompany each case element are among important determinants in the classification of predicate propositions. It must be emphasized however, that the cases used in a propositional analysis of discourse are semantic cases, not syntactic cases. Each case assumes a certain semantic relation to the predicate that it follows. Refer again to example 10 (The door opened.). A syntactic analysis of this sentence would suggest that the noun "door" is the subject of the sentence. Semantic case grammar specifies an object as an entity that undergoes change as the result of an action. Therefore the semantic case of "door" in this sentence is the object case.

#### 2.1.2 NOMINAL PROPOSITIONS

Nominal propositions should be distinguished from those propositions that involve case relations. Nominal propositions express set membership. They assert that some noun belongs to a set named by another noun. Such propositions are of the form:

(ISA, A1, A2)

where arguments A1 and A2 are nouns, and where A1 is subordinated to



A2, as an element of the set A2. Note that the presence of the copula "ISA" differs from Kintsch (1974). Its use was adopted in an effort to clarify the precise relationships involved in these propositions. The copula will be applied only in propositions denoting set membership. Example 5 above illustrates the use of the copula "ISA:"

5. A collie is a dog.  
     (ISA, COLLIE, DOG)

A collie is a member of the superset category "dog."

Most sets with which psychologists and linguists must deal are "fuzzy" and defined only by one or more "typical members." The superset class is delineated only vaguely. Nominal propositions, then, must be inferred on the basis of probabilistic and ambiguous information. Semantic memory does not contain an exhaustive list of category memberships. It is likely that some typical examples are stored in semantic memory, while others must be derived from other information in memory.

### 2.1.3 REFERENCE

A referential proposition is one which states that the referent of one argument is the same as that of the second argument. While this form of proposition is not common in short texts, it is necessary in a long one to maintain the coherence of the text. Propositions of reference are frequently implied propositions. The referent word may emphasize one or another aspect of the referent. Pronouns are also an example of this type of proposition. Reference propositions take the form:

(REFERENCE, A1, A2),

where A1 and A2 are different expressions for the same entity. This

may be expressed linguistically or indicated by extralingual information through reference to objects or events physically or psychologically present. While the relation of reference can take arguments which are properties or relations, they most commonly take arguments which are nouns, usually persons, animals, or objects, as opposed to abstractions.

A proposition of reference is often inferred from the text, from knowledge of the language's syntax, or from world knowledge. It is considered good literary practice to use more than one expression to indicate the same referent. Thus it is often necessary for the reader to form referential implications in order to comprehend the text.

#### EXAMPLES:

16. Clark Kent is Superman.  
(REFERENCE, CLARK KENT, SUPERMAN)

Both arguments refer to the same individual. There is only one Clark Kent, as there is only one Superman. Although the identity is the same, the arguments may be used in different contexts. When he is working on his job at the Daily Planet, this particular individual is likely to be referred to as Clark Kent, but when flying over the city of Metropolis, he is Superman.

17. The subjects used in the experiment attended Colorado University...  
The participants were divided into groups...The students reported  
what they saw...  
(REFERENCE, PARTICIPANTS, SUBJECTS)  
(REFERENCE, STUDENTS, (REFERENCE, PARTICIPANTS, SUBJECTS))

The words "subjects" and "participants" are loosely interchangeable in psychological literature. The same set of individuals are referenced through the two expressions.

Not all subjects are students, nor students, subjects. However, in the context provided by this example, a particular set of subjects and a particular set of students are identical. In one case, they are referred to by their role as subjects, while in the other, by their role as students. Again, the referent is identical, regardless of the word used.

## 2.1.4 THE USE OF "HAVE" AND "IS"

The verbs "to have" and "to be" can be used as auxiliary verbs to form a number of tenses from other verbs. Propositions do not have tenses, therefore when used as verbal auxiliaries, "have" and "is" are not represented in the propositional text base. The following example shows a number of tense transformations of the text base (EAT, A:STEVE, O:(QUALITY OF, CAKE, CHOCOLATE)):

- 18a. Steve eats chocolate cake.
- b. Steve is eating chocolate cake.
- c. Steve was eating chocolate cake.
- d. Steve has been eating chocolate cake.
- e. Steve has eaten chocolate cake.
- f. Steve had eaten chocolate cake.
- g. Steve will have eaten chocolate cake.
- h. Steve is to eat chocolate cake.
- i. Steve was to have eaten chocolate cake.
- j. Chocolate cake was eaten by Steve.
- k. Chocolate cake is being eaten by Steve.

Tenses are assumed to be a product of the syntax of the language. They express distinctions in time. This temporal information is derived from contextual information or explicit time references given in the text base. When transforming propositional information into sentences, the tenses are chosen according to linguistic rules and available temporal references.

In order to avoid the possible pitfall of propositionalizing function words, the convention has been adopted here of eliminating the use of "have" and "is." For uses of "have" which involve a semantic component, other relations such as "POSSESS," "QUALITY OF," or "PART OF" (see Modification section below) will be substituted. For instance,

19. Lori has a book.  
(POSSESS, A:LORI, O:BOOK)
20. A cat has whiskers.  
(PART OF, CAT, WHISKERS)

21. Sally has red hair.  
(QUALITY OF, SALLY, (RED, HAIR))
22. Bob has to go home.  
(QUALIFY, (GO, A:BOB, G:HOME), NECESSARY)

When the verb "to be" is used with semantic intent, some of the possible substitute relations are "REFERENCE," "ISA," and "QUALITY OF."

23. The person Lynn met was Ray.  
(REFERENCE, RAY, (MEET, A:LYNN, O:PERSON))
24. A trout is a fish.  
(ISA, TROUT, FISH)
25. Milton is fat.  
(QUALITY OF, MILTON, FAT)

"Have" and "is" can take many other meanings. Care should be taken to determine their meaning in the text before attempting to construct a text base. When necessary, the use of a good dictionary is recommended for this purpose.

## 2.2 MODIFICATION

The four most important types of modifier propositions are Qualifiers, Quantifiers, Partitives, and Negations. Other modifying relations are possible. However, the four outlined here are broad enough to be sufficient for most purposes. Qualifier propositions limit or restrict the scope of an argument or proposition by expressing a quality or attribute of it. Quantifier propositions can express a definite or indefinite quantity, either of the numerosity of distinct entities or of the extent of a single entity. Partitive propositions indicate a part of a collective whole. Propositions of negation express the complement of the asserted proposition. Modifier propositions can be expressed in this way:

(MODIFYING RELATION, A1, A2),

where A1 is modified by the argument A2 in the way indicated by the modifying relation. Qualifier and quantitative propositions can be simplified to

(A2, A1),

if the meaning of the proposition remains clear without the explicit relation.

### 2.2.1 QUALIFIER PROPOSITIONS

There are two forms of qualifier propositions. One form, "QUALIFY," restricts the scope of the modified argument, while the other form, "QUALITY OF," indicates an attribute of the modified argument.

Some qualifier propositions reduce the scope of a general concept or proposition to a more particular or restricted form. They take the form

(QUALIFY, A1, A2),

where the argument A1 is limited or altered by the argument A2.

When the proposition is a qualification of an argument through the naming of an attribute or feature, it takes the following form:

(QUALITY OF, A1, A2),

where the argument A1 has a certain quality or feature, namely A2. Both kinds of qualifier propositions can be written

(A2, A1).

This second notation alone is used in Kintsch (1972a, 1974). The forms using the relation "QUALIFY" or "QUALITY OF" make explicit the underlying relationship of qualification. Qualifiers can use adjectives

or adverbs as the modifying argument A2, depending on the nature of the argument which is being modified, A1. If A1 is a noun phrase, A2 will be an adjective; if A1 is a verb, adjective, or a proposition, A2 will be an adverb.

#### Adjectival Propositions

An adjective defines some attribute of the noun phrase that it modifies. Written in less formal manner, these propositions are similar to example 25:

25. Milton is fat.  
                   (QUALITY OF, MILTON, FAT)

or more simply

                  (FAT, MILTON)

This is an implicit statement of comparison. Milton's fatness has no relevance unless one compares his size to that of others of his class. He is only fat when compared to a population that is thinner than he is. Explicit comparisons, such as

26. Milton is fatter than Dan.

are discussed below in the section on Contrast in Connective Propositions.

The concepts named by adjectives are often fuzzy. Most qualities can take on any number of values, which are usually based on an underlying, continuous distribution. In many cases, these attributed values are assigned according to some property present in the object or individual being modified. For example, the attribute of fatness is usually assigned to a person with reference to the amount of fatty tissue that the person has. This "objective" value is still an implicit comparison to a standard. A person who would have been considered to

have a healthy, normal weight in an age which regarded plumpness as a virtue, would probably be considered fat by today's standards.

Other, more subjective considerations may add to the attribution of a quality. The adjective may make an implied comparison of the modified object or individual to a particular set of objects or individuals, instead of to a quasi-objective scale. If Milton were slightly overweight and standing next to a group of Olympic Marathon runners, one would be quicker to call him "fat" than if he were in the company of a number of people hospitalized for obesity.

Point of view can also effect quality attribution. An extremely thin person would probably be more willing to call a moderately overweight Milton "fat" than a person who is wildly obese. Again, the application of the adjective is relative, here to the standard of the attributer of the quality.

#### Adverbial Propositions

Adverbial propositions delimit some quality of a adjective, verb, or of a whole proposition. It is possible to modify an adjectival proposition with an adverb:

27. It is a very large nose.  
       (QUALITY OF, NOSE, LARGE)  
       (QUALIFY, (QUALITY OF, NOSE, LARGE), VERY);

or more simply:

(LARGE, NOSE)  
 (VERY, (LARGE, NOSE)).

Note that in this example, "VERY" qualifies the entire proposition (LARGE, NOSE).

Adverbial propositions can also qualify predicate propositions:

28. The snow is melting slowly.  
 (MELT, O:SNOW)  
 (QUALIFY, (MELT, SNOW), SLOW)

or

(SLOW, (MELT, SNOW))

The second proposition here describes the manner in which the action of the snow melting occurs. Adverbial propositions can be used to modify entire clauses, as in example 29:

29. Unfortunately, Myrna came to the party.  
 (COME, A:MYRNA, G:PARTY)  
 (QUALIFY, (COME, MYRNA, PARTY), UNFORTUNATE)

or

(UNFORTUNATE, (COME, MYRNA, PARTY))

#### Hedges

Hedges are considered to be qualifications of concepts or of propositions. They are used to soften or limit a statement. Their "job is to make things fuzzier or less fuzzy" (Lakoff, 1972).

#### EXAMPLES:

30. Richard Nixon is technically a Quaker.  
 (ISA, RICHARD NIXON, QUAKER)  
 (QUALIFY, (ISA, RICHARD NIXON, QUAKER), TECHNICALLY)

or

(TECHNICALLY, (ISA, RICHARD NIXON, QUAKER))

This excellent example is from Lakoff (1972). It illustrates a hedge of a proposition.

31. Hank dug with sort of a shovel.  
 (QUALIFY, SHOVEL, SORT OF)

or

(SORT OF, SHOVEL)  
 (DIG, A:HANK, I:(SORT OF, SHOVEL))



This is a hedge of a concept. The thing that Hank is digging with could be considered a shovel, but it is a "fuzzy" member of that concept category.

### 2.2.2 QUANTIFIER PROPOSITIONS

Some quantifiers are used to express the numerosity of distinct entities. These propositions are of the form

(NUMBER OF, A1, A2),

where A2 specifies the number, whether definite or indefinite, of the argument A1. Example 32 is an example in which a precise number is given. Example 33 expresses an imprecise quantity.

32. Ed bought three tickets.

(NUMBER OF, TICKETS, THREE)

(BUY, A:ED, O:(NUMBER OF, TICKETS, THREE))

33. All men are mortal.

(NUMBER OF, MAN, ALL)

(QUALITY OF, (NUMBER OF, MAN, ALL), MORTAL)

This can be simplified to

(ALL, MAN)

(MORTAL, (ALL, MAN)).

Among the quantifiers used to express an imprecise number are "all," "some," "many," "few" and "several." "Many" and "few" are defined relative to the size of the entire set being discussed. "Several" usually specifies some number between two and twelve, but is not defined in relative terms.

Another form of quantifier proposition is used to express the amount, the extent, or the degree of a single entity. This proposition takes the form

(EXTENT OF, A1, A2)

where the argument A1 is the whole entity being quantified, and A2 is its quantifier.

EXAMPLES:

34. Rita runs a long distance.  
 (EXTENT OF, DISTANCE, LONG)  
 (RUN, A:RITA, G:(EXTENT OF, DISTANCE, LONG))
35. Lyle stayed in Paris for three months.  
 (STAY, A:LYLE, O:PARIS)  
 (NUMBER OF, MONTHS, THREE)  
 (EXTENT OF,(STAY, LYLE, PARIS),(NUMBER OF, MONTHS, THREE))

2.2.3 PARTITIVE PROPOSITIONS

Partitive propositions express a part/whole relationship. They can take one of two forms. The first of these is

(PART OF, A1, A2),

where the argument A1 is a composite unit which has as one of its parts the argument A2. The other form for partitives is

(SOME OF, A1, A2),

where the argument A1 is a set of units, while the argument A2 is a subset of the set A1.

