String Program Reports No. 15

COMPUTER-BASED INVESTIGATION INTO THE STRUCTURE OF INFORMATION

Naomi Sager, Daniel Gordon, Elaine Marsh,
Lynette Hirschman, Carol Friedman, Michiko Kosaka,
Catherine Christenson, and Carolyn White

New York University

Final Report on Research Grant #IST79-20788
Division of Information Science and Technology
National Science Foundation

NYU Linguistic String Project
251 Mercer Street
New York, New York 10012

December 1982

Abstract

In this study, quantitative measures of the information content of textual material have been developed based upon analysis of the linguistic structure of the sentences in the text. It has been possible to measure such properties as:

1) The amount of information contributed by a sentence to the discourse.

2) The complexity of the information within the sentence, including the overall logical structure and the contributions of local modifiers.

3) The density of information based on the ratio of the number of words in a sentence to the number of information-contributing operators.

The methods used were computer-based, ensuring that the rules for linguistic analysis and measurement of information content were carried out uniformly.

Two contrasting types of texts were used to develop the measures. The measures were then applied to contrasting sentences within one type of text.

The results show that quantitative measures of properties of textual information can be developed which accord with intuitively perceived differences in the informational complexity of the material.

COMPUTER-BASED INVESTIGATION INTO THE STRUCTURE OF INFORMATION

<u>Appendices</u>

The Measurement of Properties of Textual Information

Daniel Gordon and Naomi Sager

Introduction

The problem of measuring information is not new. Yet the impressive results in this domain (e.g. information theory) have dealt with statistical properties of messages as they bear on the capacity of channels to transmit messages rather than with the content of the messages themselves. Thus, Shannon writes in his Introduction to The Mathematical Theory of Communication [1]:

> "The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point. Frequently the messages have *meaning*; that is, they refer to or are correlated according to some system with certain physical or conceptual entities. These semantic aspects of communication are irrelevant to the engineering problem. The significant aspect is that the actual message is one *selected from a set* of possible messages." (Italics in original.)

The approach taken in this study is somewhat different. Rather than investigate the degradation in the amount of information contained in a message, we have endeavoured to develop quantitative measures of the amount of information contained in a message. Such measures would make it possible

1) to compare the amount of information in two different texts

2) to characterize a text by its quantity of information

3) to compute and compare the densities and complexity of information in various texts.

Shannon is correct that such an approach cannot be taken without reference to "...some system with certain physical or conceptual entities." Text must be interpreted in the light of a theory which accounts for <u>how</u> the text carries information.

We are able to approach these questions because of advances over the past 20 years or more in the description and computation of language structure. Measuring structural properties of textual information depends on the ability to separate the information-bearing portions (or processes) in sentences from all other linguistic operations that enter into the final form.

The first step toward such an analysis was the discovery of linguistic transformations by Z. Harris [2] with the result that every sentence could be analyzed as composed of elementary sentences and operations on them, and the further result that these transformations were themselves composed of elementary operators that on the whole were either incremental (information-bearing) or paraphrastic [3]. The isolation of informational processes in language was formalized in a theory of grammar in which the information-contributing operations in building a sentence are a separate component of the grammar and produce a predicational structure for sentences that corresponds, recognizably, to the information carried by the sentence [4]. This structure can be represented by a parenthesized structure or by an operator-argument tree, in which each higher operator is interpreted as a predication on the argument(s) it immediately dominates. For example, from [4] (as reprinted, p. 674]:

relate to (I, grow (tree), increase (rain))

= I relate the tree's growth to the rain's increase.

Here, the incremental (informational) operator relate to has the three arguments I, grow, increase; the arguments grow and increase are themselves informational operators with the respective arguments tree and rain.

The remaining portion of the operator grammar (the non-information-contributing part) is found to consist of "reductions," i.e. shortenings of the full information form by such devices as substituting a distinguished word (pronoun), dropping repeated or indefinite reconstructable words (zeroing), shortening words to become affixes, etc. [5,6]. These operations, in various combinations, when applied to the information form of the sentence, result in the sentence as we see it. For example, if the above example sentence had as subject the indefinite someone in place of I, then the permitted reduction of indefinites to zero, in combination with certain other operations, would provide the passive form: The tree's growth is related to the rain's increase (or: to the increase in rain), which contains no additional substantive information over the purely informational form.

Harris shows in [6] that this type of analysis can be applied to the whole language, laying bare, in the composition of each sentence, the contribution of each informational operator and its relation to the other informational operators (and primitive arguments) in the sentence. This result provides us, finally, with units of information related to the structure of the information carried by the sentence, and provides a basis for quantitative measures of properties of information that are related to its structure. At the simplest, one would expect that sentences carrying, intuitively, more information than others, will contain more operators of the informational type. Other informational properties should also be measurable based on the number, types and configurations of the informational operators.

During this same period that a theory of linguistic information was developing, the techniques for computerized natural language processing were also increasing in extent and power [7]. Comprehensive bibliographies are to be found in several recent books, where the work is viewed in the context of information retrieval [8], artificial intelligence and cognitive science [9], semantic representation in limited domains [10], text processing and informational representation [11], discourse processes [12], and speech understanding [13], to name a few.

Regardless of the ultimate purpose, workers have found that to obtain useful representations of language material and to process it with some degree of generality, regularities within the language data have to be formulated and made part of the processing procedures. In our own case (the LSP system of NYU) and in several other systems (e.g. the IBM REQUEST/TQA system [14,15,16]), a first stage of parsing that captures "surface" grammatical regularities is followed by a stage of transformational analysis that produces a yet more regular, underlying, or "deep," structure of the sentence [17].

The use of linguistic transformations to regularize parse tree structures in both the systems cited above is in the service of reaching a final representation suited to a particular application. TQA is a natural language question-answering system, serving as a front-end to numerical databases. In the LSP case, the purpose is to obtain a database of structured information from free-text input [18], and, in related research, to develop a natural language query system for such a database [19]. For this purpose, it has not been necessary in the material processed to date to carry out a full operator-argument type decomposition in the transformational stage of processing. Still, many of the operations are the same or similar to those that would result in an operator-argument type tree, so that, with adjustments, we have been able to use the LSP transformational implementation to obtain the structures that figure in the present investigation.

Methods

1. Measures based on operator counts.

In this study, we have based our measures on the operator-argument type of analysis, although we have not used the full operator grammar of [6], e.g. affixes that contribute a particular semantic effect have not been treated as separate operators in most cases. What we have tried to do, however, is to apply the principles of operator-argument analysis and to adjust our results for the different types of texts so that comparisons are valid. As noted in the Introduction, we decided to work with the operator-argument representation of the texts to obtain quantitative results because such a representation is both uniform and informationally revealing.

In particular, the number and distribution of the operators in the operator-argument tree was felt to be a measure of the information contained in a sentence. Moreover, the conversion of a text into its operator-argument representation depends mainly on general properties of natural language rather than domain-specific information. If texts similar in complexity should turn out to have similar measures based on an operator-argument analysis, then a quite general method would exist to compare measures derived from the operator-argument representation of sentences derived from different materials.

We decided to measure the following about each sentence and its operator-argument tree:

1) The number of words in each sentence.

2) The number of operators in each sentence.

3) The ratio of the number of words to the number of operators (the inverse of the density operator).

4a) The maximum depth of nesting of each operator-argument tree, including that of local modifiers.

4b) The maximum depth not including that of local modifiers.

5) The product of 2) and 4), as a measure of overall complexity.

Our motivation for these measures was straightforward. (2), the number of operators in a sentence, is an index of the amount of predication contained in a sentence. Since the predications installed by a sentence are at least one sense of the "meaning" of a sentence, we reasoned that (2) would be one measure of amount of information carried by a sentence. The inverse of (3) is at least a rough measure of the density of information per sentence. We felt that (4) would measure one feature of informational complexity, viz., the extent to which the predications of the sentence were predicated "on top of" each other. A sentence could have a large number of operators without those operators being deeply nested; measuring (4) would give an independent estimate of the complexity of predication within a sentence. The product of (2) and (4), (5) would be a more satisfactory measure of complexity because it would "reward" a sentence for having either high operator content or deep nesting, and would serve as a composite index of the relative effects of the two.

2. Measures developed for contrasting types of text.

One obvious problem in developing measures of such properties of information as amount and complexity is the need for an independent source of judgment as to whether the measurements do indeed capture our intuitive (human-perceived) evaluation of these properties. Suppose the results of our calculations of complexity were to show that sentence S1 is twice as complex as S2. Who would judge this result, and on what basis say "yes, S1 is twice as complex as S2" or "no, S1 is only 1 1/2 times as complex as S2"?

To avoid the large task of setting up scaled subjective judgments of human observers, we divided our problem into two stages. First we would use highly contrastive types of text, say T1 and T2, for which there was general agreement because of the type of text, that T2 was carrying more, and more complex, information than T1, and we would develop measures whose numbers clearly reflected this contrast. Then, having gained a certain amount of confidence in the measures, we would examine sentences within the same type of text, whose values of the measure indicated they should have differing properties, and see if our human reaction as readers discerned these differences. A controlled experiment with human subjects would have been the ideal situation, but

*

in the absence of this possibility, the techniques we used had gratifyingly consistent results that confirmed our measures.

The contrasting types of texts used in the investigation were:

T1 Factual reports -- documents containing the narrative of the events in one particular situation.

T2 Theoretical texts -- portions of published articles describing mechanisms of action of complex processes.

For the T1 material, we chose texts extracted from hospital discharge summaries. Fig. 1.1 shows an example of this type of text. Typically, the sentences in this material are short and declarative, and their concatenation into discourse represents only an accumulation of facts.

For the T2 material, we chose two articles reviewing advances in the theory of lipid metabolism. In sharp contrast to the discharge summary material, these sentences are individually complex, and the discourse as a whole is marked by the reference in later sentences to earlier sentences. An example text is shown in Fig. 1.2.