The syntax of the language often does not differentiate the part/whole relationship from that of simple possession. The phrases "Lucy's book" and "the car's bumper" look similar on superficial examination. However, the car without its bumper is not a complete car, while Lucy without her book is still the complete Lucy. Thus care must be taken to distinguish partitive relationships from possessives.

## EXAMPLES:

36. Mike dented the car's bumper.  
 (PART OF, CAR, BUMPER)  
 (DENT, A:MIKE, O:(PART OF, CAR, BUMPER))

## 2.2.4 NEGATIVE PROPOSITIONS

Negation expresses the modification of a proposition such that the existence or truth of that proposition is denied or is caused to become invalid. The logical definition of negation is the opposite of affirmation. A negated proposition, therefore, is the complement of that proposition. Only the proposition which is negated is not true. All other possibilities are left open. In logical terminology, this is expressed

$$(\text{NEGATE}, P) = 1 - P,$$

where 1 is the universe of possibilities and P is the proposition which is negated. Negative propositions take the form

$$(\text{NEGATE}, A1),$$

where the argument A1 is always a proposition. Only entire propositions can be negated.

37. Jay does not play the piano.  
 (PLAY, A:JAY, O:PIANO)  
 (NEGATE, (PLAY, A:JAY, O:PIANO))

The truth of the entire proposition is negated, not merely one of its arguments or its relation. Negating this proposition does not exclude other possibilities. Jay might be able to play the violin. He may be able to build pianos. His sister, Maggie, may be able to play the piano. None of these are eliminated from the realm of possibilities

by the proposition (NEGATE, (PLAY, JAY, PIANO)).

38a. We have no bananas.  
 (NUMBER OF, BANANAS, NONE)  
 (POSSESS, A:WE, O:(NUMBER OF, BANANAS, NONE))

The above example appears to have a negated argument, but on closer examination, it is a quantitative proposition with A2, the specified quantity, being null.

38b. We do not have bananas.  
 (POSSESS, A:WE, O:BANANAS)  
 (NEGATE, (POSSESS, A:WE, O:BANANAS))

Here the entire proposition is negated.

### 2.3 CONNECTIVE PROPOSITIONS

Connective propositions express the relationship between facts or propositions in a text. Their purpose is to coordinate the text propositions and to provide coherence to the text. Thus, the arguments of the connective propositions, more often than not, are other propositions in the text.

The relations used by these propositions are conjunctions, adverbs, or adverbial phrases. The relation that is generally used in the proposition is that which is given in the surface structure of the text. Not all connective relations are explicit. Sometimes they are deleted in the process of translating a propositional base structure into a text. For example,

39. The volcano erupted. Pompeii was buried in lava.

The connective proposition of causality has been omitted from the surface structure of this text. It is necessary to generate this proposition which relates (ERUPT, A:VOLCANO) and (BURY, A:VOLCANO,

O:POMPEII, I:LAVA) in order to understand this text. The section below entitled Causality deals with this kind of proposition.

At times, the given relation may appear to be ambiguous, due to multiple meanings of the word(s) in the text or to uncertain referents. If the meaning intended in the text can be clarified through the context or through common sense, then the surface representation of the relation can be replaced in the proposition by a word (or words) indicating the unambiguous meaning. (See the section below on multiple meanings of connective propositions.) If it is not possible to separate plausible meanings of the relation, propositions should be formed to represent all possible meanings. The text base, then, would branch out to the possible meanings in parallel, returning to a single text base when ambiguity ceases.

The eight major classes of connectives are:

I.	CONJUNCTION	expressing	union, association, combination
II.	DISJUNCTION	expressing	alternatives, opposition
III.	CAUSALITY	expressing	cause & effect, correlated events
IV.	PURPOSE	expressing	purpose, reason, intent
V.	CONCESSION	expressing	admission of a point, yielding
VI.	CONTRAST	expressing	divergence, comparison
VII.	CONDITION	expressing	prerequisites, restriction, qualification
VIII.	CIRCUMSTANCE	expressing	A. time B. location C. manner or mode of action

Each class can be divided into subclasses, such as the above divisions of the connective class Circumstance. The subdivisions stress one or another aspect of the superordinate class. Most subsets do not differ from each other to a degree warranting their notation. The division of Circumstance into time, location, and manner is the exception.

The arguments in a connective proposition must be within certain logical and pragmatic constraints. The context provides a set of

propositions or facts which are appropriate. Only those propositions which are relevant to the context may be connected.

The propositions or facts which are connected must be related to each other in some possible world. Their relationship may be through the identity of an individual in the arguments of the connected propositions, as in 40; through properties in common, as in 41; or through common actions, as in 42.

40. Gil caught a cab and went home.  
 (CATCH, A:GIL, O:CAB)  
 (GO, A:GIL, G:HOME)  
 (CONJUNCTION: AND, (CATCH, GIL, CAB), (GO, GIL, HOME))
41. The water is cold. So is the air.  
 (QUALITY OF, WATER, COLD)  
 (QUALITY OF, AIR, COLD)  
 (CONJUNCTION: [AL]SO, (QUALITY OF, WATER, COLD), (QUALITY OF, AIR, COLD))
42. Ron and Jeff are skiing.  
 (CONJUNCTION: AND, RON, JEFF)  
 (SKI, (CONJUNCTION: AND, RON, JEFF))

The connection may be temporal.

43. Yesterday Alice went to the library and drove to Denver for dinner.  
 (GO, A:ALICE, G:LIBRARY)  
 (DRIVE, A:ALICE, G:DENVER)  
 >(OBTAIN, A:ALICE, O:DINNER)  
 (PURPOSE: FOR, (DRIVE, A:ALICE, G:DENVER), >(OBTAIN, A:ALICE, O:DINNER))  
 (CONJUNCTION: AND, (GO, A:ALICE, G:LIBRARY), (DRIVE, A:ALICE, G:DENVER))  
 (TIME: DURING, (CONJ: AND, (GO, A:ALICE, G:LIBRARY), (DRIVE, A:ALICE, G:DENVER)), YESTERDAY)

— (The symbol ">" indicates an inferred proposition.)

The arguments may also be connected through their relevance to a particular topic of conversation.

44. Allan needed a wrench, but the hardware store was closed.  
 (NEED, A:ALLAN, O:WRENCH)  
 (CLOSE, A:\$, O:HARDWARE STORE)  
 (CONCESSION: BUT, (NEED, A:ALLAN, O:WRENCH), (CLOSE, A:\$, O:HARDWARE STORE))

The compatibility of these two conjuncts is due to their relationship to the topic of "buying a wrench." (But see pp. 50-53 for a more precise analysis of this kind of example.)

Each section below will give some of the particular constraints placed on possible arguments for the connective propositions of that class. In addition to these constraints which are peculiar to a given connective class, the arguments of these propositions must be relevant to the context as well as compatible with each other on logical and pragmatic grounds. For a more detailed discussion of relevance in connectives, see van Dijk (1976).

### 2.3.1 CONSTRUCTION OF CONNECTIVE PROPOSITIONS

#### Conjunction

Conjunctive propositions express the general relationship of connection or union of the specified arguments. Any number of arguments may be included in a conjunctive proposition, provided there are at least two. These propositions take the form

(CONJUNCTION: RELATION, A1, A2,...,An),

where the list of arguments are connected in the manner indicated by the relation. If the arguments of the relation are ordered or if they are unequal in importance, they should be listed in the specified order or in the order of lessening importance.

The most common conjunctive relation is the word "and." Some other examples of conjunctive relations are "in addition to," "also," and "along with." The arguments that are conjoined by the relation may be a number of simple arguments (nouns, verbs, etc.) or a number of propositions.

## EXAMPLES:

45. Dick and Jane went to the store.  
 (CONJUNCTION: AND, DICK, JANE)  
 (GO, A:(CONJ: AND, DICK, JANE), G:STORE)

This is an example of a conjunction which has simple arguments. The agent of the predicate relation "GO" is the conjunction of the two people.

46. Jimmy ate an orange and a banana.  
 (EAT, A:JIMMY, O:ORANGE)  
 (EAT, A:JIMMY, O:BANANA)  
 (CONJUNCTION: AND, (EAT, A:JIMMY, O:ORANGE), (EAT, A:JIMMY O: BANANA))

Note that the conjunction is between the two propositions of (EAT, A:JIMMY, O:\$). Presumably Jimmy did not eat the two fruits simultaneously, but first ate one and then the other. If he ate a fruit salad made of both oranges and bananas, the connective of this sentence would read, (CONJ: AND, ORANGE, BANANA).

47. Linda ran for Senator. Don did also.  
 (RUN, A:LINDA, G:SENATOR)  
 (RUN, A:DON, G:SENATOR)  
 (CONJUNCTION: ALSO, (RUN, A:LINDA, G:SENATOR), (RUN, A:DON, G:SENATOR))

Note that the connective here is between two sentences. Since the word "also" indicates the addition of one to another, the order of the two propositions is (RUN, A:LINDA, G:SENATOR), then (RUN, A:DON, G:SENATOR).

48. \*Carol went to the concert and Amy is a chiropodist.  
 (GO, A:CAROL, G:CONCERT)  
 (ISA, AMY, CHIROPODIST)  
 \*(CONJUNCTION: AND, (GO, CAROL, CONCERT),(ISA, AMY, CHIROPODIST))

(The asterisk indicates unacceptability.)  
 This sentence and the connective proposition are unacceptable because the two arguments of the connective are unrelated. Conjunctions must take arguments which are relevant to a particular topic and which are related to each other in some possible world.



Disjunction

Disjunctive propositions express the relationship of alternatives. In most cases of disjunction, the correct alternative is not known. Like conjunctives, these propositions are enumerative in nature, and may have two or more arguments. Disjunctive propositions take the form

(DISJUNCTION: RELATION, A1, A2, ..., An),

where the list of arguments are alternatives.

There are two kinds of disjunction, inclusive and exclusive. In inclusive disjunction, at least one of the arguments must be or can be true, and it is possible that more than one may be or can be true. An example of this would be the sentence

49. Sue and Bob are going to the movies, or perhaps just to dinner.

Sue and Bob are going to dinner, and perhaps also to the movies.

Exclusive disjunction refers to cases where one and only one of the possible alternatives must be or can be true. Most disjunctive propositions are exclusive. An example of this is

50. Marilyn is in London or New York.

Clearly Marilyn cannot be in both places at the same time. One and only one of these alternatives is true. (It is implied that one is true by the nature of disjunction.) The arguments which form the alternatives may be simple arguments, as in Conjunction, or they may be whole propositions, as in examples 49 and 50.

The most common disjunctive relation is the word "or." This may be used in either exclusive or inclusive disjunction. The most common relation indicative of inclusive disjunction is the form "and/or." The form "either...or" is used to denote a proposition which is

explicitly exclusive.

EXAMPLES:

51. Either Jack or Sarah broke the plate.  
 (BREAK, A:JACK, O:PLATE)  
 (BREAK, A:SARAH, O:PLATE)  
 (DISJUNCTION: EITHER...OR, (BREAK, A:JACK, O:PLATE),  
 (BREAK, A:SARAH, O:PLATE))

The sentence indicates that the plate is broken, that it is not known exactly who broke it, and that the suspects are Jack and Sarah.

Causality

Propositions expressing the relation of cause and effect or the relation between certain regularly correlated events are propositions of causality. They usually have two arguments, although occasionally they can take a third. The form taken by causal propositions is

(CAUSALITY: RELATION, A1, A2, [A3]),

where A1 is the causal agent or event, A2 is the effect or consequence of A1, and A3, if present, is the instrument of the causal relationship. Causality presupposes no intention on the part of the agent.

There are a large number of words and phrases which denote a causal relationship. The order of cause and effect in the surface structure of the sentence is dependent on the relation word or words which are used. "A1 causes A2," but "A2, because of A1." Similarly, "the consequence of A1 is A2", while "A1, consequently A2," and "A2 is the consequent of A1." Some other common relations which indicate causality are "as a result," "thus," or "therefore." Some relations which can represent causality, such as "that" or "as," have multiple meanings. Not all of the meanings of "that" or "as" express causality. At times, this can cause ambiguity in the text. Care should be taken to identify the meaning intended by the text, if possible, when propositionalizing texts containing

multi-meaning words.

EXAMPLES:

- 52a. Merinos are popular sheep because they have fine-quality wool.  
 (ISA, MERINOS, SHEEP)  
 (QUALITY OF, (ISA, MERINOS, SHEEP), POPULAR)  
 (QUALITY OF, WOOL, FINE-QUALITY)  
 (POSSESS, (ISA, MERINOS, SHEEP), (QUALITY OF, WOOL, FINE-QUALITY))  
 (CAUSALITY: BECAUSE, A1:(POSSESS, (ISA, MERINOS, SHEEP), (QUALITY OF, WOOL, FINE-QUALITY)), A2:(QUALITY OF, (ISA, MERINOS, SHEEP), POPULAR))

(The argument markers "A1," "A2," and when necessary "A3," is used for the rest of the connective propositions given below for the sake of clarity. It is not obligatory to include these markers when writing connective propositions.)

Although (QUALITY OF, (ISA, MERINOS, SHEEP), POPULAR) occurs earlier in the sentence, it is the second argument of the proposition. Using the identical surface relation, the same propositional text base can be expressed as

- 52b. Because Merinos have fine-quality wool, they are popular sheep.

Using the same connective class, and thus retaining the original meaning, many other texts are possible from this propositional text base:

- 52c. Because of their fine-quality wool, Merinos are popular sheep.  
 52d. Merino sheep have fine-quality wool, and consequently are popular.  
 52e. Merino sheep are popular as a result of their fine-quality wool.  
 52f. The fine-quality wool of the Merino sheep causes their popularity.

All of these sentences have identical bases, with the exception of the surface relation. The meaning is preserved throughout the transformation of surface structure. There may be slight differences in their connotation, however. The changes in the surface structure can also be used to place a particular emphasis on one or another argument or for stylistic purposes.

Note that there is no intention involved. The second argument is a logical consequent of the first argument.

53. Donna made Ted unhappy by refusing to come to the party.  
 (COME, A:DONNA, G:PARTY)  
 (REFUSE, A:DONNA, O:(COME, A:DONNA, G:PARTY))  
 (UNHAPPY, E:TED)  
 (CAUSALITY: MAKE, A1:DONNA, A2:(UNHAPPY, E:TED), A3:(REFUSE, A:DONNA, O:(COME, A:DONNA, G:PARTY)))

The optional third argument of the CAUSALITY frame is filled in this case. The means by which DONNA, the causal agent, causes (UNHAPPY, E:TED) is her refusal to come to the party. No intention to make Ted unhappy is attributed to Donna here, although it is possible to make an inference to that effect.

The inference of purpose, if formed, would be a separate proposition from the proposition of causality.

### Purpose

A connective proposition of purpose indicates an intentional act or event carried out (or to be carried out) for an express reason or to obtain a particular result. This differs from causality because it requires intention. The framework for purposive propositions is

(PURPOSE: RELATION, A1, A2),

where A1 indicates the act or event which is directed at the fulfillment of the intention, and A2 indicates the act or event which is the intended outcome or result.

Some relations which are used to express purpose are "in order to," "for the purpose of," "to," "for," "that," and "so." Often the same relational word can represent both causality and purpose. The difference between the two classes is, again, in the intention involved in Purpose. One act or event is deliberately carried out in order to permit the occurrence of a second act or event. In Causality, however, intention is not involved.

### EXAMPLES:

- 54a. Mark went to the beach to swim.  
 (GO, A:MARK, G:BEACH)  
 (SWIM, A:MARK)  
 (PURPOSE: TO, A1:(GO, A:MARK, G:BEACH), A2:(SWIM, A:MARK))

As in causal propositions, propositions of purpose can be expressed using any number of relations within the superordinate category of Purpose. The same text base can be expressed in many ways by varying the surface relation used. For example,

- 54b. Mark went to the beach in order to swim.  
 54c. The reason Mark went to the beach is to swim.  
 54d. Mark went to the beach so that he could swim.

Concession

A connective proposition of concession is one which expresses the admission of a point claimed in an argument or the yielding of a disputable contention. It sets up an argument, and then yields to the other side. The propositions take the form

(CONCESSION: RELATION, A1, A2),

where A1 is an argument which does not prevent the argument A2. A presupposition to concession is that A1 normally does prevent A2.

Some common relation words indicating concession are "but," "although," "while," "however," and "yet." The placement of the relation word in the text's surface structure may vary, while the argument order does not.

EXAMPLES:

55a. Although Doug studied for the exam, he failed.

(STUDY, A:DOUG, G:EXAM)

(FAIL, A:DOUG, O:EXAM)

(CONCESSION: ALTHOUGH, A1:(STUDY, A:DOUG, G:EXAM),  
A2:(FAIL A:DOUG, O:EXAM))

If Doug studies for the exam, one would presuppose that he would pass it. That he did not is an admission of a point against this presupposition.

Other surface variations are possible within the superordinate category of Concession:

55b. Doug studied for the exam, yet he failed it.

55c. In spite of the fact that Doug studied for it, he failed the exam.

55d. Doug studied for the exam, but he failed it.

55e. Even though he studied, Doug failed the exam.

56a. Notwithstanding his incompetence, Grant became President.

(QUALITY OF, GRANT, INCOMPETENCE)

(BECOME, A:GRANT, O:PRESIDENT)

(CONCESSION: NOTWITHSTANDING, A1:(QUALITY OF, GRANT,  
INCOMPETENCE), A2:(BECOME, A:GRANT, O:PRESIDENT))

Again, the same meaning can be expressed using other concessional relations.

56b. Grant is incompetent. Nevertheless he became President.

56c. Whereas Grant is incompetent, he became President.

- 56d. Grant is incompetent, however he became President.  
 56e. While he is incompetent, Grant became President.

Example 56e would be an ambiguous connective if the superordinate category of Concession were not clearly indicated or if the information conveyed by tense was not present.

Note that the order of arguments does not change with the transformations of surface structure.

### Contrast

Contrastive propositions express the relationship of divergence between objects or actions. The arguments being contrasted either have qualities belonging to the same category or are associated in some actual or assumed relationship. It is through comparison of like objects that divergent qualities or differences in degree are made noticeable. The frame for propositions of contrast is

(CONTRAST: RELATION, A1, A2),

where the result of the comparison of A1 and A2 is indicated by the relation, and where A1 is related to A2 by the specified relation.

The contrastive relationship rates one argument on a particular dimension using the second argument as the standard for that dimension. Possible relations that express comparative outcomes are

Greater than or More than;  
 Less than;  
 Equal to or the Same as;  
 Close to or Similar to;  
 Far from or Different from.

The comparison is often indicated by the suffix -er added to the dimension which is being contrasted.

The superlative is a special case of contrast, where the standard to which A1 is compared is a particular set, such as "all the boys in the third grade" or "all cities in the United States." The suffix

-est is often used to indicate the superlative, when it is added to the dimension being contrasted.

EXAMPLES:

57. Milton is fatter than Dan.  
(CONTRAST: FATTER, A1:MILTON, A2:DAN)

When Milton and Dan are compared on the dimension of fatness, Milton (A1) is fatter (is rated higher on the dimension of fatness) than Dan (A2).

Note that the comparison here is explicit. This differs from adjectival and adverbial modifiers which make implicit comparisons, such as in example 25.

58. Leslie has more money than Karen.  
(POSSESS, A:LESLIE, O:MONEY)  
(POSSESS, A:KAREN, O:MONEY)  
(CONTRAST: MORE THAN, A1:(POSSESS, A:LESLIE, O:MONEY),  
A2:(POSSESS, A:KAREN, O:MONEY))

59. Chicago is the largest city in the Midwest.  
(LOCATION: IN, A1:CITIES, A2:MIDWEST)  
(CONTRAST: LARGEST, A1:CHICAGO, A2:(LOCATION: IN,  
A1:CITIES, A2:MIDWEST))

In this connective proposition, Chicago's size (largeness) is being contrasted with the size of each of the other cities in the set of all cities located in the Midwest, and is rated higher than each on the dimension of largeness.

Condition

A conditional proposition expresses the relationship in which something limits or modifies the existence or the character of something else, or in which one thing's existence is a prerequisite for the occurrence of something else. The form taken by a conditional proposition is

(CONDITION: RELATION, A1, A2),

where A1 is a condition for the occurrence of A2 in the manner specified by the relation, or where A1 is a restriction or qualification placed on the occurrence or character of A2 as specified by the relation.

The most common expression of condition is "if...then" or simply

"if." Some other possible relations are "unless," "unless...not," or "except." As in Causality, the order of the arguments is dependent on the surface relation used.

EXAMPLES:

60. If Martha goes, Peter will go.  
 (GO, A:MARTHA G:\$)  
 (GO, A:PETER, G:\$)  
 (CONDITION:IF, A1:(GO, A:MARTHA, G:\$), A2: (GO, A:PETER, G:\$))
61. Lewis will not dance unless forced to.  
 (DANCE, A:LEWIS)  
 (FORCE, A:\$, O:(DANCE, A:LEWIS))  
 (CONDITION:UNLESS..NOT, A1:(FORCE, A:\$, O:(DANCE, A:LEWIS)), A2:(DANCE, A:LEWIS))

Circumstance

A proposition of circumstance expresses the state of affairs surrounding and affecting an agent or an action. Circumstance is divided into three subclasses: Time; Location; and Manner. Time relates one argument to a temporal framework. Location sets its arguments into a spatial framework. Manner describes the mode of one argument's occurrence. This last category appears to be somewhat superfluous for the present purposes, as will be discussed below in the section on Manner.

Time A proposition of time relates an event or occurrence to a temporal reference. The reference may be another event or occurrence, or it may be a specific point in time such as a date (July 4, 1976) or a pseudo-date (yesterday). The reference time may also be unspecified (some time, often). These propositions take the form

(TIME: RELATION, A1, A2),

where A1 occurs in the specified temporal relation to A2. A2 may be



another event or a temporal reference point, such as a date or pseudo-date.

The relation used for temporal propositions is indicative of the precise relationship of the argument A1 to the argument A2. Time relations can be roughly divided into Past, Present and Future, but some temporal concepts do not fit neatly into these categories (e.g. "yet," which spans past and present, "From now on," which spans present and future, and "always," which spans all three). Another relevant aspect of time is duration, of both A1 and A2 (e. g. "simultaneously," "continuously," "as"). A third aspect is the span of time separating the two arguments (e. g. "momentarily," "long before," "eventually"). Still another is frequency ("never," "always," "intermittently"). Because of these variations in the expression of the concept of time, it is not possible to randomly substitute one time relation for another, as is sometimes the case with other connective propositions such as those of concession or of causality.

#### EXAMPLES:

62. Larry went home after Nancy went home.  
 (GO, A:LARRY, G:HOME)  
 (GO, A:NANCY, G:HOME)  
 (TIME: AFTER, A1:(GO, A:LARRY, G:HOME), A2:(GO, A:NANCY, G:HOME)).

Note that the relationship between (GO, A:LARRY, G:HOME) and (GO, A:NANCY, G:HOME) is "after." To express the same meaning using the relation word "before," the proposition would read

(TIME: BEFORE, A1:(GO, A:NANCY, G:HOME), A2:(GO, A:LARRY, G:HOME)).

The order of the arguments is reversed here in order to maintain the correct temporal relationship between them.

63. Sandy was born in 1943.  
 (BORN, A:SANDY)  
 (TIME: IN, A1:(BORN, A:SANDY), A2:1943)

The relation word "in" means "during the inclusive period of"

in this example.

64. When Tim came home, Vicki had not yet returned from work.  
 (COME, A:TIM, G:HOME)  
 (RETURN, A:VICKI, S:WORK)  
 (TIME: WHEN...NOT YET, A1:(COME, A:TIM, G:HOME),  
 A2:(RETURN, A:VICKI, S:WORK))

The relation "When...not yet" indicated that at the occurrence of A1, A2 had not occurred, but that A2 would occur in the future.

Location Like time relations, relations expressing location are varied in the relationships they express. Locative propositions are used to place objects or events into a spatial framework. They take the form

(LOCATION: RELATION, A1, A2),

where A1 is in the specified relationship to A2. A2 is generally an object or location, while A1 could also be a proposition.

Connective propositions which are locative should not be confused with the source or goal cases which some predicate relations accept as a part of their verb frame. In particular, verbs denoting movement, such as "to come" or "to go," can take these cases. Example 65 provides an illustration. Note the difference between 65a and 65b.

EXAMPLES:

- 65a. Eileen bought a couch at Bloomingdale's.  
 (BUY, A:EILEEN, O:COUCH)  
 (LOCATION:AT, A1:(BUY, A:EILEEN, O:COUCH), A2:  
 BLOOMINGDALE'S)
- 65b. Eileen went to Bloomingdale's to buy a couch.  
 (GO, A:EILEEN, G:BLOOMINGDALE'S)  
 (BUY, A:EILEEN, O:COUCH)  
 (PURPOSE: TO, A1:(GO, A:EILEEN, G:BLOOMINGDALE'S),  
 A2:(BUY, A:EILEEN, O:COUCH))
66. Hans is standing behind the counter.  
 (STAND, A:HANS)  
 (LOCATION: BEHIND, A1:(STAND, A:HANS), A2:COUNTER)

Manner Connective propositions of manner express the mode or manner of an action or procedure. While Manner is certainly a major subordinate of the category of Circumstance, propositions of manner appear to be expressed equally well using propositions of qualification or of contrast. Therefore the use of manneristic propositions will be severely restricted in favor of those of qualification or contrast. Propositions of manner take the form

(MANNER: RELATION, A1, A2),

where A1 is a proposition which proceeds in the manner A2, as prescribed by the relation.

Manner refers specifically to a mode of action, procedure, or state. Thus the argument A1 is usually a predicate proposition. A2 is often a predicate proposition as well. At times, the arguments might be connective propositions, where the connected arguments are predicate propositions.

The relationship of manner is frequently one of comparison of procedures or states, where one argument occurs in the same or a similar way as the second argument's occurrence. Some possible relation words are "as," "like," "likewise," or "so." This relationship can be expressed using a contrastive proposition.