The conversion of the discharge summary material to an operator-argument form was performed by computer. For this we made use of the LSP natural-language processing system. Although the system as a whole is designed to convert natural-language texts into database

BELLEVUE

DISCHARGE SUMMARY

Ped. Ser.

HIS:
This is the 2nd BHA adm. for this 13 and 10/12 yr. old Hispanic male who presented with a history of one day of headache, vomiting, fever, anorexia and the onset of a stiff neck on the morning of adm.

The patient was in his usual state of good health until the morning prior to adm. when he developed a severe bifrontal headache with no diplopia, photophobia, tinnitus or seizure activity. Shortly after the onset of the headache, vomiting ensued with 8 episodes by history. No blood or coffeeground material or bilious material was noted. The patient also had a low grade fever on the morning of adm. He noted the onset of a stiff neck and a rash that day. He was rushed to the PES, seen and immediately admitted to rule out meningitis.

The patient apparently had a URI last week which resolved. On transfer to the floor, he was lethargic and complaining of a very severe stiff neck.

PAST MED. HIS:
The patient was previously admitted to BH Center secondary to an auto accident in which he sustained a blowout fracture of the right orbit and a fracture of the floor of the left orbit, as well as a fractured zygoma. This required extensive repair. He has had residual ptosis and blurred vision.

In Aug. of 1975 the patient had headaches, vomiting, facial rash, enuresis and again a stiff neck while in Florida. He was admitted to ORANGE County Hospital in Orlando, Florida and diagnosed by spinal tap to have a meningitis; etiology not quite clear. The patient was treated with 10 days of IV antibiotics.

ALLERGIES:
None known.

IMMUNIZATIONS:
Complete.

ACUTE INFECTION DISEASE:
The patient has had the chickenpox, measles and the mumps.

ROS:
Noncontributory.

Figure 1.1: Sample of FACT Textual Material.

Figure 1.2: Sample of Theoretical Textual Material.

Regulation of Plasma Cholesterol by Lipoprotein Receptors

Michael S. Brown, Petri T. Kovanen, Joseph L. Goldstein

Summary. The lipoprotein transport system holds the key to understanding the mechanisms by which genes, diet, and hormones interact to regulate the plasma cholesterol level in man. Crucial components of this system are lipoprotein receptors in the liver and extrahepatic tissues that mediate the uptake and degradation of cholesterol-carrying lipoproteins. The number of lipoprotein receptors, and hence the efficiency of disposal of plasma cholesterol, can be increased by cholesterol-lowering drugs. Regulation of lipoprotein receptors can be exploited pharmacologically in the therapy of hypercholesterolemia and atherosclerosis in man.

Recent advances in the genetics and cellular biology of cholesterol metabolism have provided new insights into the control of plasma cholesterol levels in man. It is now apparent that normal humans possess efficient mechanisms for the removal of cholesterol from plasma (*1*). This disposal process depends on receptors located on the surface of cells in the liver and extrahepatic tissues. The receptors bind circulating lipoproteins that transport cholesterol in the bloodstream, thereby initiating a process by which the lipoproteins are taken up and degraded by cells, yielding their cholesterol for cellular uses (*2*).

The lipoprotein receptors are components of an integrated transport system that shuttles cholesterol continuously among intestine, liver, and extrahepatic tissues (*1*). An interesting feature of the system is that the lipoproteins are degraded as they deliver their cholesterol to tissues, while the cholesterol survives, eventually to be excreted from the tissues bound to new lipoprotein carriers. Exit of cholesterol from the body occurs only when the sterol is transported to the liver for excretion into the bile.

Because of the continuous cycling of cholesterol into and out of the bloodstream, the plasma cholesterol concentration is not a simple additive function of dietary cholesterol intake and endogenous cholesterol synthesis. Rather, it reflects the rates of synthesis of the cholesterol-carrying lipoproteins and the efficiency of the receptor mechanisms that determine their catabolism. Recent experiments show that the number of lipoprotein receptors in the liver, and hence the rate of removal of cholesterol from plasma, is under regulation and that the number of receptors can be increased by certain cholesterol-lowering drugs (*3*, *4*). This new information eventually may reveal why different individuals respond differently to dietary cholesterol.

In this article we review the general features of the lipoprotein transport system worked out in studies from many laboratories over the past 25 years (*5*). We then focus on the lipoprotein receptors of liver and extrahepatic tissues and their role in regulating plasma cholesterol levels. Finally, we raise the possibility that pharmacologic manipulation of the hepatic and extrahepatic receptors may have therapeutic importance in the treatment of hypercholesterolemia and atherosclerosis in man.

The Lipoprotein Transport System

The lipoprotein transport system carries two classes of hydrophobic lipids, triglycerides (esters of glycerol and long-chain fatty acids) and cholesteryl esters (esters of cholesterol and long-chain fatty acids). Before they can be used by cells, the triglycerides and cholesteryl esters must be hydrolyzed to liberate fatty acids and unesterified cholesterol, respectively (*1*). Triglycerides are delivered primarily to adipose tissue and muscle where the fatty acids are stored or oxidized for energy. Cholesteryl esters are transported to all body cells where the unesterified sterol is used as a structural component of plasma membranes. These esters also supply cholesterol for synthesis of steroid hormones and bile acids.

For transport in plasma, triglycerides and cholesteryl esters are packaged into lipoprotein particles in which they form a hydrophobic core surrounded by a surface monolayer of polar phospholipids. The surface coat also contains unesterified cholesterol in relatively small amounts together with proteins called apoproteins (*5*). Through interactions with enzymes and cell surface receptors, the apoproteins direct each lipoprotein to its site of metabolism.

The lipoprotein transport pathway can be divided conceptually into exogenous and endogenous systems that transport lipids of dietary and hepatic origin, respectively (Fig. 1). Both systems begin with the secretion of triglyceride-rich lipoproteins—intestinal chylomicrons in the exogenous system and hepatic very-low-density lipoproteins (VLDL) in the endogenous system. Each of these particles contains an apoprotein called apoB, which maintains its structural integrity and which remains with the particle throughout its interconversions in the plasma. Recent studies indicate that the apoB of chylomicrons is not identical to that of VLDL (*6*).

Exogenous lipid transport. A typical American adult absorbs about 100 grams of triglyceride and 250 milligrams of cholesterol from the diet daily (Fig. 1). The intestine incorporates these lipids into chylomicrons, huge lipoproteins (diameter, 800 to 5000 angstroms) that are secreted into the lymph and from there enter the bloodstream. Inasmuch as chylomicrons are too large to cross the endothelial barrier, they must be metabolized while still in the bloodstream. For this purpose the chylomicrons bind to lipoprotein lipase, an enzyme (E.C. 3.1.1.34) that is fixed to the luminal surface of the endothelial cells that line capillaries of adipose and muscle tissues (Fig. 2A). The chylomicrons contain an apoprotein (C-II) that activates the lipase, which liberates free fatty acids and monoglycerides. The fatty acids enter the adjacent muscle or adipose cells where they are either oxidized or reesterified for storage (*1*). As the triglyceride core is depleted, the chylomicron

Drs. Brown and Goldstein are the Paul J. Thomas Professors in the Departments of Molecular Genetics and Internal Medicine at the University of Texas Health Science Center, 5323 Harry Hines Boulevard, Dallas 75235. Dr. Kovanen, now at the University of Helsinki, was formerly a Visiting Scientist at the University of Texas Health Science Center.

 0036-8075/81/0508-0628$02.00/0

entries, the first two components--the English parse and the English transformational decomposition--are perfectly general and were adaptable to our task.

The English parse [11,20] produces a structure for each sentence representing the "surface" grammatical features of the sentence. The output of the parser for one sentence from the discharge summary material is shown in Fig. 1.3

The transformational component operates on such a parse tree. Essentially, its effect is to undo the paraphrastic (non-information-bearing) reductions from the surface text, leaving a tree with only the information-bearing relations represented in the structure. The "decomposition" tree for our example sentence is shown in Fig. 1.4.

From the standpoint of the laws governing the "packing" of information into natural-language text, the details of these transformations, their mode of operation and their integration into a consistent sequence of processing, is a subject of some interest. The accompanying article, The Implementation of English Transformations, covers this in detail.

Although the detail of the decomposition tree of Fig. 1.4 is important for further computer processing in the full LSP system, what we needed for our measurements was a "skeleton" of this tree, where only the important operators, arguments, and their relations of