Propositions of manner can specify the particular way in which an argument occurs, without making a comparison. This is often expressed by a prepositional phrase using "with" or by an adverb. These propositions can be handled sufficiently with the use of a proposition of qualification.

## EXAMPLES:

67. Betsy skates like a professional.  
 (SKATE, A:BETSY)  
 (SKATE, A:PROFESSIONAL)  
 (MANNER:LIKE, A1:(SKATE, BETSY), A2:(SKATE, PROFESSIONAL))

or

(CONTRAST: LIKE, A1:(SKATE, BETSY), A2:(SKATE,  
 PROFESSIONAL))

Note that a proposition of contrast can be used here as well as one of manner. Manner makes explicit the dimension which is being compared, i.e. the mode of skating.

68. David answered the lawyer's questioning with care.  
 (QUESTION, A:LAWYER, O:DAVID)  
 (ANSWER, A:DAVID, O:(QUESTION, LAWYER, DAVID))  
 (MANNER: WITH, A1:(ANSWER, DAVID, (QUESTION, LAWYER,  
 DAVID)), A2:CARE)

or

(QUALIFY, (ANSWER, DAVID, (QUESTION, LAWYER, DAVID)),  
 CAREFULLY)

Since "with care" restricts the mode in which David answers the questions, the proposition of qualification is able to express the meaning as well as does the proposition of manner.

### 2.3.1 MULTIPLE MEANINGS--AMBIGUOUS CONNECTIVES

Some connectives with more than one meaning may fit into more than one class of connectives. This can cause the interpretation of the text to be ambiguous in some cases. In texts with more than one possible meaning, the potential reader's text base branches out into parallel propositional bases to account for all possible meanings.

An assumption has been made that the ideas of a speaker/writer are unambiguously represented in the base structure. The base structure is the propositional text base from which the text itself is derived. The actual ideas of the speaker/writer may be confused or contradictory. However, the base structure itself is certain in its intended meaning.

One source which allows text ambiguity to arise is the translation

of the base structure into the spoken or written text. Some propositions may be deleted from text representation because they are obvious to the speaker/writer from the context, because of stylistic considerations that the speaker/writer may have, or because of a desire on the part of the speaker/writer to be obscure.

In the derivation of a text from its base structure, particular words are chosen as tokens for particular word concepts. Since one word can conceivably represent more than one underlying concept, it is possible to choose a word to represent more than one meaning. In this case, the intended meanings are all represented unambiguously in the speaker/writer's base structure. Hence no ambiguity is present in the base structure of the text.

The listener/reader's interpretation of the surface text becomes the second possible source of ambiguity. In comprehending the text, the listener/reader forms a text base which is an interpretation of the original base structure of the text. During this process, misinterpretations can occur. The same token word may evoke different word concepts in different people. Omissions by the listener/reader may lead to differences between the base structure and the listener/reader's text base. The general world knowledge and semantic information contributed by persons in their attempt to comprehend a text may differ in both quantity and quality, thus leading to differences in the elaboration of the text in their text bases.

Surface expressions are usually not ambiguous when considered in the appropriate context, both linguistic and extralinguistic. Removing linguistic expressions from their context can produce ambiguity. This is not caused by ambiguity in the base structure which underlies the text, but by the absence of relevant contextual information which aids in the determination of the correct interpretation of the text.

Let us consider an example in detail. The word "as" has a number of meanings, which range from time and manner to purpose and causality. The following is a partial list of possible meanings for the word "as."

- 1:when, while
- 2:for the reason that
- 3:in or to the same degree in which
- 4:in the way or manner that

Suppose one wanted to propositionalize the following sentence:

69. As John came home, Mary left for work.  
 (LEAVE, A:MARY, G:WORK)  
 (COME, A:JOHN, G:HOME)  
 (CONNECTIVE?: AS, A1:?, A2:?)

The two predicate propositions are unambiguous. However, due to the several meanings of the word "as," the precise nature of the connection between the two is not clear. Since there is no degree involved in the predicate propositions, meanings 3 and 4 can be eliminated as not applicable. Without further context, neither 1 nor 2 can be ruled out. If the word "as" were to take the first meaning, the relationship between (LEAVE, A:MARY, G:WORK) and (COME, A:JOHN, G:HOME) would be a temporal one, indicating that the two events overlapped in time. The connective proposition would then read

(TIME: AS, A1:(LEAVE, A:MARY, G:WORK), A2:(COME, A:JOHN  
 G:HOME)).

If, on the other hand, the word "as" meant "for the reason that," the relationship between the two propositions is one of causality.

(CAUSALITY: AS, A1:(COME, A:JOHN, G:HOME), A2:(LEAVE,  
 A:MARY, G:WORK))

A text containing this example without clarification of the intended meaning for the word "as" would of necessity involve a branching out

of the text base into parallel segments. A text base for this example would appear as follows, where P represents a proposition in the text:

```

P1...
P2...
P3...
.
.
.
P(n)...
P(n+1):(LEAVE, A:MARY, G:WORK)
P(n+2):(COME, A:JOHN, G:HOME)
P(n+3):(TIME: AS, A1:(LEAVE, MARY, WORK), A2:(COME, JOHN, HOME))
P(n+3)^(CAUSALITY: AS, A1:(COME, JOHN, HOME), A2:(LEAVE, MARY, WORK))
P(n+4)...
.
.
.

```

Footnote: The superordinate connective category may be abbreviated when writing a connective proposition. The abbreviations to be used below are: I. CONJ; II. DISJ; III. CAUS; IV. PURP; V. CONC; VI. CONT; VII. COND; VIII. A. TIME; B. LOC; C. MANN.

### 3 THE STEP-BY-STEP CONSTRUCTION OF A TEXT BASE

A text is a coherent, connected discourse which is united by a common theme or topic. A text base is an ordered set of interrelated propositions representing the meaning underlying the text. It can represent the ideas of the speaker/writer which gave rise to the text, or the meaning derived from the text by a listener/reader. Unfortunately, these text bases are destined to remain to some extent unknown. Not all of the ideas of the speaker/writer find an expression in the surface representation of the text. Nor are all surface expressions intended. (People often say or write things that are not precisely what they mean, or that are subject to interpretations which they did not intend.) Listener/readers, for their part, do not process all parts of the text with equal care. In addition, they are subject to their own biases or misinterpretations.

Why, then, construct a text base at all? One cannot pretend to know all that the speaker/writer had in mind in constructing the text, nor can one unequivocally determine what meaning will be derived from a text by a given listener/reader. It is possible, however, to construct a text base which represents the meaning which should be derived by the ideal listener/reader. This text base would contain all propositions which are expressed directly in the text, as well as all the necessary inferences which must be made in order to connect the text propositions into a coherent and connected text base. (It will be assumed that this idealized listener/reader will make at least these inferences, in addition to some number of less crucial ones, in the course of translating the text into its meaning.) This ideal text base can then serve several functions, for example, as a template for a listener/reader's report of the meaning of the text, or as a tool to be used in the study of the theoretical issues involved in text structure and comprehension.



To construct a template text base, then, all one has to do is to identify the meaning of a text. However this is not as easy as it may sound. Problems of interpretation somehow always manage to arise. It cannot be claimed that the system of propositionalizing a text which is outlined here can account for, explain, or describe all meaning in all texts to everyone's satisfaction. If it could, many of the troublesome problems in the area of language comprehension would be solved. In the absence of a complete understanding of the processes involved in language comprehension, some problems will have to remain problems for the present.

The following section will detail the analysis of a number of texts into their template text bases. The major emphasis will be on the process of construction, although several problem areas will be touched on during the course of the analyses.

The specific procedure used below need not be followed to obtain satisfactory results. The order in which the propositions are extracted from the text has been systematized here for the sake of organization and clarity. The policy of making underlying assumptions and relations explicit will be continued for the most part. Some shortcuts will be introduced throughout the section. Not all of them will be used here, however, so as to avoid possible confusion.

### 3.1 THE ORDERING OF PROPOSITIONS IN A TEXT BASE

When attempting to propositionalize a text, it is best to first read over the entire text carefully. Then, if it is a long text, reread the section of the text which will be analyzed first. Begin with the first sentence. If it is a complex sentence (one which has a main clause and at least one subordinate clause) or a compound sentence (one which has more than one main clause), do not attempt to work on all of it at once.

Choose one main clause as a starting point. Loosely determine the propositions within it. Then write them down according to the format set forth in the section on the construction of propositions.

The order in which propositions will be written down in this work is systematized to some extent. The basic rule which has been followed is

→ Ordering Rule: A proposition "A" which is embedded in another proposition "B" is written before that proposition "B" whenever possible.

This rule can be expanded to an algorithm with an nine-branch loop:

- 1) MODIFIED ARGUMENTS OF PREDICATE PROPOSITIONS  
Write all modifier propositions which are used in a predicate proposition to modify an argument of that proposition. Modifiers of the predicate proposition itself are not included here.
- 2) CONNECTED ARGUMENTS OF PREDICATE PROPOSITIONS  
Write all connective propositions which are used as arguments for a predicate proposition, and which occupy the same case position.
- 3) PREDICATE PROPOSITIONS  
Write the predicate proposition(s). There is rarely more than one per clause.
- 4) MODIFIERS OF PREDICATE PROPOSITIONS  
Write all modifiers of the predicate proposition(s).
- 5) MODIFIED ARGUMENTS OF CIRCUMSTANTIAL PROPOSITIONS  
Write all modifiers of arguments of propositions of circumstance.
- 6) CIRCUMSTANTIAL PROPOSITIONS  
Write all connective propositions of circumstance, provided they do not connect arguments between clauses.
- 7) OTHER CONNECTIVE PROPOSITIONS WITHIN CLAUSE  
Write all connective propositions which connect arguments within the clause, and which have not been transcribed earlier.
- 8) CONNECTIVE PROPOSITIONS BETWEEN CLAUSES  
Write all connective propositions which connect an argument or arguments from the present clause with an argument or arguments in a previous clause or clauses. Forward-looking connectives from the previous clause are included here.
- 9) REPEAT  
Repeat steps 1-8 for the next unit of text.

To construct a template text base, then, all one has to do is to identify the meaning of a text. However this is not as easy as it may sound. Problems of interpretation somehow always manage to arise. It cannot be claimed that the system of propositionalizing a text which is outlined here can account for, explain, or describe all meaning in all texts to everyone's satisfaction. If it could, many of the troublesome problems in the area of language comprehension would be solved. In the absence of a complete understanding of the processes involved in language comprehension, some problems will have to remain problems for the present.

The following section will detail the analysis of a number of texts into their template text bases. The major emphasis will be on the process of construction, although several problem areas will be touched on during the course of the analyses.

The specific procedure used below need not be followed to obtain satisfactory results. The order in which the propositions are extracted from the text has been systematized here for the sake of organization and clarity. The policy of making underlying assumptions and relations explicit will be continued for the most part. Some shortcuts will be introduced throughout the section. Not all of them will be used here, however, so as to avoid possible confusion.

### 3.1 THE ORDERING OF PROPOSITIONS IN A TEXT BASE

When attempting to propositionalize a text, it is best to first read over the entire text carefully. Then, if it is a long text, reread the section of the text which will be analyzed first. Begin with the first sentence. If it is a complex sentence (one which has a main clause and at least one subordinate clause) or a compound sentence (one which has more than one main clause), do not attempt to work on all of it at once.

Choose one main clause as a starting point. Loosely determine the propositions within it. Then write them down according to the format set forth in the section on the construction of propositions.

The order in which propositions will be written down in this work is systematized to some extent. The basic rule which has been followed is

→ Ordering Rule: A proposition "A" which is embedded in another proposition "B" is written before that proposition "B" whenever possible.

This rule can be expanded to an algorithm with an nine-branch loop:

- 1) MODIFIED ARGUMENTS OF PREDICATE PROPOSITIONS  
Write all modifier propositions which are used in a predicate proposition to modify an argument of that proposition. Modifiers of the predicate proposition itself are not included here.
- 2) CONNECTED ARGUMENTS OF PREDICATE PROPOSITIONS  
Write all connective propositions which are used as arguments for a predicate proposition, and which occupy the same case position.
- 3) PREDICATE PROPOSITIONS  
Write the predicate proposition(s). There is rarely more than one per clause.
- 4) MODIFIERS OF PREDICATE PROPOSITIONS  
Write all modifiers of the predicate proposition(s).
- 5) MODIFIED ARGUMENTS OF CIRCUMSTANTIAL PROPOSITIONS  
Write all modifiers of arguments of propositions of circumstance.
- 6) CIRCUMSTANTIAL PROPOSITIONS  
Write all connective propositions of circumstance, provided they do not connect arguments between clauses.
- 7) OTHER CONNECTIVE PROPOSITIONS WITHIN CLAUSE  
Write all connective propositions which connect arguments within the clause, and which have not been transcribed earlier.
- 8) CONNECTIVE PROPOSITIONS BETWEEN CLAUSES  
Write all connective propositions which connect an argument or arguments from the present clause with an argument or arguments in a previous clause or clauses. Forward-looking connectives from the previous clause are included here.
- 9) REPEAT  
Repeat steps 1-8 for the next unit of text.