* HIDSM 1. 1.13

* THERE WAS NO HX OF STIFF NECK , ABNORMAL

* POSTURING , DIARRHEA OR OTHER FOCAL SIGNS

* COMPATIBLE WITH MENINGITIS .

```
((SENTENCE (TEXTLET (OLD-SENTENCE (INTRODUCER (NULL)) (CENTER (ASSERTION
NULL)) (SA (NULL)) (VERB (LV (NULL)) (VVAR (TV = 'WAS' % (PAST SINGULAR
VBE OBJLIST [(OBJECTBE)]) ^ ((('ARE') TV [(PLURAL VBE OBJLIST [(OBJECTBE
)])]) (('BE') V [(VBE OBJLIST [(OBJECTBE)])]) (('BEEN') VEN [(VBE
OBJLIST [(OBJECTBE)])]) (('BEING') VING [(VBE OBJLIST [(OBJECTBE)])]) ((
'IS') TV [(SINGULAR VBE OBJLIST [(OBJECTBE)])]) (('WAS') TV [(PAST
SINGULAR VBE OBJLIST [(OBJECTBE)])]) (('WERE') TV [(PAST PLURAL VBE
OBJLIST [(OBJECTBE)])]) (('AM') TV [(SINGULAR VBE NOTNSUBJ [(NONHUMAN
NSENT1 NSENT2 NSENT3 NTIME1 NTIME2)] OBJLIST [(OBJECTBE)])])))) (RV (
NULL))) (SA (NULL)) (OBJECT (OBJECTBE (OBJBE (NSTG (LNR (LN (TPOS (LTR (
LT (NULL)) (T = 'NO' % (NEGATIVE H-NEG) ¨ ((('NO') T [(NEGATIVE H-NEG)])
)))) (QPOS (NULL)) (APOS (NULL)) (NSPOS (NULL)) (NPOS (NULL))) (NVAR (N
= 'HX' % (SINGULAR NONHUMAN H-REP H-PER PAST H-RECORD) ¨ ((('HISTORY') N
 [(SINGULAR NONHUMAN H-REP H-PER PAST H-RECORD)]) (('HIS.') N [(SINGULAR
 NONHUMAN H-REP H-PER PAST H-RECORD)]) (('HX') N [(SINGULAR NONHUMAN
H-REP H-PER PAST H-RECORD)]) (('HISTORIES') N [(PLURAL NONHUMAN H-REP
H-PER PAST H-RECORD)])))) (RN (RNP (PN (LP (NULL)) (P = 'OF' % ('OF') ¨
((('OF') P [('OF')]))) (NSTGO (NSTG (LNR (LN (TPOS (LTR (LT (NULL)) (
NULL))) (QPOS (NULL)) (APOS (ADJADJ (LAR1 (LA (NULL)) (AVAR (LCDA (NULL)
) (ADJ = 'STIFF' % (H-INDIC) ^ ((('STIFF') ADJ [(H-INDIC)])))) (RA1 (
NULL))))) (NSPOS (NULL)) (NPOS (NULL))) (NVAR (N = 'NECK' % (SINGULAR
NCOUNT1 NONHUMAN H-BODYPART) ¨ ((('NECK') N [(SINGULAR NCOUNT1 NONHUMAN
H-BODYPART)]) (('NECKS') N [(PLURAL NCOUNT1 NONHUMAN H-BODYPART)])))) (
COMMASTG (',' = ',' ¨ NIL) (NOTOPT (NULL)) (Q-CONJ (LN (TPOS (LTR (LT (
NULL)) (NULL))) (QPOS (NULL)) (APOS (ADJADJ (LAR1 (LA (NULL)) (AVAR (
LCDA (NULL)) (ADJ = 'ABNORMAL' % (ASENT1 [(AFORTO)] H-INDIC) ¨ (((
'ABNORMAL') ADJ [(ASENT1 [(AFORTO)] H-INDIC)])))) (RA1 (NULL))))) (NSPOS
 (NULL)) (NPOS (NULL))) (NVAR (VING = 'POSTURING' % (NOTNSUBJ [(NONHUMAN
 NSENT1 NSENT2 NSENT3 NTIME1 NTIME2)] NOTNOBJ [(NSENT1 NSENT2 NSENT3
NTIME1 NTIME2)] OBJLIST [(NULLOBJ)] H-BODYFUNC) ¨ ((('POSTURE') N [(
SINGULAR NCOUNT1 NONHUMAN H-BODYFUNC)] V [(NOTNSUBJ [(NONHUMAN NSENT1
NSENT2 NSENT3 NTIME1 NTIME2)] NOTNOBJ [(NSENT1 NSENT2 NSENT3 NTIME1
NTIME2)] OBJLIST [(NULLOBJ)] H-BODYFUNC)] TV [(PLURAL NOTNSUBJ [(
NONHUMAN NSENT1 NSENT2 NSENT3 NTIME1 NTIME2)] NOTNOBJ [(NSENT1 NSENT2
NSENT3 NTIME1 NTIME2)] OBJLIST [(NULLOBJ)] H-BODYFUNC)]) (('POSTURED')
TV [(PAST NOTNSUBJ [(NONHUMAN NSENT1 NSENT2 NSENT3 NTIME1 NTIME2)]
NOTNOBJ [(NSENT1 NSENT2 NSENT3 NTIME1 NTIME2)] OBJLIST [(NULLOBJ)]
H-BODYFUNC)] VEN [(NOTNSUBJ [(NONHUMAN NSENT1 NSENT2 NSENT3 NTIME1
NTIME2)] NOTNOBJ [(NSENT1 NSENT2 NSENT3 NTIME1 NTIME2)] OBJLIST [(
NULLOBJ)] H-BODYFUNC)]) (('POSTURES') N [(PLURAL NCOUNT1 NONHUMAN
H-BODYFUNC)] TV [(SINGULAR NOTNSUBJ [(NONHUMAN NSENT1 NSENT2 NSENT3
NTIME1 NTIME2)] NOTNOBJ [(NSENT1 NSENT2 NSENT3 NTIME1 NTIME2)] OBJLIST [
(NULLOBJ)] H-BODYFUNC)]) (('POSTURING') VING [(NOTNSUBJ [(NONHUMAN
```

Figure 1.3: Parse Tree of HIDSM 1.1.13 in Parenthesized Form.

NSENT1 NSENT2 NSENT3 NTIME1 NTIME2)] NOTNOBJ [(NSENT1 NSENT2 NSENT3
NTIME1 NTIME2)] OBJLIST [(NULLOBJ)] H-BODYFUNC)])))) (COMMASTG (',' =
',' ^ NIL) (NOTOPT (NULL)) (Q-CONJ (LN (TPOS (LTR (LT (NULL)) (NULL))) (
QPOS (NULL)) (APOS (NULL)) (NSPOS (NULL)) (NPOS (NULL))) (NVAR (N =
'DIARRHEA' % (SINGULAR NONHUMAN H-INDIC) ¨ ((('DIARRHEA') N [(SINGULAR
NONHUMAN H-INDIC)])))) (ORSTG ('OR' = 'OR' ¨ NIL) (SACONJ (SA (NULL))) (
Q-CONJ (LN (TPOS (LTR (LT (NULL)) (NULL))) (QPOS (NULL)) (APOS (ADJADJ (
LAR1 (LA (NULL)) (AVAR (LCDA (NULL)) (ADJ = 'OTHER' % (APREQ) ¨ (((
'OTHER') ADJ [(APREQ)])))) (RA1 (NULL))))) (NSPOS (NULL)) (NPOS (NULL)))
(NVAR (N = 'FOCAL' 'SIGNS' % (PLURAL NONHUMAN NCOUNT1 H-INDIC) ¨ (((
'FOCAL' 'SIGN') N [(SINGULAR NONHUMAN NCOUNT1 H-INDIC)]) (('FOCAL'
'SIGNS') N [(PLURAL NONHUMAN NCOUNT1 H-INDIC)]))))))))))) (RN (ADJINRN (
LAR (LA (NULL)) (AVAR (LCDA (NULL)) (ADJ = 'COMPATIBLE' % (ASENT1 [(
ATHAT ASHOULD AFORTO)] H-CONN PVAL [('WITH')]) ¨ ((('COMPATIBLE') ADJ [(
ASENT1 [(ATHAT ASHOULD AFORTO)] H-CONN PVAL [('WITH')]]))))) (RA (PN (LP
(NULL)) (P = 'WITH' % ('WITH' VHAVE INSTR CONJ-LIKE) ¨ ((('WITH') P [(
'WITH' VHAVE INSTR CONJ-LIKE)] CS4 CS5 CS6 CS7))) (NSTGO (NSTG (LNR (LN
(TPOS (LTR (LT (NULL)) (NULL))) (QPOS (NULL)) (APOS (NULL)) (NSPOS (NULL
)) (NPOS (NULL))) (NVAR (N = 'MENINGITIS' % (SINGULAR NONHUMAN H-DIAG) ¨
((('MENINGITIS') N [(SINGULAR NONHUMAN H-DIAG)])))) (RN (NULL))))))))
(NULL)))))) (NULL)))))) (RV (NULL)) (SA (NULL)))) (ENDMARK ('.' = '.'
¨ NIL))) (MORESENT (NULL)))) [((81 [COMPUTED-ATT] H-INDIC [H-BODYPART])
(81 [N-TO-LN-ATT] . 67) (81 [SELECT-ATT] H-BODYPART) (73 [SELECT-ATT]
H-INDIC) (111 [COMPUTED-ATT] H-RESULT [H-BODYFUNC]) (111 [N-TO-LN-ATT]
. 97) (111 [SELECT-ATT] H-BODYFUNC) (103 [SELECT-ATT] H-INDIC) (80 [
POSTCONJELEM] . 110) (110 [PRECONJELEM] . 80) (111 [COMPUTED-ATT]
H-RESULT [H-BODYFUNC]) (111 [SELECT-ATT] H-BODYFUNC) (81 [COMPUTED-ATT]
H-INDIC [H-BODYPART]) (81 [SELECT-ATT] H-BODYPART) (111 [MATCHED] . 81)
(110 [POSTCONJELEM] . 131) (131 [PRECONJELEM] . 110) (87 [
NOT-DISTR-LN-ATT] . 131) (132 [SELECT-ATT] H-INDIC) (111 [COMPUTED-ATT]
H-RESULT [H-BODYFUNC]) (111 [SELECT-ATT] H-BODYFUNC) (132 [MATCHED] .
111) (131 [POSTCONJELEM] . 162) (162 [PRECONJELEM] . 131) (163 [
SELECT-ATT] H-INDIC) (132 [SELECT-ATT] H-INDIC) (163 [MATCHED] . 132) (
87 [POSTCONJELEM] . 139) (139 [PRECONJELEM] . 87) (57 [POSTCONJELEM] .
87) (87 [PRECONJELEM] . 57) (177 [PVAL-ATT] . 172) (172 [SELECT-ATT]
H-CONN) (196 [SELECT-ATT] H-DIAG) (81 [SELECT-ATT] H-BODYPART) (81 [
COMPUTED-ATT] H-INDIC [H-BODYPART]) (172 [SELECT-ATT] H-CONN) (196 [
SELECT-ATT] H-DIAG) (163 [SELECT-ATT] H-INDIC) (172 [SELECT-ATT] H-CONN)
(196 [SELECT-ATT] H-DIAG) (132 [SELECT-ATT] H-INDIC) (172 [SELECT-ATT]
H-CONN) (196 [SELECT-ATT] H-DIAG) (111 [SELECT-ATT] H-BODYFUNC) (111 [
COMPUTED-ATT] H-RESULT [H-BODYFUNC]) (47 [COMPUTED-ATT] H-INDIC [H-PER])
(47 [N-TO-RN-ATT] . 56) (47 [SELECT-ATT] H-REP H-PER) (81 [SELECT-ATT]
H-BODYPART) (81 [COMPUTED-ATT] H-INDIC [H-BODYPART]) (87 [POSTCONJELEM]
. F) (87 [POSTCONJELEM] . 117) (117 [PRECONJELEM] . 87) (139 [
PRECONJELEM] . F) (117 [POSTCONJELEM] . 139) (139 [PRECONJELEM] . 117) (
87 [NOT-DISTR-LN-ATT] . F))])

Figure 1.3, continued: Parse Tree of HIDSM 1.1.13 in Parenthesized Form.

```
* HIDSM    1. 1.13

* THERE WAS NO HX OF STIFF NECK , ABNORMAL

* POSTURING , DIARRHEA OR OTHER FOCAL SIGNS

* COMPATIBLE WITH MENINGITIS .
```

```
(PARSE-CONN (CONJOINED (SA (NULL)) (LCONNR (LCONN (NULL)) (HEADCONN (','
= ',')) (RCONN (NULL))) (SA (NULL)))) (PARSE-CONN (REL-CLAUSE (SA (NULL)
) (LCONNR (LCONN (NULL)) (HEADCONN (N = 'TRN-FILLIN')) (RCONN (NULL))) (
SA (NULL)))) (ASSERTION (SA (NULL)) (SUBJECT (NSTG (LNR (LN (TPOS (LTR (
LT (NULL)) (T = 'NO' % (NEGATIVE H-NEG)))) (QPOS (NULL)) (APOS (NULL)) (
NSPOS (NULL)) (NPOS (NULL))) (NVAR (N = 'HISTORY' % (SINGULAR NONHUMAN
H-REP H-PER PAST H-RECORD)) (N = 'PAST')) (RN (RNP (PN (LP (NULL)) (P =
'OF' % ('OF')) (NSTGO (NSTG (LNR (LN (TPOS (LTR (LT (NULL)) (NULL))) (
QPOS (NULL)) (APOS (ADJADJ (LAR1 (LA (NULL)) (AVAR (LCDA (NULL)) (ADJ =
'STIFF' % (H-INDIC))) (RA1 (NULL))))) (NSPOS (NULL)) (NPOS (NULL))) (
NVAR (N = 'NECK' % (SINGULAR NCOUNT1 NONHUMAN H-BODYPART))) (RN (NULL)))
)))) (NULL))))) (SA (NULL)) (TENSE (NULL)) (SA (NULL)) (VERB (LV (NULL))
 (VVAR (V = 'EXIST' % (H-EVID)) (N = 'PAST')) (RV (NULL))) (SA (NULL)) (
OBJECT (NULLOBJ)) (RV (NULL)) (SA (NULL))) (PARSE-CONN (RELATION (SA (
NULL)) (LCONNR (LCONN (NULL)) (HEADCONN (ADJ = 'COMPATIBLE' % (ASENT1 [(
ATHAT ASHOULD AFORTO)] H-CONN PVAL [('WITH')])) (LPR (LP (NULL)) (P =
'WITH' % ('WITH' VHAVE INSTR CONJ-LIKE)) (RP (NULL)))) (RCONN (NULL))) (
SA (NULL)))) (FRAGMENT (SA (NULL)) (NSTG (LNR (LN (TPOS (LTR (LT (NULL))
 (NULL))) (QPOS (NULL)) (APOS (ADJADJ (LAR1 (LA (NULL)) (AVAR (LCDA (
NULL)) (ADJ = 'STIFF' % (H-INDIC))) (RA1 (NULL))))) (NSPOS (NULL)) (NPOS
 (NULL))) (NVAR (N = 'NECK' % (SINGULAR NCOUNT1 NONHUMAN H-BODYPART))) (
RN (NULL)))) (SA (NULL))) (FRAGMENT (SA (NULL)) (NSTG (LNR (LN (TPOS (
LTR (LT (NULL)) (NULL))) (QPOS (NULL)) (APOS (NULL)) (NSPOS (NULL)) (
NPOS (NULL))) (NVAR (N = 'MENINGITIS' % (SINGULAR NONHUMAN H-DIAG))) (RN
 (NULL)))) (SA (NULL))) (PARSE-CONN (CONJOINED (SA (NULL)) (LCONNR (
LCONN (NULL)) (HEADCONN (',' = ',')) (RCONN (NULL))) (SA (NULL)))) (
PARSE-CONN (REL-CLAUSE (SA (NULL)) (LCONNR (LCONN (NULL)) (HEADCONN (N =
 'TRN-FILLIN')) (RCONN (NULL))) (SA (NULL)))) (ASSERTION (SA (NULL)) (
SUBJECT (NSTG (LNR (LN (TPOS (LTR (LT (NULL)) (T = 'NO' % (NEGATIVE
H-NEG)))) (QPOS (NULL)) (APOS (NULL)) (NSPOS (NULL)) (NPOS (NULL))) (
NVAR (N = 'HISTORY' % (SINGULAR NONHUMAN H-REP H-PER PAST H-RECORD)) (N
= 'PAST')) (RN (RNP (PN (LP (NULL)) (P = 'OF' % ('OF')) (NSTGO (NSTG (
LNR (LN (TPOS (LTR (LT (NULL)) (NULL))) (QPOS (NULL)) (APOS (ADJADJ (
LAR1 (LA (NULL)) (AVAR (LCDA (NULL)) (ADJ = 'ABNORMAL' % (ASENT1 [(
AFORTO)] H-INDIC))) (RA1 (NULL))))) (NSPOS (NULL)) (NPOS (NULL))) (NVAR
(VING = 'POSTURING' % (NOTNSUBJ [(NONHUMAN NSENT1 NSENT2 NSENT3 NTIME1
NTIME2)] NOTNOBJ [(NSENT1 NSENT2 NSENT3 NTIME1 NTIME2)] OBJLIST [(
NULLOBJ)] H-BODYFUNC))) (RN (NULL)))))) (NULL))))) (SA (NULL)) (TENSE (
NULL)) (SA (NULL)) (VERB (LV (NULL)) (VVAR (V = 'EXIST' % (H-EVID)) (N =
 'PAST')) (RV (NULL))) (SA (NULL)) (OBJECT (NULLOBJ)) (RV (NULL)) (SA (
NULL))) (PARSE-CONN (RELATION (SA (NULL)) (LCONNR (LCONN (NULL)) (
HEADCONN (ADJ = 'COMPATIBLE' % (ASENT1 [(ATHAT ASHOULD AFORTO)] H-CONN
PVAL [('WITH')])) (LPR (LP (NULL)) (P = 'WITH' % ('WITH' VHAVE INSTR
CONJ-LIKE)) (RP (NULL)))) (RCONN (NULL))) (SA (NULL)))) (FRAGMENT (SA (
NULL)) (NSTG (LNR (LN (TPOS (LTR (LT (NULL)) (NULL))) (QPOS (NULL)) (
APOS (ADJADJ (LAR1 (LA (NULL)) (AVAR (LCDA (NULL)) (ADJ = 'ABNORMAL' % (
ASENT1 [(AFORTO)] H-INDIC))) (RA1 (NULL))))) (NSPOS (NULL)) (NPOS (NULL)
```