Not all these steps are necessary for a given unit of text. One should return to step 1 of the loop as soon as all propositions in a text unit have been listed. Again, this system is used here mainly for the sake of an orderly discussion of the process by which a text base is derived. There is no claim made that it is the only method, nor even the best one. It is merely one way of organizing the ordering of propositions to result in a consistent order.

### 3.2 A NIGHT ON THE TOWN

The first text to be analyzed is a simple story. It will be used to illustrate the order of analysis.

#### 3.2.1 TEXT

Louise and Ann went to the movies last night. They met Charlie there. Afterwards they all went for a chocolate sundae, but the ice cream parlor was closed.

#### 3.2.2 PROCEDURE

First the text is read. The first sentence is reread. A loose analysis determines that this sentence consists of a conjunction of agents (i.e. a compound subject), a predicate proposition, and a modified temporal proposition. Since there are no modified arguments of the predicate proposition, step 1 is skipped. Step 2 governs connected arguments. By applying this step, the first proposition is

1. (CONJ: AND, LOUISE, ANN).

The convention of numbering the propositions in the text base with their line number has been carried over from Kintsch (1974). This notation facilitates the referencing of a proposition. When embedding a proposition in another, the number of the embedded proposition will be used in

its place.

Step three concerns predicate propositions. The predicate relation in the sentence is the verb "To Go." Its verb frame is

(GO, A:\$, [I:\$], S:\$, G:\$),

that is, "Someone goes by some means from somewhere to somewhere." The agent case is obligatory, while the instrumental case is optional. The use of either the source or the goal case is necessary, but the inclusion of both cases is optional. The agent in the sentence is the conjunction of Louise and Ann. In the surface structure of the text, the instrument and source cases have not been included. The goal is the movies.

2. (GO, A:1, I:\$, S:\$, G:MOVIES)

Optional unfilled cases need not be included in the proposition.

Therefore Proposition 2 (P2) may be simplified to read

2. (GO, A:1, G:MOVIES).

Often the cases of the arguments in a proposition are obvious without reference to the case markers (e.g. A:, I:, O:, etc.). If the proposition is unambiguous without using case markers, they may be omitted. For clarity, case markers will continue to be used in the present work. Argument markers (e.g. A1:, A2:) need not be included.

The application of steps 5 and 6 focuses on the temporal phrase "last night." A modifier proposition,

3. (QUALIFY, NIGHT, LAST)

is needed to specify which night is being discussed. While P3 clearly is temporally related to P2, the specific relationship is not mentioned

in the surface structure. However, by default, the temporal relationship is assumed to be "during" or "at the same time."

4. (TIME: >DURING, 2, 3)

The symbol ">" indicates that the relation that is used here is inferred (This symbol can also be used to indicate that an entire proposition is inferred.) This proposition could also be written

4. (TIME: Ø, 2, 3)

where the symbol "Ø" indicates that the actual relation is absent in the surface structure. However, it is reasonable to assume from a knowledge of the language that the absence of a specific temporal relation indicates simultaneity.

The propositions of the first sentence are all transcribed at this point. The analysis proceeds to the next unit of text, beginning with step 1.

The next sentence (They met Charlie there.) is expanded, that is, the pronouns are replaced with referents<sup>1</sup>, to read

Louise and Ann met Charlie at the movies.

This sentence can be analyzed into three propositions: the conjuncted agent (Louise and Ann), which has been included in our proposition list already and need not be included again; a predicate proposition; and a

---

1

Strictly speaking, replacement of pronouns with their referents involves making inferences as to their identity. These inferences are based on the syntax and on world knowledge. This would involve the inclusion of two additional implied propositions in the text base.

>A. (REF, THEY, 1)  
>B. (REF, THERE, MOVIES)

In most cases, the referents are clear and unambiguous. Therefore, the problem of inferred referents of pronouns will be ignored hereafter.

locative proposition.

Step 1 (modified arguments) and step 2 (connected arguments) do not apply to this sentence. Step 3 directs attention to the predicate proposition, with the relation "to meet." The frame for this verb is

(MEET, A:\$, O:\$),

or "Someone meets someone else." The location of the meeting is not a part of the verb frame and must be included as a separate proposition, using step 6 (connectives of circumstance). Steps 4 and 5 do not apply.

5. (MEET, A:1, O:CHARLIE)
6. (LOC: AT, 5, MOVIES)

The propositions in the second sentence are all transcribed at this point. Steps 7 and 8 do not apply.

The next unit of text is a complex sentence. The subject "they all" is replaced with its referent, so that the sentence reads

Afterwards Louise, Ann, and Charlie went for a chocolate sundae, but the ice cream parlor was closed.

A loose analysis reveals that the sentence consists of a main clause (they went for a chocolate sundae), a subordinate clause (the ice cream parlor was closed), a connective between clauses (but), and a connective between sentences (afterwards). The main clause will be examined first.

On closer examination, the main clause consists of a purposive action (going somewhere) and a goal (a chocolate sundae). The actual goal is some further action, that of obtaining and eating the sundae, as one familiar with the joys of ice cream sundaes will recognize. Therefore, underlying the phrase "for a chocolate sundae" is a subordinate clause. We will delay consideration of this clause until



the main clause has been analyzed.

Using steps 2 and 3, the propositions of the main clause are

7. (CONJ: AND, LOUISE, ANN, CHARLIE)
8. (GO, A:7, >S:MOVIES, >G:(QUALIFY, PARLOR, ICE CREAM))

The source and goal cases in P8 are inferred from other parts of the text. Sentence 2 locates the three at the movies. In the sentence following they go somewhere to get a sundae. Since no intervening action is mentioned, the implication is that they left directly from the movies. The implied goal will be justified later.

We return now to the subordinate phrase "for a chocolate sundae." Expanding the phrase and clarifying the underlying meaning will yield

...so that Louise, Ann, and Charlie could get a chocolate sundae.

This makes it explicit that the goal is not the sundae, but getting (and eating) the sundae. It also shows that the relationship between going somewhere (to the ice cream parlor) and getting the sundae is one of purpose. Using steps 1, 3, and 8, the propositions are

9. (QUALITY OF, SUNDAE, CHOCOLATE)
- >10. (GET, A:7, O:9)
11. (PURP: FOR, 8, >10)

P10 is a necessary inference in order to understand the purpose of going to the ice cream parlor.

The last segment of the text is the subordinate clause "but the ice cream parlor was closed." Although there has been no previous mention of an ice cream parlor, the definite article is used. This indicates that the first mention of the ice cream parlor has been deleted from the text's surface structure. The second mention, the first in the text, retains the definite article. In the main clause, the three people go somewhere to get ice cream. It is reasonable to assume that they went

to an ice cream parlor. Therefore the mention of one here is not surprising, and it is inferred that the goal case in P8 is the ice cream parlor.

A loose analysis of the clause shows that it consists of a modifier proposition (ice cream parlor), a predicate proposition (the ice cream parlor was closed), and a connective proposition (but...). Using steps 1 and 3, the first two propositions are

- 12. (QUALIFY, PARLOR, ICE CREAM)
- 13. (CLOSE, A:\$, O:12).

The two word phrase "ice cream" is presumed to be a single concept.

The two parts of the sentence, represented by P11 and P13, are connected by the concession relation "but" in the surface structure. However, the connection as it stands violates the constraints of a concession. Recall that a concession consists of two parts; the first sets up an argument, while the second part yields to the other side. A presupposition is that the first argument normally prevents the second. Going for a sundae does not prevent the ice cream parlor from being closed. In order to make sense out of this concessional relationship, several propositions must be inferred. First, they did not get sundaes after all, and second, the fact that the ice cream parlor was closed had that consequence. With these two inferences it is now possible to meet the constraints of the concession. The argument that is set up is that they went to the ice cream parlor with the intention of getting sundaes. The argument that is yielded is that they did not get any sundaes. The proposition of causality is necessary to maintain a cohesive text base, connecting P12 with the inferred proposition that they did not get sundaes. Using steps 3 and 7, the inferred propositions are

- >14. (NEGATE, >10)
- >15. (CAUS: BECAUSE, 13, >14).

The concession, by step 8, is

16. (CONC: BUT, 11, >14).

Lastly, the temporal connective between sentences (afterwards) links the first topic (going to the movies) with the second topic (going for sundaes). The exact proposition to be used as the first argument is unclear, however. What can be known about the temporal sequence of events? P5 (MEET, A:1, O:CHARLIE) could have occurred before, after, or during the movies. From general knowledge about movie going, it can be assumed that Louise, Ann, and Charlie stayed at the movie until its end. Therefore the text could mean that they went for sundaes after the meeting, after the movie's end, or both. A text base which aims at determining the meaning of this text, would have to allow for all of these possibilities.

17. (TIME: AFTERWARDS, 5, 11)

18. ∅

19. ∅

20. ∅

>17a. (END, MOVIES)

18a. (TIME: AFTERWARDS, >17a, 11)

19a. ∅

20a. ∅

>17b. (END, MOVIES)

18b. (TIME: AFTERWARDS, 5, 11)

19b. (TIME: AFTERWARDS, >17b, 11)

20b. (CONJ: ALSO, 18b, 19b)

All the propositions of the text are now transcribed.

### 3.2.3 THE TEMPLATE TEXT BASE

The entire text base for this text is:

1. (CONJ: AND, LOUISE, ANN)
2. (GO, A:1, G:MOVIES)
3. (QUALIFY, NIGHT, LAST)
4. (TIME: >DURING, 2, 3)
5. (MEET, A:1, O:CHARLIE)
6. (LOC: AT, 5, MOVIES)

- 7. (CONJ: AND, LOUISE, ANN, CHARLIE)
- 8. (GO, A:7, >S:MOVIES, >G:12)
- 9. (QUALITY OF, SUNDAE, CHOCOLATE)
- >10. (GET, A:7, O:9)
- 11. (PURP: FOR, 8, >10)
- 12. (QUALIFY, PARLOR, ICE CREAM)
- 13. (CLOSE, A:\$, O:12)
- >14. (NEGATE, >10)
- >15. (CAUS: BECAUSE, 13, >14)
- 16. (CONC: BUT, 11, >14)
- 17. (TIME: AFTERWARDS, 5, 11)
- 18. Ø
- 19. Ø
- 20. Ø
  
- >17a. (END, A:MOVIES)
- 18a. (TIME: AFTERWARDS, >17a, 11)
- 19a. Ø
- 20a. Ø
  
- >17b. (END, A:MOVIES)
- 18b. (TIME: AFTERWARDS, 5, 11)
- 19b. (TIME: AFTERWARDS, 11, >17b)
- 20b. (CONJ: ALSO, 18b, 19b)

#### 3.2.4 THE BASE STRUCTURE

The base structure from which a text is derived generally contains a great deal more information than is contained in the text itself. A possible base structure for the above text, for example, might contain such information as where Louise and Ann left from to go to the movies, how they got there, what movie they saw, exactly when and where they met Charlie, and so on. Presumably, the propositions which are inferred in the template text base above would be present in the base structure. The temporal ambiguity would not be present in the base structure. The base structure would contain many more specifics and greater detail.

In general, the actual base structure of a text remains unknown. Some ideas may be deleted from expression in the surface text. Other ideas may take on forms which do not reflect the intended meaning. It should be clear that speculation on precise base structures is beyond the scope of this paper. Instead, the focus is on the template text base, which is to a greater extent knowable.

### 3.3 MERINO SHEEP

The next text to be analyzed is a simple descriptive text. It represents the type of paragraph which is often used in a memory experiment.

#### 3.3.1 TEXT

Three fourths of Australia's sheep are pure Merinos. They are popular because of the large amount of semiarid land. The Merino thrives on the grasses and low bushes which grow on semiarid plains. They are known for their heavy fleeces of fine-quality wool which bring a high price from textile manufacturers.

#### 3.3.2 PROCEDURE

In the first sentence, the possessive "Australia's sheep" is an embedded predicate proposition, forming as a unit one argument. As such it is written first.

1. (POSSESS, A:AUSTRALIA, O:SHEEP)

The rest of the sentence consists of two modifier propositions--one a qualifier, one a quantifier--and two set inclusion propositions.

2. (NUMBER OF, 1, THREE FOURTHS)
3. (QUALIFY, MERINOS, PURE)
4. (ISA, 2, 3)
5. (ISA, MERINOS, SHEEP)

The second sentence is made up of two clauses which are connected by a causal relation. The first clause is a simple predicate proposition.

6. (QUALITY OF, 3, POPULAR)

The proposition

- \*6. (QUALITY OF, 4, POPULAR)

is considered unacceptable as an alternative for several reasons. First, the closest potential referent in the surface structure is "pure Merinos." Second, it seems more sensible to talk about a breed of sheep in general being popular as opposed to a specific set of sheep belonging to this breed. Therefore the proposition \*P6 is not as plausible an interpretation as P6.

The next clause, expanded, is

[There is a] large amount of semiarid land.

The propositions of the clause are:

- 7. (QUALIFY, LAND, SEMIARID)
- 8. (EXTENT OF, 7, LARGE)

The connective is simply

- 9. (CAUS: BECAUSE OF, 8, 6).

The next sentence is composed of a main clause and a subordinate modifying clause. The main clause consists of a modifier proposition (low bushes), a conjunction (grasses and bushes), and a predicate proposition (Merinos thrive...). Since the means by which Merinos thrive is the particular grasses and bushes that are on the semiarid plains, the subordinate clause will be propositionalized before the predicate proposition (Merinos thrive...). The subordinate clause contains a modifier (semiarid plains), and a locative (they grow on plains). The word "grow" is a locative relation and does not carry any additional semantic content here.

- 10. (QUALIFY, BUSHES, LOW)
- 11. (CONJ: AND, GRASSES, 10)
- 12. (QUALIFY, PLAINS, SEMIARID)
- 13. (LOC: ON, 11, 12)
- 14. (THRIVE, A:3, I:13)

The last sentence has a main and a subordinate clause. The main clause consists of three modifiers (fine-quality wool, heavy fleeces, fleeces of wool), a possessive (Merinos have fleeces), and a predicate proposition (Merinos are known for...).

- 15. (QUALITY OF, WOOL, FINE-QUALITY)
- 16. (QUALITY OF, FLEECES, HEAVY)
- 17. (QUALITY OF, 16, 15)
- 18. (POSSESS, A:2, O:17)
- 19. (KNOW, A:\$, O:18)

The subordinate clause, modifying the fleeces of wool, is made up of a qualifier (high price), and two predicate propositions (textile manufacturers, fleeces bring prices). Although the phrase "textile manufacturers" superficially corresponds to a modifier, it does not at a semantic level. It does not mean that a quality of the manufacturers is "textile," but that textiles are what is made by the manufacturers.

- 20. (QUALIFY, PRICE, HIGH)
- 21. (MAKE, A:MANUFACTURERS, O:TEXTILES)
- 22. (BRING, A:18, O:20, S:21)

All the propositions in the text have been written down at this point.

This is a relatively simple text which does not draw heavily on the inferential processes. Often it is desired to construct texts of this nature for experiments in psychology. A text to be used as material should be simple and straightforward to meet most required specifications. The analysis of a text can guide in the task of rewriting a text to fit an experimenter's needs.

## 3.3.3 THE TEMPLATE TEXT BASE

The complete text base is as follows:

1. (POSSESS, A:AUSTRALIA, O:SHEEP)
2. (NUMBER OF, 1, THREE FOURTHS)
3. (QUALIFY, MERINO, PURE)
4. (ISA, 2, 3)
5. (ISA, MERINOS, SHEEP)
6. (QUALITY OF, 3, POPULAR)
7. (QUALIFY, LAND, SEMIARID)
8. (EXTENT OF, 7, LARGE)
9. (CAUS: BECAUSE OF, 8, 6)
10. (QUALIFY, BUSHES, LOW)
11. (CONJ: AND, GRASSES, 10)
12. (QUALIFY, PLAINS, SEMIARID)
13. (LOC: ON, 11, 12)
14. (THRIVE, A:3, I:13)
15. (QUALITY OF, WOOL, FINE-QUALITY)
16. (QUALITY OF, FLEECES, HEAVY)
17. (QUALITY OF, 16, 15)
18. (POSSESS, A:2, O:17)
19. (KNOW, A:\$, O:18)
20. (QUALIFY, PRICE, HIGH)
21. (MAKE, A:MANUFACTURERS, O:TEXTILES)
22. (BRING, A:18, O:20, S:21)



### 3.4 THE SUN ALSO RISES (EXCERPT)

The examples detailed above have illustrated the step-by-step construction of a text base in relatively unequivocal terms. Those paragraphs were written expressly for use in a propositional analysis of connected discourse. Most texts with which readers must deal however, are less precise and less well-defined than are those analyzed above. This is particularly true of non-technical and/or less rigidly-structured material (e.g. literary prose, dialogue, folklore). In those cases, the relationship between the syntax of the passage and its underlying template text base is not always a direct one. Often, much more must be interpreted or inferred by the reader before full comprehension can be attained. Authors of literary fiction, for example, may make seemingly incredible demands on a reader. These authors are not bound by journalistic constraints which have the purpose of simplifying the communication process. They are free to employ various literary devices, such as metaphors, symbolism, and other idiosyncratic expressions. For this reason, propositional analysis of such material can often become complicated. Many inferences must be generated by the reader. Any major ambiguity must be satisfactorily resolved. Interpretation of literary devices must be undertaken to accomplish full comprehension. In other words, the task before the reader becomes a more active process when dealing with a text which is less rigidly constructed. We will not attempt a detailed examination of inferential processes or literary analysis. Such an undertaking is clearly beyond the scope of this manual. Rather, we will utilize a passage from Ernest Hemingway's The Sun Also Rises to demonstrate the more problematic, but still possible construction of a text base from literary prose.

### 3.4.1 TEXT

...I unbolted the door and went out. It was cool outside in the early morning, and the sun had not yet dried the dew that had come when the wind died down. I hunted around in the shed behind the inn and found a sort of mattock, and went down toward the stream to try and dig some worms for bait. The stream was clear and shallow but it did not look trouty. On the grassy bank where it was damp, I drove the mattock into the earth and loosened a chunk of sod. There were worms underneath. They slid out of sight as I lifted the sod and I dug carefully and got a good many.

### 3.4.2 PROCEDURE

We will analyze this paragraph alone and out of context. Clearly, some connections must exist between this passage and those that precede and follow it, but they will be disregarded as irrelevant for our purposes.

The first sentence is relatively explicit, consisting of two predicate propositions (I unbolted the door, I went out), and an ambiguous connective between them (and).

The connective "and" can be interpreted in many different ways. The most common meaning of the word is simply to indicate connection or addition. However, it has other potential meanings. One is to indicate a temporal sequence. Another is to indicate a causal relationship. Still another is to indicate purpose. These uses of the word "and" are not often encountered in the writings of professional journalists, textbook writers, or literary authors. Nonetheless, in everyday communication, they are quite common. In particular, they are often found in the recall protocols of subjects in text memory experiments.

The use of "and" in this particular passage is ambiguous. It may simply conjoin the two propositions to indicate that the author undertook two separate actions. Alternatively, it may indicate a

temporal sequence of events, i.e. that the author's exit occurred after he had unbolted the door. In this case, the relation "after" can also be inferred from our knowledge of the world. The door must be ajar before one can go out. Either of these interpretations are valid ones, thus both should be included in the text base.

1. (UNBOLT, A:I, O:DOOR)
2. (GO, A:I, G:OUT)
3. (CONJ: AND, 1, 2) &/OR (TIME: AFTER, 2, 1)

Sentence two is more complex: there are two main clauses (It was cool outside in the early morning, the sun had not yet dried the dew), a subordinate clause modifying the second (that had come when the wind died down), and a connective joining the two (and). Beginning with the first main clause, there is a qualifier (early morning), a temporal connective (in the morning), a locative (outside), and an idiomatic expression for the weather (it was cool). Idioms are acceptable English expressions that do not conform to the rules of language construction. The semantic interpretation of the idiom is used in the propositional text base. The surface form of the idiom is ignored.

---

2

Idiomatic expressions may be of two types:

A. An expression that is peculiar to itself in grammatical construction, and that cannot be analyzed or justified grammatically. The "quality of weather" expression is one such idiom. Other examples:

- Example 1. "... for many a year."  
 Example 2. "No, it isn't me."

B. An expression whose meaning cannot be derived from the combined meaning of its elements. These idioms cannot be interpreted literally. Examples:

- Example 3. "I completely lost my head."  
 Example 4. "How are you?"

When such expressions occur in the text, they should be proposition-

4. (QUALITY OF, WEATHER, COOL)
5. (QUALIFY, MORNING, EARLY)
6. (TIME: IN, 4, 5)
7. (LOC: OUTSIDE, 6, \$) OR (LOC: OUTSIDE, 6, >INN)

The second main clause consists of a predicate proposition (the sun dried the dew), and a temporal connective (...not yet...). The subordinate clause modifies the dew in the main clause. It is made up of two predicate propositions (wind died down, dew came [to the ground]), and a temporal connective between the two predicates (when the wind died down, the dew came). Since the subordinate clause modifies part of the main clause, its propositions will be written first.

8. (DIE DOWN, A:WIND)
9. (COME, A:DEW, >G:GROUND)
10. (TIME: WHEN, 9, 8)
11. (DRY, A:SUN, O:10)
12. (TIME: NOT YET, 2, 11)
13. (CONJ: AND, 7, 12) &/OR (CAUS: CAUSE, 12, 7)

P9 contains the implied argument "ground," since the verb frame of

---

alized in terms of their semantic content, and not in a literal manner.  
Examples:

Example 1. "...for many a year."  
(NUMBER OF, YEARS, MANY)

Example 2. "No, it isn't me."  
2a. (REF, IT, ME)  
2b. (NEGATE, 2a)

Example 3. "I completely lost my head."  
3a. (LOSE, A:I, O:CONTROL)  
3b. (QUALIFY, 3a, COMPLETELY)

The idiom "to lose one's head" conveys the meaning that someone lost reasoning facilities or control capabilities. The proposition is constructed to reflect the semantics, not the syntax or surface expression.

Example 4. "How are you?"  
4a. (POSSESS, A:\$, O:STATE OF HEALTH)  
4b. (QUESTION: WHAT, 4a)

The speaker's intention here is to determine the state of the listener's health. The text base will include only this semantic explanation. The literal form of the idiom has been discarded.

"to come" requires the presence of a source and/or goal case. Often, the source or goal case can be inferred from the context.

The time relation expressed in P12 is a complicated one--it states a temporal relationship between the drying of the dew and the exiting of the narrator (refer to P2). When the narrator went out, the sun had not yet dried the dew. The causal connective of P13 presupposes some knowledge of weather conditions. If it was early morning and the sun had not dried the dew, then the air was undoubtedly damp and the sun was not warm. Consequently, the narrator felt coolness.

Sentence three (I hunted around in the shed behind the inn and found a sort of mattock and went down toward the stream to try and dig some worms for bait.) consists of four clauses and a number of ambiguous connectives. The first clause has a predicate proposition (I hunted around...) and two locatives (in the shed, behind the inn). The predicate "to hunt around" requires the use of an instrument and/or goal case. One hunts around either by means of some instrument (I hunted around with a flashlight.) or one hunts around in order to find some object (I hunted around for the shovel.).

- 14. (HUNT AROUND, A:I, G:\$)
- 15. (LOC: BEHIND, SHED, INN)
- 16. (LOC: IN, 14, 15)

The second clause has a hedge (sort of mattock), a predicate proposition (I found...), and an ambiguous connective to the previous clause (and). The two arguments of the connective are the actions of the narrator, the acts of "hunting around" and of "finding." Since these are distinct, although not independent, activities, it is possible to interpret the connective "and" as a relatively unprocessed simple conjunctive relation. (The search continued until the mattock was discovered, so the temporal relation "until" is also a legitimate connective

between the two acts.) Finally, some causal relationship exists between the narrator's hunting in the shed and his discovery of the mattock, so a relationship of causality is a third possible interpretation of the connective. Careful analysis of the connective in this sentence indicates that any or all of the three interpretations are psychologically valid. Since the word "and" can be used to indicate any of these meanings, and all are plausible, all are listed in the template text base.

- 17. (QUALIFY, MATTOCK, SORT OF)
- 18. (FIND, A:I, O:17)
- 19. (CONJ: AND, 16, 18) &/OR (TIME: UNTIL, 16, 18)  
&/OR (CAUS: AS A RESULT, 16, 18)

The third clause in the sentence has a predicate proposition (I went [from the shed]), a modified locative (down towards the stream), and another ambiguous connective with the previous clause (and). The source case of the predicate is inferred to be the shed, as there is no intervening action. The goal is not necessarily the stream here, as the narrator only mentions going in the direction of the stream.

- 20. (GO, A:I, >S:SHED, G:\$)
- 21. (LOC: TOWARD, 20, STREAM)
- 22. (QUALIFY, 21, DOWN)
- 23. (CONJ: AND, 19, 22) &/OR (TIME: AFTER, 22, 19)

The last clause is the explanation of the narrator's actions of the rest of the sentence. It consists of a quantifier (some worms), a reference proposition (worms = bait), two predicate propositions, one of which is inferred (I dig worms, >I obtain worms), and two connective propositions of purpose (obtaining worms to be used as bait, doing the previous actions to obtain worms). The purpose of the narrator's digging was to obtain worms for later use as bait. The worms obtained are the same worms to be used as bait. When the word "and" is used to connect one finite verb (e.g. try, come) to another, the result is an

indication of purpose. The phrase "try and dig" is logically equivalent to an infinitive of purpose, that is, equivalent to "in order to dig."

- 24. (NUMBER OF, WORMS, SOME)
- 25. (REFERENCE, 24, BAIT)
- >26. (OBTAIN, A:I, O:24)
- 27. (DIG, A:I, >I:MATTOCK)
- 28. (PURP: TO, 27, >26)
- 29. (PURP: TO, >26, 25)
- 30. (PURP: TO, 23, 29)

The fourth sentence is made up of two main clauses (stream was clear and shallow, stream did not look trouty), and a concession connecting them (but). The first clause consists of two qualifiers (clear stream, shallow stream) and a conjunction between them (and).