Figure 1.4: Decomposition Tree of HIDSM 1.1.13 in Parenthesized Form.

```
)) (NVAR (VING = 'POSTURING' % (NOTNSUBJ [(NONHUMAN NSENT1 NSENT2 NSENT3
 NTIME1 NTIME2)] NOTNOBJ [(NSENT1 NSENT2 NSENT3 NTIME1 NTIME2)] OBJLIST
 [(NULLOBJ)] H-BODYFUNC))) (RN (NULL)))) (SA (NULL))) (FRAGMENT (SA (
NULL)) (NSTG (LNR (LN (TPOS (LTR (LT (NULL)) (NULL))) (QPOS (NULL)) (
APOS (NULL)) (NSPOS (NULL)) (NPOS (NULL))) (NVAR (N = 'MENINGITIS' % (
SINGULAR NONHUMAN H-DIAG))) (RN (NULL)))) (SA (NULL))) (PARSE-CONN (
CONJOINED (SA (NULL)) (LCONNR (LCONN (NULL)) (HEADCONN ('AND' = 'AND'))
(RCONN (NULL))) (SA (NULL)))) (PARSE-CONN (REL-CLAUSE (SA (NULL)) (
LCONNR (LCONN (NULL)) (HEADCONN (N = 'TRN-FILLIN')) (RCONN (NULL))) (SA
(NULL)))) (ASSERTION (SA (NULL)) (SUBJECT (NSTG (LNR (LN (TPOS (LTR (LT
(NULL)) (T = 'NO' % (NEGATIVE H-NEG)))) (QPOS (NULL)) (APOS (NULL)) (
NSPOS (NULL)) (NPOS (NULL))) (NVAR (N = 'HISTORY' % (SINGULAR NONHUMAN
H-REP H-PER PAST H-RECORD)) (N = 'PAST')) (RN (RNP (PN (LP (NULL)) (P =
'OF' % ('OF')) (NSTGO (NSTG (LNR (LN (TPOS (LTR (LT (NULL)) (NULL))) (
QPOS (NULL)) (APOS (NULL)) (NSPOS (NULL)) (NPOS (NULL))) (NVAR (N =
'DIARRHEA' % (SINGULAR NONHUMAN H-INDIC))) (RN (NULL)))))) (NULL))))) (
SA (NULL)) (TENSE (NULL)) (SA (NULL)) (VERB (LV (NULL)) (VVAR (V =
'EXIST' % (H-EVID)) (N = 'PAST')) (RV (NULL))) (SA (NULL)) (OBJECT (
NULLOBJ)) (RV (NULL)) (SA (NULL))) (PARSE-CONN (RELATION (SA (NULL)) (
LCONNR (LCONN (NULL)) (HEADCONN (ADJ = 'COMPATIBLE' % (ASENT1 [(ATHAT
ASHOULD AFORTO)] H-CONN PVAL [('WITH')])) (LPR (LP (NULL)) (P = 'WITH' %
 ('WITH' VHAVE INSTR CONJ-LIKE)) (RP (NULL)))) (RCONN (NULL))) (SA (NULL
)))) (FRAGMENT (SA (NULL)) (NSTG (LNR (LN (TPOS (LTR (LT (NULL)) (NULL))
) (QPOS (NULL)) (APOS (NULL)) (NSPOS (NULL)) (NPOS (NULL))) (NVAR (N =
'DIARRHEA' % (SINGULAR NONHUMAN H-INDIC))) (RN (NULL)))) (SA (NULL))) (
FRAGMENT (SA (NULL)) (NSTG (LNR (LN (TPOS (LTR (LT (NULL)) (NULL))) (
QPOS (NULL)) (APOS (NULL)) (NSPOS (NULL)) (NPOS (NULL))) (NVAR (N =
'MENINGITIS' % (SINGULAR NONHUMAN H-DIAG))) (RN (NULL)))) (SA (NULL))) (
PARSE-CONN (REL-CLAUSE (SA (NULL)) (LCONNR (LCONN (NULL)) (HEADCONN (N =
 'TRN-FILLIN')) (RCONN (NULL))) (SA (NULL)))) (ASSERTION (SA (NULL)) (
SUBJECT (NSTG (LNR (LN (TPOS (LTR (LT (NULL)) (T = 'NO' % (NEGATIVE
H-NEG)))) (QPOS (NULL)) (APOS (NULL)) (NSPOS (NULL)) (NPOS (NULL))) (
NVAR (N = 'HISTORY' % (SINGULAR NONHUMAN H-REP H-PER PAST H-RECORD)) (N
= 'PAST')) (RN (RNP (PN (LP (NULL)) (P = 'OF' % ('OF')) (NSTGO (NSTG (
LNR (LN (TPOS (LTR (LT (NULL)) (NULL))) (QPOS (NULL)) (APOS (ADJADJ (
LAR1 (LA (NULL)) (AVAR (LCDA (NULL)) (ADJ = 'OTHER' % (APREQ))) (RA1 (
NULL))))) (NSPOS (NULL)) (NPOS (NULL))) (NVAR (N = 'FOCAL' 'SIGN' % (
SINGULAR NONHUMAN NCOUNT1 H-INDIC)) (N = 'PLURAL')) (RN (NULL)))))) (
NULL))))) (SA (NULL)) (TENSE (NULL)) (SA (NULL)) (VERB (LV (NULL)) (VVAR
 (V = 'EXIST' % (H-EVID)) (N = 'PAST')) (RV (NULL))) (SA (NULL)) (OBJECT
 (NULLOBJ)) (RV (NULL)) (SA (NULL))) (PARSE-CONN (RELATION (SA (NULL)) (
LCONNR (LCONN (NULL)) (HEADCONN (ADJ = 'COMPATIBLE' % (ASENT1 [(ATHAT
ASHOULD AFORTO)] H-CONN PVAL [('WITH')])) (LPR (LP (NULL)) (P = 'WITH' %
 ('WITH' VHAVE INSTR CONJ-LIKE)) (RP (NULL)))) (RCONN (NULL))) (SA (NULL
)))) (FRAGMENT (SA (NULL)) (NSTG (LNR (LN (TPOS (LTR (LT (NULL)) (NULL))
) (QPOS (NULL)) (APOS (NULL)) (NSPOS (NULL)) (NPOS (NULL))) (NVAR (N =
'FOCAL' 'SIGN' % (SINGULAR NONHUMAN NCOUNT1 H-INDIC)) (N = 'PLURAL')) (
RN (NULL)))) (SA (NULL))) (FRAGMENT (SA (NULL)) (NSTG (LNR (LN (TPOS (
LTR (LT (NULL)) (NULL))) (QPOS (NULL)) (APOS (NULL)) (NSPOS (NULL)) (
NPOS (NULL))) (NVAR (N = 'MENINGITIS' % (SINGULAR NONHUMAN H-DIAG))) (RN
 (NULL)))) (SA (NULL)))) (ENDMARK ('.' = '.'))) (MORESENT (NULL)))) [((
89 [COMPUTED-ATT] H-INDIC [H-BODYPART]) (163 [COMPUTED-ATT] H-INDIC [
H-BODYPART]) (89 [N-TO-LN-ATT] . 75) (163 [N-TO-LN-ATT] . 149) (89 [
SELECT-ATT] H-BODYPART) (163 [SELECT-ATT] H-BODYPART) (81 [SELECT-ATT]
H-INDIC) (155 [SELECT-ATT] H-INDIC) (275 [COMPUTED-ATT] H-RESULT [
H-BODYFUNC]) (349 [COMPUTED-ATT] H-RESULT [H-BODYFUNC]) (275 [
```

Figure 1.4, cont: Decomposition Tree of HIDSM 1.1.13 in Parenthesized Form.

N-TO-LN-ATT] . 261) (349 [N-TO-LN-ATT] . 335) (275 [SELECT-ATT]
H-BODYFUNC) (349 [SELECT-ATT] H-BODYFUNC) (267 [SELECT-ATT] H-INDIC) (
341 [SELECT-ATT] H-INDIC) (88 [POSTCONJELEM] . 274) (162 [POSTCONJELEM]
. 274) (88 [POSTCONJELEM] . 348) (162 [POSTCONJELEM] . 348) (274 [
PRECONJELEM] . 88) (274 [PRECONJELEM] . 162) (348 [PRECONJELEM] . 88) (
348 [PRECONJELEM] . 162) (275 [COMPUTED-ATT] H-RESULT [H-BODYFUNC]) (349
[COMPUTED-ATT] H-RESULT [H-BODYFUNC]) (275 [SELECT-ATT] H-BODYFUNC) (
349 [SELECT-ATT] H-BODYFUNC) (89 [COMPUTED-ATT] H-INDIC [H-BODYPART]) (
163 [COMPUTED-ATT] H-INDIC [H-BODYPART]) (89 [SELECT-ATT] H-BODYPART) (
163 [SELECT-ATT] H-BODYPART) (275 [MATCHED] . 89) (275 [MATCHED] . 163)
(349 [MATCHED] . 89) (349 [MATCHED] . 163) (274 [POSTCONJELEM] . 451) (
348 [POSTCONJELEM] . 451) (274 [POSTCONJELEM] . 516) (348 [POSTCONJELEM]
. 516) (451 [PRECONJELEM] . 274) (451 [PRECONJELEM] . 348) (516 [
PRECONJELEM] . 274) (516 [PRECONJELEM] . 348) (251 [NOT-DISTR-LN-ATT] .
451) (325 [NOT-DISTR-LN-ATT] . 451) (251 [NOT-DISTR-LN-ATT] . 516) (325
[NOT-DISTR-LN-ATT] . 516) (452 [SELECT-ATT] H-INDIC) (517 [SELECT-ATT]
H-INDIC) (275 [COMPUTED-ATT] H-RESULT [H-BODYFUNC]) (349 [COMPUTED-ATT]
H-RESULT [H-BODYFUNC]) (275 [SELECT-ATT] H-BODYFUNC) (349 [SELECT-ATT]
H-BODYFUNC) (452 [MATCHED] . 275) (452 [MATCHED] . 349) (517 [MATCHED]
. 275) (517 [MATCHED] . 349) (451 [POSTCONJELEM] . 615) (516 [
POSTCONJELEM] . 615) (451 [POSTCONJELEM] . 681) (516 [POSTCONJELEM] .
681) (615 [PRECONJELEM] . 451) (615 [PRECONJELEM] . 516) (681 [
PRECONJELEM] . 451) (681 [PRECONJELEM] . 516) (616 [SELECT-ATT] H-INDIC)
(682 [SELECT-ATT] H-INDIC) (452 [SELECT-ATT] H-INDIC) (517 [SELECT-ATT]
H-INDIC) (616 [MATCHED] . 452) (616 [MATCHED] . 517) (682 [MATCHED] .
452) (682 [MATCHED] . 517) (251 [POSTCONJELEM] . 592) (325 [POSTCONJELEM
] . 592) (592 [PRECONJELEM] . 251) (592 [PRECONJELEM] . 325) (65 [
POSTCONJELEM] . 251) (139 [POSTCONJELEM] . 251) (65 [POSTCONJELEM] . 325
) (139 [POSTCONJELEM] . 325) (251 [PRECONJELEM] . 65) (251 [PRECONJELEM]
. 139) (325 [PRECONJELEM] . 65) (325 [PRECONJELEM] . 139) (127 [
PVAL-ATT] . 123) (313 [PVAL-ATT] . 309) (490 [PVAL-ATT] . 486) (655 [
PVAL-ATT] . 651) (123 [SELECT-ATT] H-CONN) (309 [SELECT-ATT] H-CONN) (
486 [SELECT-ATT] H-CONN) (651 [SELECT-ATT] H-CONN) (188 [SELECT-ATT]
H-DIAG) (374 [SELECT-ATT] H-DIAG) (542 [SELECT-ATT] H-DIAG) (708 [
SELECT-ATT] H-DIAG) (89 [SELECT-ATT] H-BODYPART) (163 [SELECT-ATT]
H-BODYPART) (89 [COMPUTED-ATT] H-INDIC [H-BODYPART]) (163 [COMPUTED-ATT]
H-INDIC [H-BODYPART]) (123 [SELECT-ATT] H-CONN) (309 [SELECT-ATT]
H-CONN) (486 [SELECT-ATT] H-CONN) (651 [SELECT-ATT] H-CONN) (188 [
SELECT-ATT] H-DIAG) (374 [SELECT-ATT] H-DIAG) (542 [SELECT-ATT] H-DIAG)
(708 [SELECT-ATT] H-DIAG) (616 [SELECT-ATT] H-INDIC) (682 [SELECT-ATT]
H-INDIC) (123 [SELECT-ATT] H-CONN) (309 [SELECT-ATT] H-CONN) (486 [
SELECT-ATT] H-CONN) (651 [SELECT-ATT] H-CONN) (188 [SELECT-ATT] H-DIAG)
(374 [SELECT-ATT] H-DIAG) (542 [SELECT-ATT] H-DIAG) (708 [SELECT-ATT]
H-DIAG) (452 [SELECT-ATT] H-INDIC) (517 [SELECT-ATT] H-INDIC) (123 [
SELECT-ATT] H-CONN) (309 [SELECT-ATT] H-CONN) (486 [SELECT-ATT] H-CONN)
(651 [SELECT-ATT] H-CONN) (188 [SELECT-ATT] H-DIAG) (374 [SELECT-ATT]
H-DIAG) (542 [SELECT-ATT] H-DIAG) (708 [SELECT-ATT] H-DIAG) (275 [
SELECT-ATT] H-BODYFUNC) (349 [SELECT-ATT] H-BODYFUNC) (275 [COMPUTED-ATT
] H-RESULT [H-BODYFUNC]) (349 [COMPUTED-ATT] H-RESULT [H-BODYFUNC]) (54
[COMPUTED-ATT] H-INDIC [H-PER]) (240 [COMPUTED-ATT] H-INDIC [H-PER]) (
426 [COMPUTED-ATT] H-INDIC [H-PER]) (581 [COMPUTED-ATT] H-INDIC [H-PER])
(54 [N-TO-RN-ATT] . 64) (54 [N-TO-RN-ATT] . 138) (240 [N-TO-RN-ATT] .
64) (240 [N-TO-RN-ATT] . 138) (426 [N-TO-RN-ATT] . 64) (426 [N-TO-RN-ATT
] . 138) (581 [N-TO-RN-ATT] . 64) (581 [N-TO-RN-ATT] . 138) (54 [
N-TO-RN-ATT] . 250) (54 [N-TO-RN-ATT] . 324) (240 [N-TO-RN-ATT] . 250) (
240 [N-TO-RN-ATT] . 324) (426 [N-TO-RN-ATT] . 250) (426 [N-TO-RN-ATT] .