- 31. (QUALITY OF, STREAM, CLEAR)
- 32. (QUALITY OF, STREAM, SHALLOW)
- 33. (CONJ: AND, 31, 32)

The second clause in this sentence uses figurative language, and thus requires that a number of inferences be made before it can be understood. The problem is: what is a "trouty-looking" stream? It makes no sense in this context to compare the physical appearance of the stream and the trout. A seasoned trout fisherman, however, might note a particular set of qualities about streams in which trout prefer to live. These streams could be called "trouty," in that they possess trout. It can be inferred from this sentence that two of the qualities defining a "trouty" stream are clearness and shallowness. While the stream in question is clear and shallow, it is lacking some unspecified quality necessary to "trouty" streams. These inferences must be made to understand what is meant here by a "trouty-looking" stream.

- >34. (POSSESS, A:STREAM, O:TROUT)
- >35. (COND: IF, >34, 33)
- 36. (QUALITY OF, STREAM, "TROUTY")
- >37. (COND: IF, >34, 36)
- 38. (NEGATE, 36)

39. (CONC: BUT, 33, 36)

The fifth sentence begins with a highly ambiguous locative clause (On the grassy bank where it was damp). Three interpretations of this phrase are possible. First it is conceivable that there was dampness somewhere on the grassy bank, but that some other part of the bank was free of moisture. In this case, these propositions follow:

- 40. (QUALIFY, BANK, GRASSY)
- 41. (QUALIFY, AREA, DAMP)
- 42. (LOC: ON, 41, 40)

Conversely, a grassy bank may have been situated on a larger area of dampness.

- 40a. (QUALIFY, BANK, GRASSY)
- 41a. (QUALIFY, AREA, DAMP)
- 42a. (LOC: ON, 40a, 41a)

Finally, it may be that the grassy bank directly overlaps the damp area.

- 40b. (CONJ: AND, GRASSY, DAMP)
- 41b. (QUALITY OF, BANK, 40b)
- 42b.  $\emptyset$

All three of these interpretations are conceivable ones, and there is no way to disambiguate the phrase at the text base level. Thus, the three alternatives must all be included in a template text base. The remainder of the sentence consists of a qualifier (chunk of sod), two predicate propositions (I drove the mattock, I loosened a chunk), a locative (on the grassy bank), and an ambiguous connective.

- 43. (DRIVE, A:I, I:MATTOCK, G:EARTH)
- 44. (LOC: ON, 43, (42 or 42a or 41b))
- 45. (QUALIFY, CHUNK, SOD)
- 46. (LOOSEN, A:I, O:45)
- 47. (CONJ: AND, 44, 46) &/OR (TIME: AFTER, 46, 44) &/OR  
(PURP: TO, 44, 46)



The connective "and" affords three possible interpretations here: that the narrator both drove the mattock and loosened the sod, that he loosened the sod after driving the mattock, and that he drove the mattock in order to loosen the sod.

The sixth sentence is a simple locative statement.

48. (LOC: UNDERNEATH, WORMS, 46)

Four clauses and their connectives comprise the final sentence. The first two clauses are simple predicate propositions (They slid out of sight, I lifted the sod) connected by a temporal proposition (as).

49. (SLIDE, A:48, G:OUT OF SIGHT)  
 50. (LIFT, A:I, O:45)  
 51. (TIME: AS, 49, 50)

The third clause is a modified predicate proposition (I dug carefully). The fourth has a quantifier (a good many worms) and a predicate (I got worms).

52. (DIG, A:I, O:EARTH, I:MATTOCK)  
 33. (QUALIFY, 32, CAREFULLY)  
 54. (NUMBER OF, WORMS1, MANY)  
 55. (GET, A:I, O:54)  
 56. (CONJ: AND, 53, 55) &/OR (CAUS: CAUSE, 53, 55) &/OR  
 (PURP: TO, 53, 55) &/OR (TIME: AFTER, 55, 53)  
 57. (CONJ: AND, 51, 56) &/OR (TIME: AFTER, 56, 51)

The concept of "worms" being referred to in P54 is not equivalent to the concept expressed in P48 (LOC: UNDERNEATH, WORMS, 44). P54 specifies only those worms that the narrator actually caught, while P48 refers to the total population of worms that the narrator spotted beneath the sod. Hence, the argument must be subscripted in P54 to differentiate it from the similar concept of P48.

## 3.4.3 THE TEMPLATE TEXT BASE

The complete text base underlying this passage is presented below:

1. (UNBOLT, A:I, O:DOOR)
2. (GO, A:I, G:OUT)
3. (CONJ: AND, 1, 2) &/OR (TIME: AFTER, 2, 1)
4. (QUALITY OF, WEATHER, COOL)
5. (QUALIFY, MORNING, EARLY)
6. (TIME: IN, 4, 5)
7. (LOC: OUTSIDE, 6, \$)
8. (DIE DOWN, A:WIND)
9. (COME, A:DEW, >G:GROUND)
10. (TIME: WHEN, 9, 8)
11. (DRY, A:SUN, O:10)
12. (TIME: NOT YET, 2, 11)
13. (CONJ: AND, 7, 12) &/OR (CAUS: AS A RESULT, 12, 7)
14. (HUNT AROUND, A:I, G:\$)
15. (LOC: BEHIND, SHED, INN)
16. (LOC: IN, 14, 15)
17. (QUALIFY, MATTOCK, SORT OF)
18. (FIND, A:I, O:17)
19. (CONJ: AND, 16, 18) &/OR (TIME: DURING, 18, 16)  
&/OR (CAUS: AS A RESULT, 16, 18)
20. (GO, A:I, >S:SHED)
21. (LOC: TOWARD, 20, STREAM)
22. (QUALIFY, 21, DOWN)
23. (CONJ: AND, 19, 22) &/OR (TIME: AFTER, 22, 19)
24. (NUMBER OF, WORMS, SOME)
25. (REFERENCE, 24, BAIT)
- >26. (OBTAIN, A:I, O:24)
27. (DIG, A:I, >I:MATTOCK)
28. (PURP: TO, 27, >26)
29. (PURP: TO, >26, 25)
30. (PURP: TO, 23, 29)
31. (QUALITY OF, STREAM, CLEAR)
32. (QUALITY OF, STREAM, SHALLOW)
33. (CONJ: AND, 31, 32)
- >34. (POSSESS, A:STREAM, O:TROUT)
- >35. (COND: IF, >34, 33)
36. (QUALITY OF, STREAM, "TROUTY")
- >37. (COND: IF, >34, 36)
38. (NEGATE, 36)
39. (CONC: BUT, 33, 36)
40. (QUALIFY, BANK, GRASSY)
41. (QUALIFY, AREA, DAMP)
42. (LOC: ON, 41, 40)
  
- 40a. (QUALIFY, BANK, GRASSY)
- 41a. (QUALIFY, AREA, DAMP)
- 42a. (LOC: ON, 40a, 41a)
  
- 40b. (CONJ: AND, GRASSY, DAMP)
- 41b. (QUALITY OF, BANK, 40b)
- 42b. Ø

- 43. (DRIVE, A:I, I:MATTOCK, G:EARTH)
- 44. (LOC: ON, 43, (42 or 42a or 41b))
- 45. (QUALIFY, CHUNK, SOD)
- 46. (LOOSEN, A:I, O:45)
- 47. (CONJ: AND, 44, 46) &/OR (TIME: AFTER, 46, 44)  
&/OR (PURP: TO, 44, 46)
- 48. (LOC: UNDERNEATH, WORMS, 46)
- 49. (SLIDE, A:48, G:OUT OF SIGHT)
- 50. (LIFT, A:I, O:45)
- 51. (TIME: AS, 49, 50)
- 52. (DIG, A:I, I:MATTOCK, O:EARTH)
- 53. (QUALIFY, 52, CAREFULLY)
- 54. (NUMBER OF, WORMS1, MANY)
- 55. (GET, A:I, O:54)
- 56. (CONJ: AND, 53, 55) &/OR (CAUS: CAUSE, 53, 55)  
&/OR (PURP: TO, 53, 55) &/OR (TIME: AFTER, 55, 53)
- 57. (CONJ: AND, 51, 56) &/OR (TIME: AFTER, 56, 51)

#### 3.4.4 A NOTE ON LITERARY PASSAGES

This passage has been selected to demonstrate that material not contrived for use in an experimental situation can be readily subjected to analysis by propositionalization. It is true that literary devices (e.g. idioms, metaphors) and ambiguity in the text must be dealt with in order to arrive at a template text base representation of the passage in question. But this is not an impossible task. Idioms are expressed in terms of their semantic content; metaphors are usually decomposed into a number of presuppositions and inferences that a reader must make to understand the figurative language; and ambiguity is handled by the inclusion of multiple interpretations of an equivocal expression. It is unreasonable to suggest that any one reader will make all possible interpretations of the ambiguity. However, when there are a number of alternative meanings that might be feasible, all should be included in the text base.

#### 4 SCORING RECALL PROTOCOLS

The study of complex cognitive processes such as knowledge acquisition, text comprehension, and memory presents numerous problems to researchers in these areas. Not the least of these problems is finding a reliable and consistent method to determine the representation of meaning. The previous section has dealt with the construction of an idealized text base for any given text. This template text base contains all propositions explicitly represented in the surface structure of the text, as well as the inferences necessary to ensure a connected text base. It is conceded that this text base is sometimes a poor approximation to the original base structure. Yet it is all that can be determined without resorting to arbitrary or intuitive formulations.

Should this same approach be applied when scoring recall protocols of subjects? An affirmative answer to this question seems reasonable at first. After all, the original material and the subject's recall both constitute "texts." However, in practice, a major difference between the two exists. While the text is generally a thought out, coherent, and organized whole, the protocol of a subject is likely to be a disorganized list of ideas. While it is plausible to assume that the propositions of the speaker/writer are highly interconnected, the subject will not always integrate the ideas of the text, nor will the subject spend as much time thinking about the text as did the speaker/writer. Hence it is not wise to assume connections were made by the subject where none were given in the subject's protocol. Inter-scorer reliability would be threatened if propositions which are only implied by the subject's protocol are scored as given.

Another area of concern is the tremendous variety of expressions possible which mean the same--or nearly the same--thing. If asked to recall verbatim, subjects may not respond rather than give incorrect

wording. Yet when asked to recall in their own words, subjects can produce a large variety of expressions which have an ambiguous or questionable relationship to the original ideas.

Among the other considerations in scoring is the question of criterion. Where should one draw the line between acceptable and unacceptable representations of propositions? How lenient or strict should one be? The question of criterion must be decided by each individual researcher. An important qualification of this is that one must be consistent to whatever standards one reaches. In general, we have found that changes in the strictness of the criterion in scoring do not lead to changes in the overall pattern of results, other than to depress or elevate the level of recall slightly. Changes may be noted at the level of the individual proposition, however.

#### 4.1 PURPOSES

There are three basic measures possible when scoring by propositional analysis. Quantitative scoring is used if one wishes to determine the amount of recall. Propositions are scored by comparing the propositions of the protocol to those in the template text base. Scoring of propositions is often all-or-none. Scoring can also take into account the qualitative aspects of recall. The propositions can be scored as to the type of recall they represent, such as reproduction, inference elaboration, abstraction, error, etc. Lastly, recognition of individual propositions can be tested, using such measures as proportion of propositions recognized, reaction time, etc.

The propositional method of analysis is a useful aid in research using complex linguistic materials. Materials can be reduced to a list of connected ideas. The protocols can then be compared with the set of text propositions, thus simplifying the task of scoring.

#### 4.2 SCORING CRITERION--A PRIORI CONSIDERATIONS

The major consideration in scoring is consistency. One should choose a criterion and be consistent. It is not advisable to adapt too extreme a criterion. An important point to consider is that the words used in the template propositions are tokens representing abstract word concepts. These words may be expressed in the surface text as words or phrases. Many expressions may be chosen to represent a particular word concept. Some decisions must be made, therefore, as to the degree of surface variation that is acceptable.

Reproductive recall is the simplest form of recall to score. If one wishes to score for inferences, elaborations, errors, etc., it is necessary to define how these additional propositions are to be classified. A list of possible constructive propositions may be added to the list of text propositions before scoring. Of course, subjects are inventive, and will undoubtedly come up with many constructive propositions that are not on the list. These may be added to the list to ensure consistency in the scoring of these propositions.

It is possible to score for macrostructure propositions. However, rules for macrostructure formation from microstructure are tentative at this time. While scoring for macrostructure propositions could be (and is being) done, it will not be discussed here.

It is possible to score for partial recall of propositions. Optional cases which are specified in the text may be omitted in the recall. Some researchers may wish to give credit to subjects who produced partial recall of propositions. This can lead to problems in the reliability between scorers.

Researchers interested in text hierarchy may wish to give weights to propositions according to their importance in the text. An analysis of levels (see Kintsch, 1974) may be done in order to construct the text's

hierarchy. Weights may then be assigned to the different levels. Another approach is to score all propositions equally, then determine the proportion of propositions recalled at each level of the text.