Fig. 1.4, cont: Decomposition Tree of HIDSM 1.1.13 in Parenthesized Form.

324) (581 [N-TO-RN-ATT] . 250) (581 [N-TO-RN-ATT] . 324) (54 [
N-TO-RN-ATT] . 436) (54 [N-TO-RN-ATT] . 501) (240 [N-TO-RN-ATT] . 436) (
240 [N-TO-RN-ATT] . 501) (426 [N-TO-RN-ATT] . 436) (426 [N-TO-RN-ATT] .
501) (581 [N-TO-RN-ATT] . 436) (581 [N-TO-RN-ATT] . 501) (54 [
N-TO-RN-ATT] . 591) (54 [N-TO-RN-ATT] . 666) (240 [N-TO-RN-ATT] . 591) (
240 [N-TO-RN-ATT] . 666) (426 [N-TO-RN-ATT] . 591) (426 [N-TO-RN-ATT] .
666) (581 [N-TO-RN-ATT] . 591) (581 [N-TO-RN-ATT] . 666) (54 [SELECT-ATT
] H-REP H-PER) (240 [SELECT-ATT] H-REP H-PER) (426 [SELECT-ATT] H-REP
H-PER) (581 [SELECT-ATT] H-REP H-PER) (89 [SELECT-ATT] H-BODYPART) (163
[SELECT-ATT] H-BODYPART) (89 [COMPUTED-ATT] H-INDIC [H-BODYPART]) (163
[COMPUTED-ATT] H-INDIC [H-BODYPART]) (251 [POSTCONJELEM] . F) (325 [
POSTCONJELEM] . F) (251 [POSTCONJELEM] . 437) (325 [POSTCONJELEM] . 437)
(251 [POSTCONJELEM] . 502) (325 [POSTCONJELEM] . 502) (437 [PRECONJELEM
] . 251) (437 [PRECONJELEM] . 325) (502 [PRECONJELEM] . 251) (502 [
PRECONJELEM] . 325) (592 [PRECONJELEM] . F) (437 [POSTCONJELEM] . 592) (
502 [POSTCONJELEM] . 592) (592 [PRECONJELEM] . 437) (592 [PRECONJELEM]
. 502) (251 [NOT-DISTR-LN-ATT] . F) (325 [NOT-DISTR-LN-ATT] . F) (37 [
INDEX] . N282) (223 [INDEX] . N282) (409 [INDEX] . N282) (564 [INDEX] .
N282) (40 [INDEX] . N283) (226 [INDEX] . N283) (412 [INDEX] . N283) (567
[INDEX] . N283) (44 [INDEX] . N284) (230 [INDEX] . N284) (416 [INDEX]
. N284) (571 [INDEX] . N284) (54 [INDEX] . N285) (240 [INDEX] . N285) (
426 [INDEX] . N285) (581 [INDEX] . N285) (61 [INDEX] . N286) (247 [INDEX
] . N286) (433 [INDEX] . N286) (588 [INDEX] . N286) (63 [INDEX] . N287)
(249 [INDEX] . N287) (435 [INDEX] . N287) (590 [INDEX] . N287) (81 [
INDEX] . N288) (155 [INDEX] . N288) (89 [INDEX] . N289) (163 [INDEX] .
N289) (15 [INDEX] . N290) (267 [INDEX] . N291) (341 [INDEX] . N291) (275
[INDEX] . N292) (349 [INDEX] . N292) (201 [INDEX] . N293) (452 [INDEX]
. N294) (517 [INDEX] . N294) (608 [INDEX] . N296) (616 [INDEX] . N297)
(682 [INDEX] . N297) (123 [INDEX] . N298) (309 [INDEX] . N298) (486 [
INDEX] . N298) (651 [INDEX] . N298) (127 [INDEX] . N299) (313 [INDEX] .
N299) (490 [INDEX] . N299) (655 [INDEX] . N299) (171 [INDEX] . N300) (
357 [INDEX] . N300) (525 [INDEX] . N300) (691 [INDEX] . N300) (188 [
INDEX] . N301) (374 [INDEX] . N301) (542 [INDEX] . N301) (708 [INDEX] .
N301) (714 [INDEX] . N302) (103 [TENSE-ATT] PAST) (289 [TENSE-ATT] PAST)
(466 [TENSE-ATT] PAST) (631 [TENSE-ATT] PAST) (54 [TENSE-ATT] PAST) (
240 [TENSE-ATT] PAST) (426 [TENSE-ATT] PAST) (581 [TENSE-ATT] PAST) (251
[PRECONJELEM] . F) (325 [PRECONJELEM] . F) (65 [POSTCONJELEM] . F) (139
[POSTCONJELEM] . F) (274 [PRECONJELEM] . F) (348 [PRECONJELEM] . F) (88
[POSTCONJELEM] . F) (162 [POSTCONJELEM] . F) (437 [PRECONJELEM] . F) (
502 [PRECONJELEM] . F) (251 [POSTCONJELEM] . F) (325 [POSTCONJELEM] . F)
(451 [PRECONJELEM] . F) (516 [PRECONJELEM] . F) (274 [POSTCONJELEM] . F
) (348 [POSTCONJELEM] . F) (592 [PRECONJELEM] . F) (437 [POSTCONJELEM]
. F) (502 [POSTCONJELEM] . F) (615 [PRECONJELEM] . F) (681 [PRECONJELEM
] . F) (451 [POSTCONJELEM] . F) (516 [POSTCONJELEM] . F) (436 [
POSTCONJELEM] . 591) (501 [POSTCONJELEM] . 591) (436 [POSTCONJELEM] .
666) (501 [POSTCONJELEM] . 666) (591 [PRECONJELEM] . 436) (591 [
PRECONJELEM] . 501) (666 [PRECONJELEM] . 436) (666 [PRECONJELEM] . 501)
(250 [POSTCONJELEM] . 436) (324 [POSTCONJELEM] . 436) (250 [POSTCONJELEM
] . 501) (324 [POSTCONJELEM] . 501) (436 [PRECONJELEM] . 250) (436 [
PRECONJELEM] . 324) (501 [PRECONJELEM] . 250) (501 [PRECONJELEM] . 324)
(64 [POSTCONJELEM] . 250) (138 [POSTCONJELEM] . 250) (64 [POSTCONJELEM]
. 324) (138 [POSTCONJELEM] . 324) (250 [PRECONJELEM] . 64) (250 [
PRECONJELEM] . 138) (324 [PRECONJELEM] . 64) (324 [PRECONJELEM] . 138) (
250 [PRECONJELEM] . F) (324 [PRECONJELEM] . F) (64 [POSTCONJELEM] . F) (
138 [POSTCONJELEM] . F) (436 [PRECONJELEM] . F) (501 [PRECONJELEM] . F)
(250 [POSTCONJELEM] . F) (324 [POSTCONJELEM] . F) (591 [PRECONJELEM] . F

Figure 1.4, cont: Decomposition Tree of HIDSM 1.1.13 in Parenthesized Form.

) (666 [PRECONJELEM] . F) (436 [POSTCONJELEM] . F) (501 [POSTCONJELEM] . F) (435 [POSTCONJELEM] . 590) (590 [PRECONJELEM] . 435) (249 [POSTCONJELEM] . 435) (435 [PRECONJELEM] . 249) (63 [POSTCONJELEM] . 249) (249 [PRECONJELEM] . 63) (249 [PRECONJELEM] . F) (63 [POSTCONJELEM] . F) (435 [PRECONJELEM] . F) (249 [POSTCONJELEM] . F) (590 [PRECONJELEM] . F) (435 [POSTCONJELEM] . F) (434 [POSTCONJELEM] . 589) (589 [PRECONJELEM] . 434) (248 [POSTCONJELEM] . 434) (434 [PRECONJELEM] . 248) (62 [POSTCONJELEM] . 248) (248 [PRECONJELEM] . 62) (248 [PRECONJELEM] . F) (62 [POSTCONJELEM] . F) (434 [PRECONJELEM] . F) (248 [POSTCONJELEM] . F) (589 [PRECONJELEM] . F) (434 [POSTCONJELEM] . F) (430 [POSTCONJELEM] . 585) (585 [PRECONJELEM] . 430) (244 [POSTCONJELEM] . 430) (430 [PRECONJELEM] . 244) (58 [POSTCONJELEM] . 244) (244 [PRECONJELEM] . 58) (244 [PRECONJELEM] . F) (58 [POSTCONJELEM] . F) (430 [PRECONJELEM] . F) (244 [POSTCONJELEM] . F) (585 [PRECONJELEM] . F) (430 [POSTCONJELEM] . F) (429 [POSTCONJELEM] . 584) (584 [PRECONJELEM] . 429) (243 [POSTCONJELEM] . 429) (429 [PRECONJELEM] . 243) (57 [POSTCONJELEM] . 243) (243 [PRECONJELEM] . 57) (243 [PRECONJELEM] . F) (57 [POSTCONJELEM] . F) (429 [PRECONJELEM] . F) (243 [POSTCONJELEM] . F) (584 [PRECONJELEM] . F) (429 [POSTCONJELEM] . F) (428 [POSTCONJELEM] . 583) (583 [PRECONJELEM] . 428) (242 [POSTCONJELEM] . 428) (428 [PRECONJELEM] . 242) (56 [POSTCONJELEM] . 242) (242 [PRECONJELEM] . 56) (242 [PRECONJELEM] . F) (56 [POSTCONJELEM] . F) (428 [PRECONJELEM] . F) (242 [POSTCONJELEM] . F) (583 [PRECONJELEM] . F) (428 [POSTCONJELEM] . F) (410 [POSTCONJELEM] . 565) (565 [PRECONJELEM] . 410) (224 [POSTCONJELEM] . 410) (410 [PRECONJELEM] . 224) (38 [POSTCONJELEM] . 224) (224 [PRECONJELEM] . 38) (224 [PRECONJELEM] . F) (38 [POSTCONJELEM] . F) (410 [PRECONJELEM] . F) (224 [POSTCONJELEM] . F) (565 [PRECONJELEM] . F) (410 [POSTCONJELEM] . F) (409 [POSTCONJELEM] . 564) (564 [PRECONJELEM] . 409) (223 [POSTCONJELEM] . 409) (409 [PRECONJELEM] . 223) (37 [POSTCONJELEM] . 223) (223 [PRECONJELEM] . 37) (223 [PRECONJELEM] . F) (37 [POSTCONJELEM] . F) (409 [PRECONJELEM] . F) (223 [POSTCONJELEM] . F) (564 [PRECONJELEM] . F) (409 [POSTCONJELEM] . F) (408 [POSTCONJELEM] . 563) (563 [PRECONJELEM] . 408) (222 [POSTCONJELEM] . 408) (408 [PRECONJELEM] . 222) (36 [POSTCONJELEM] . 222) (222 [PRECONJELEM] . 36) (222 [PRECONJELEM] . F) (36 [POSTCONJELEM] . F) (408 [PRECONJELEM] . F) (222 [POSTCONJELEM] . F) (563 [PRECONJELEM] . F) (408 [POSTCONJELEM] . F) (405 [POSTCONJELEM] . 560) (560 [PRECONJELEM] . 405) (219 [POSTCONJELEM] . 405) (405 [PRECONJELEM] . 219) (33 [POSTCONJELEM] . 219) (219 [PRECONJELEM] . 33) (582 [SELECT-ATT] H-VTENSE) (582 [HOST-ASP] . 581) (581 [N-TO-RN-ATT] . 591) (581 [HOST-ASP] . 616) (568 [HOST-ASP] . 581) (683 [SELECT-ATT] H-REP) (683 [HOST-ASP] . 682) (617 [SELECT-ATT] H-REP) (617 [HOST-ASP] . 616) (632 [SELECT-ATT] H-VTENSE) (632 [HOST-ASP] . 631) (627 [HOST-ASP] . 581) (427 [SELECT-ATT] H-VTENSE) (427 [HOST-ASP] . 426) (426 [N-TO-RN-ATT] . 436) (426 [HOST-ASP] . 452) (413 [HOST-ASP] . 426) (467 [SELECT-ATT] H-VTENSE) (467 [HOST-ASP] . 466) (462 [HOST-ASP] . 426) (241 [SELECT-ATT] H-VTENSE) (241 [HOST-ASP] . 240) (240 [N-TO-RN-ATT] . 250) (240 [HOST-ASP] . 275) (227 [HOST-ASP] . 240) (349 [N-TO-LN-ATT] . 335) (349 [HOST-ASP] . 341) (275 [N-TO-LN-ATT] . 261) (275 [HOST-ASP] . 267) (290 [SELECT-ATT] H-VTENSE) (290 [HOST-ASP] . 289) (285 [HOST-ASP] . 240) (55 [SELECT-ATT] H-VTENSE) (55 [HOST-ASP] . 54) (54 [N-TO-RN-ATT] . 64) (54 [HOST-ASP] . 89) (41 [HOST-ASP] . 54) (163 [N-TO-LN-ATT] . 149) (163 [HOST-ASP] . 155) (89 [N-TO-LN-ATT] . 75) (89 [HOST-ASP] . 81) (104 [SELECT-ATT] H-VTENSE) (104 [HOST-ASP] . 103) (99 [HOST-ASP] . 54))])

Figure 1.4, cont: Decomposition Tree of HIDSM 1.1.13 in Parenthesized Form.

dominance were expressed in the tree structure. For this purpose, additional computer processing of the discharge summary material was necessary.

Since the question of extracting, representing, and counting substructures of computer-generated parse trees is of general interest for measurements of the information contained in such trees, we decided on a program of some generality for making the "skeletons". The input to the program is a specification of which substructures are required and how they are to be represented. The "skeletonizing" program scans each parse tree, extracting the desired substructures in the desired mode of representation. This sequence of substructures is then submitted to a simple bottom-up parser for correct nesting and marking of distinguished operators.

We hope in further research in this area to link the "skeletonizing" program to a high-level description of the substructures of interest, so that researchers sophisticated in linguistics or information science but unsophisticated in computer science can use the program to query collections of parse trees. Such a system would compile the user's description into tables suitable for input to the skeletonizing program, which would then display, store, and statistically manipulate the substructures described.

For the research summarized in this Report, tables for extracting simplified operator-argument skeletons from the decomposition trees were manually constructed, and run on the decomposition trees from the discharge summary material. The skeleton of our example sentence is shown in Fig. 1.5. A full listing of the discharge summary sentences used in this research and their skeletons is shown in Appendix A1.

The lipid-review sentences of T2 were analyzed manually into an operator-argument tree-structure. A sample sentence and its tree representation are shown in Fig. 1.6, and the full set of sentences together with their representations and an explantion of methods and conventions used in the notation is shown in Appendix A2.

3. Commensurability of the Decompositions of the Two Types of Text with Respect to the Counting of Operators.

One feature of the lipid-review analyses (and one reason why they could not be processed by machine) was the distinction made in these analyses between what we may call logical complexity, which is based on the global relations between arguments and operators, and the total complexity, which reflects the contribution of modifiers whose action is purely local. As an example (drawn from Note 3 in Appendix A2)

The cholesterol ester formed in HDL ...

is analyzed as

CE * (be formed in/ CE/ HDL)

```
* HIDSM    1. 1.13

* THERE WAS NO HX OF STIFF NECK , ABNORMAL

* POSTURING , DIARRHEA OR OTHER FOCAL SIGNS

* COMPATIBLE WITH MENINGITIS .

 (
   (CONJOINED=',' )
   (
      (REL-CLAUSE='TRN-FILLIN' )
      (ASSERTION=
         (SUBJECT=
            (NEG-FIND='NO' 'HISTORY' 'PAST' 'OF' 'STIFF' 'NECK' )
         )
         (VERB='EXIST' 'PAST' )
      )
      (
         (RELATION='COMPATIBLE' 'WITH' )
         (FRAGMENT=
            (NSTG='STIFF' 'NECK' )
         )
         (FRAGMENT=
            (NSTG='MENINGITIS' )
         )
      )
   )
   (CONJOINED=',' )
   (
      (REL-CLAUSE='TRN-FILLIN' )
      (ASSERTION=
         (SUBJECT=
            (NEG-FIND='NO' 'HISTORY' 'PAST' 'OF' 'ABNORMAL' 'POSTURING' )
         )
         (VERB='EXIST' 'PAST' )
      )
      (
         (RELATION='COMPATIBLE' 'WITH' )
         (FRAGMENT=
            (NSTG='ABNORMAL' 'POSTURING' )
         )
         (FRAGMENT=
            (NSTG='MENINGITIS' )
         )
      )
   )
   (-CONJOINED='AND' )
   (
      (REL-CLAUSE='TRN-FILLIN' )
      (ASSERTION=
```

Fig. 1.5: Skeleton of HIDSM 1.1.13.

```
          (SUBJECT=
             (NEG-FIND='NO' 'HISTORY' 'PAST' 'OF' 'DIARRHEA' )
          )
          (VERB='EXIST' 'PAST' )
       )
       (
          (RELATION='COMPATIBLE' 'WITH' )
          (FRAGMENT=
             (NSTG='DIARRHEA' )
          )
          (FRAGMENT=
             (NSTG='MENINGITIS' )
          )
       )
    )
    (
       (REL-CLAUSE='TRN-FILLIN' )
       (ASSERTION=
          (SUBJECT=
             (NEG-FIND='NO' 'HISTORY' 'PAST' 'OF' 'OTHER' 'FOCAL' 'SIGN'
                       'PLURAL' )
          )
          (VERB='EXIST' 'PAST' )
       )
       (
          (RELATION='COMPATIBLE' 'WITH' )
          (FRAGMENT=
             (NSTG='FOCAL' 'SIGN' 'PLURAL' )
          )
          (FRAGMENT=
             (NSTG='MENINGITIS' )
          )
       )
    )
 )
```

Figure 1.5, continued: Skeleton of HIDSM 1.1.13.