#### 4.3 EXAMPLES OF SCORING TECHNIQUES--MERINO SHEEP

##### 4.3.1 SCORING FOR AMOUNT OF REPRODUCTIVE RECALL

It is generally the case that the emphasis in scoring is placed on reproduction of text propositions. What constitutes a reproduction is not always unequivocal. In scoring, an attempt is made to decide if the meaning in the protocol is a reproduction of the meaning from the text. Variations in expression are judged on a dichotomous scale with two values, "yes" or "no," though the actual meaning is, if not continuous, certainly multifarious. Hence the subtler nuances of meaning are lost.

When scoring is done as a yes/no decision, a cutoff must be set by the researcher. Below is presented an example illustrating three possible criteria. The Merino sheep text given above will be used as the to-be-remembered text. Three protocols have been selected to illustrate scoring for reproductive recall.

Criterion 1 is the strictest of the three. A protocol proposition will not be considered reproductive unless it is highly similar to the original text proposition. In cases where one proposition is embedded in another, the protocol must have the embedded proposition correctly given before being credited with the superimposed proposition. For example, if the template proposition were P12 (THRIVE, A:MERINO, I:(CONJ: AND, GRASSES, (QUALIFY, BUSHES, LOW))), a subject's protocol containing the statement "Merinos thrive on the vegetation." would not be given credit for P12, as the embedded propositions are not present. The only exception to this rule occurs when the embedded proposition is

a pronoun, or an abstraction or generalization of the original proposition in the surface expression of the text. See, for example, P6 of the Merino sheep text base. The surface form is "They are popular...", but the underlying meaning is that "Pure Merino sheep are popular...." A fault with this criterion is that it requires a high degree of similar wording before a protocol proposition is scored as a correct reproduction of the text proposition.

Criterion 2 is less stringent. Acceptable condensations of embedded propositions, such as "vegetation" for "grasses and low bushes" in the above example, are permitted to fill case positions of "correct" reproductions in place of the embedded proposition. However, generalizations or abstractions of relations are not. Using P12 again as an example, a subject's protocol containing the statement "Merinos eat grasses and low bushes." would not be credited as a "correct" reproduction of P12. While it can be inferred that Merinos do in fact eat the vegetation, the relation "thrive" implies much more. The meanings of "thrive" and "eat" are considered too far apart for a "correct" reproduction to be scored.

The last criterion is the most lenient one. Here "gist" is acceptable. P12 in the above example would be scored as a "correct" reproduction. The gist of the relationship between the sheep and the grasses and bushes is present, while the beneficial connotations of "thrive" are absent. (It is inferred that the thriving is produced by means of ingestion of the grasses and bushes, therefore "eat" is a generalization of the actual relation "thrive.") The looseness of this criterion can itself be a problem.



TEXT--MERINO SHEEP

Three-fourths of Australia's sheep are pure Merinos. They are popular because of the large amount of semiarid land. The Merino thrives on the grasses and low bushes which grow on semiarid plains. They are known for their heavy fleeces of fine-quality wool which bring a high price from textile manufacturers.

PROTOCOLS--MERINO SHEEP

## Protocol #1

The Merino is the kind of sheep commonly raised in Australia for wool. It constitutes 3/4 of the sheep there. They are popular because they thrive easily in this semiarid climate and like to eat the short bushes and grass. The Merino has a soft fleecy wool of high quality that brings good money in the world market.

## Protocol #2

Merinos are the finest breed of sheep in Australia. The reason they are of such high quality is because of the semiarid grass lands they graze on. The grass helps produce a heavy fleece which is important to the textile industry.

## Protocol #3

The Merino constitutes 3/4 of Australia's sheep population. It is in demand because its wool is very thick and heavy, and it is sent to manufacturers for clothing production. The Merino is raised in the area of Australia which receives 10-20'' of rain a year.\* It thrives on lowland bushes and shrubs.

\* By definition "semiarid" land receives 10-20 inches of rain a year.

SCORING PROTOCOLS

The propositions in each protocol are compared to the propositions of the template text base. It is not necessary to construct a formal text base for subject protocols in order to score them for reproductive recall. A loose analysis is sufficient for this purpose.

In the table below, the template propositions occupy the rows. The columns indicate the protocol number and the criterion used. An "X" indicates credit given for reproduction of a text proposition.

	PROTOCOLS								
	# 1			# 2			# 3		
	CRITERION	CRITERION	CRITERION	CRITERION	CRITERION	CRITERION	CRITERION	CRITERION	CRITERION
	1	2	3	1	2	3	1	2	3
1. (POSSESS, A:AUSTRALIA, O:SHEEP)			X			X		X	X
2. (NUMBER OF, 1, THREE-FOURTHS)		X	X				X	X	X
3. (QUALIFY, MERINO, PURE)									
4. (ISA, 2, 3)		X	X					X	X
5. (ISA, MERINOS, SHEEP)	X	X	X	X	X	X	X	X	X
6. (QUALITY OF, 3, POPULAR)	X	X	X						X
7. (QUALIFY, LAND, SEMIARID)		X	X	X	X	X			X
8. (EXTENT OF, 7, LARGE)									
9. (CAUS: BECAUSE OF, 8, 6)									
10. (QUALIFY, BUSHES, LOW)	X	X	X						
11. (CONJ: AND, GRASSES, 10)	X	X	X						
12. (QUALIFY, PLAINS, SEMIARID)									
13. (LOC: ON, 11, 12)			X			X			
14. (THRIVE, A:3, I:13)			X			X		X	X
15. (QUALITY OF, WOOL, FINE-QUALITY)	X	X	X						
16. (QUALITY OF, FLEECES, HEAVY)				X	X	X			
17. (QUALITY OF, 16, 15)									
18. (POSSESS, A:2, O:17)		X	X			X		X	X
19. (KNOW, A:\$, O:18)									
20. (QUALIFY, PRICE, HIGH)	X	X	X						
21. (MAKE, A:MANUFACTURERS, O:TEXTILES)						X	X		
22. (BRING, A:18, O:20, S:21)		X	X						
TOTALS	6	11	14	3	4	8	2	6	8

#### 4.3.2 SCORING FOR FORM OF RECALL

In the process of scoring for reproductive recall, much of the available information is disregarded. Information about inferences, elaborations, abstractions, generalizations, intrusions, errors, in addition to variations in the expression of the text propositions, is not considered. For some purposes, it is desirable to determine the form of recall as well as the amount. In these cases, scoring of recall should be a finer analysis than that provided by the yes/no scoring of reproductive recall.

Categories are set up to represent the various kinds of transformations or additions which are possible. These categories should reflect the scope and focus desired. For instance, one who studies inferences might wish to examine the inferences in the recall protocols in greater detail. Thus a number of categories would be set up, each indicating a

different type of inference. The rest of the categories could be reduced to the minimum, thus placing the emphasis on inferences.

Below is an example illustrating scoring for the quality of recall. The to-be-remembered text is "Merino Sheep." The passage to be analyzed is Protocol #1. Six categories are used: Correct, Generalized Argument, Generalized Relation, Inference, Elaboration, and Error.

A proposition is scored as Correct if the relation and arguments are equivalent to the relation and arguments of the corresponding text proposition. Should one of the arguments be a generalized representation of the original argument (such as "vegetation" for "grasses and low bushes" in the instrumental case of P12) it is scored as a Generalized Argument. Propositions in which the relation has been underspecified are scored as Generalized Relations. An Inference occurs when the proposition is implied by or implies text propositions. Propositions are Elaborations if they add fullness of detail which is probably correct, but which is not directly inferable from the text propositions. Errors are propositions which are incorrect or are intrusion errors.

ANALYZED PROTOCOL WITH PROPOSITIONS

## Protocol #1

The Merino is the kind of sheep commonly raised in Australia for wool. It constitutes 3/4 of the sheep there. They are popular because they thrive easily in this semiarid climate and like to eat the short bushes and grass. The Merino has a soft fleecy wool of high quality that brings good money in the world market.

PROTOCOL PROPOSITIONS	FORM OF RECALL**	TEMPLATE PROPOSITION NUMBER
1. (ISA, MERINO, SHEEP)	CORRECT	5
2. (RAISE, A:\$, O:1)	INFERENCE	
3. (QUALIFY, 2, COMMONLY)	ELABORATION	
4. (LOC: IN, 3, AUSTRALIA)	INFERENCE	
>5. (OBTAIN, A:\$, O:WOOL)	INFERENCE	
6. (PURP: FOR, 2, >5)	INFERENCE	
7. (LOC: IN, SHEEP, AUSTRALIA)	GENERALIZED RELATION	1
8. (NUMBER OF, 7, THREE-FOURTHS)	GENERALIZED ARGUMENT	2
9. (ISA, 8, 1)	GENERALIZED ARGUMENT	4
10. (QUALITY OF, 1, POPULAR)	CORRECT	6
11. (QUALIFY, CLIMATE, SEMIARID)	GENERALIZED ARGUMENT	7
12. (POSSESS, A:AUSTRALIA, O:11)	ELABORATION	
13. (THRIVE, A:1, I:12)	GENERALIZED ARGUMENT	14
14. (QUALIFY, 13, EASILY)	ELABORATION	
15. (QUALIFY, BUSHES, SHORT)	CORRECT	10
16. (CONJ: AND, 15, GRASSES)	CORRECT	11
17. (EAT, A:1, O:16)	GENERALIZED RELATION	14
18. (LIKE, A:1, O:17)	ELABORATION	
19. (CONJ: AND, 13, 17)	GENERALIZED RELATION	13
20. (CAUS: BECAUSE, 19, 10)	INFERENCE	
21. (QUALITY OF, WOOL, FLEECY)	ERROR	
22. (QUALITY OF, 21, SOFT)	ELABORATION	
23. (QUALITY OF, 22, HIGH-QUALITY)	CORRECT	15
24. (QUALIFY, PRICE, HIGH)	CORRECT	20
25. (QUALIFY, MARKET, WORLD)	ELABORATION	
26. (BRING, A:23, O:24, S:26)	GENERALIZED ARGUMENT	22
27. (POSSESS, A:1, O:26)	GENERALIZED ARGUMENT	18

\*\*Template propositions P3, P8, P9, P12, P16, P17, P19, and P21 are not represented in the protocol.

SUMMARY	FORM OF RECALL	TOTAL
	CORRECT	6
	GENERALIZED ARGUMENT	6
	GENERALIZED RELATION	3
	INFERENCE	5
	ELABORATION	6
	ERROR	1
	NULL	8

#### 4.3.3 SCORING FOR RECOGNITION

The use of a propositional analysis in a recognition study differs from its use in a recall study. The propositions are used in the former to construct testing materials which assert a single proposition, or a single new proposition. In the latter, they are used as a template for the scoring of the recall protocols.

The measures in a recognition study usually are proportion recognized, confidence ratings, reaction times, and the like. One can also give subjects a verification task, asking them to correct false statements to read as true ones. This allows the experimenter to obtain more information from a given subject and determine the person's knowledge of a given proposition in a multi-proposition sentence. For example, to test knowledge of a connective propositions, such as the causal proposition P9 in "Merino Sheep," one might construct a test sentence which reads

The popularity of Merino sheep causes the large amount of semi-arid land in Australia.

Thus a subject who recognizes this as a reversal of the actual ideas expressed in the text, and is able to show this by changing the sentence into a true sentence, yields more information about the state of his or her knowledge of what was read.

## 5 CONCLUSION

The present work is an outline of one approach to the representation of the meaning in texts. It is not intended to be a prescription. Variations are certainly possible. Different purposes may call for the use of alternative representations. Some researchers may not desire the detail given in this system. Others might wish to include pragmatic or syntactic information. Still others may wish to include macrostructure propositions in their analyses. Such decisions are up to the individual researcher.

This system has worked well as a tool in experiments which use texts as material. Providing a more explicit formulation of it should prove useful to other researchers who might wish to use propositional analysis in their work.

## REFERENCES

- van Dijk, T. Text and Context. Explorations in the Semantics and Pragmatics of Discourse. London: Longmans, 1976.
- Fillmore, C.J. The case for case. In E. Bach and R.T. Harms (Eds.), Universals in Linguistic Theory. New York: Holt, Rinehart and Winston, 1968.
- Fillmore, C.J. Types of lexical information. In D.D. Steinberg and L.A. Jakobovits (Eds.), Semantics. Cambridge: University Press, 1969.
- Fillmore, C.J. Some problems for case grammar. In C.J. Fillmore (Ed.), Working Papers in Linguistics, No. 10, Dept. of Linguistics, The Ohio State University, 1971.
- Hemingway, E. The Sun Also Rises. New York, Scribner, 1926.
- Kintsch, W. Learning, Memory and Conceptual Processes. New York: Wiley, 1970.
- Kintsch, W. Notes on the structure of semantic memory. In E. Tulving and W. Donaldson (Eds.), Organization of Memory. New York: Academic Press, 1972.
- Kintsch, W. The Representation of Meaning in Memory. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1974.
- Kintsch, W. and van Dijk, T. Comment on se rappelle et on résume des histoires. Langages, 1975, 40, 98-116.
- Lakoff, G. Hedges: A study in measuring criteria and the logic of fuzzy concepts. Papers from the Eighth Regional Meeting, Chicago Linguistics Society. Chicago: University of Chicago Linguistics Dept., 1972.