Figure 1.6: Example of Transformational Decomposition in Operator-Argument Form.

UBS 6: Cholesterol ester is removed from LDL in the splanchnic bed and LDL particle stability maintained by addition of triglyceride.

```
SBS 6:1  (and(be removed from *in the splanchnic bed
                          /CE
                          /LDL)
              (by(maintained (stability/LDL))
                 (addition of/TG)))
```

UBS 7: The cholesterol derived from LDL is reused in the synthesis of VLDL and HDL.

Note: Notational conventions and abbreviations are given in Appendix A2.

where "be formed in HDL" is a local modifier of cholesterol ester (CE). In grammatical terms, "formed in HDL" is a reduced relative clause modifying the noun CE. This construction would be represented in the output of the computerized transformational decomposition as

```
(REL-CLAUSE...
(ASSERTION=
  (SUBJECT='CE')
  rest of assertion
)
(ASSERTION=
  (VERB='FORM' 'IN' )
  (OBJECT='HDL' )
)
```

Since we wished to accumulate measures for the local modifiers ("adjuncts") separately, we needed a way to make the two representations commensurate. A diagram showing this transformation appears in Fig. 1.7. Because of the more regular format of the computer-generated information, it was easier to map trees from the discharge-material sentences to the lipid sentences rather than *vice versa*. Also, the mapping was only done virtually in the execution space of the program in order to save the original skeleton representations of the discharge-material trees.

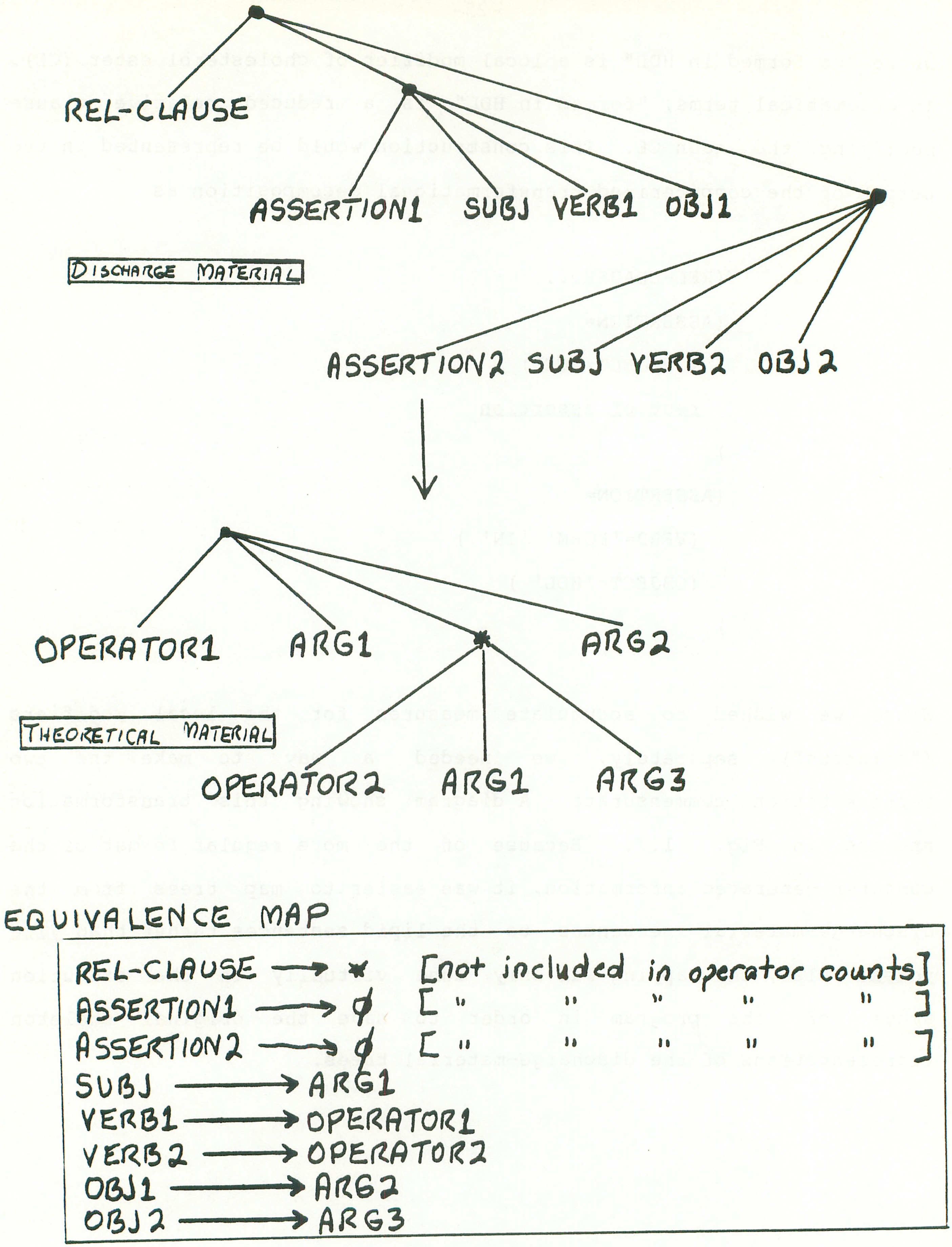


Figure 1.7: Equivalence of Operator Counts.

Another step to make the two representations commensurate was necessary. The transformational processor in the LSP system treats all conjunctions as binary operators (a sequence of conjoined ASSERTIONS has a conjunction connective between each ASSERTION), while in the lipid material all conjuncts at the same level were (more properly, from the point of view of information complexity studies) treated as n-ary operators (a sequence of conjoined subtrees is operated on by a single conjunction operator). In order to make the skeleton data conform to this analysis, the bottom-up parser of the "skeletonizer" program was used to mark the first conjunct on each level. For statistical purposes, only this distinguished conjunct was counted in the operator scores.

Other points that were considered in relation to commensurability of the two representations are the following.

1. Both vocabularies have phrases in which the component elements are not treated as operators although in other environments the same morpheme or word class may be an operator:

in FACT-REPORT texts (Appendix A1)

HIDSM 3.1.4 teeth chatter<u>ing</u> chills

HIDSM 3.1.13 alive <u>and</u> well (a standard expression in Family History)

EXDSM 1.1.3 respiration <u>rate</u>

EXDSM 1.1.21 localiz<u>ing</u> signs

in THEORY texts (Appendix A2)

UBKG 2.9.7 the 2-position of lecithin

UBKG 3.3.5 mutant FH cells

UBKG 4.4.3 membrane binding assay

These are fixed phrases in these sublanguages.

In addition, phrases that are not fixed expressions for the sublanguage, but contain adjectives or other modifiers of peripheral importance (usually also of low frequency) in the texts are not treated as operators. Thus in Appendix A1, our in our emergency room (HIDSM 1.1.8) is not counted as an operator, and in Appendix A2, excess in excess surface material (UBDG 2.9.6) is not treated as an operator.

2. In both types of texts adverbs and adverbial PN's are considered local adjuncts of the associated operator and are not counted as separate operators. This includes locatives and time expressions. Time expressions are more frequent in the FACT-type (medical) texts, whereas locatives are more frequent in the THEORY (physiological mechanism) texts. In the latter, locatives include some words that would not ordinarily be thought of as locations (e.g. diet) but these are source-sites for the substances that appear as arguments of associated operators.

Examples of adverbial modifiers are:

LOCATIVES

HIDSM 2.1.8	to the floor	(in Appendix A1)
HIDSM 2.1.13	in Orlando, Florida	(in Appendix A1)
EXDSM 2.1.4	around both eyes	(in Appendix A1)
UBKG 2.9.2	in capillaries	(in Appendix A2)
UBKG 3.3.5	in tissue culture	(in Appendix A2)

TIME

HIDSM 2.1.7	last week
UBKG 2.6.3	after a single high cholesterol meal

OTHER ADVERBS

EXDSM 2.1.22	grossly
UBKG 2.9.6	mostly
UBKG 4.6.6	intravenously

3. Negation and modal operators (e.g. _no_, _would_) are counted as operators in both data sets.

4. The conventions regarding quantity expressions are the following:

4)a) Quantity nouns (_amount_, _rate_, etc.), occurring as the argument of a quantity verb (_increase_, _decrease_, etc.) are taken as separate operators since their dominating operator requires it (but _respiration rate_ is not decomposed in Appendix 1, cf. 1 above).

4b) If a quantity verb occurs without a quantity noun as its argument, an abstract operator (such as QAMT) is supplied as its argument since one is required (this happens only rarely in the data of this study, cf. Appendix A2, UBS 1).

4c) A numerical quantity with its units is treated as a separate operator. For example, in the FACT texts of Appendix 1,

EXDSM 3.1.5 RESP. 14.

contains one operator. Similarly (with the verb be as carrier of the operator)

HIDSM 1.1.16 CHILD WAS 5 LBS. 1 OZ.

is counted as containing just one operator. Instances of quantity operators in the THEORY texts of Appendix 2 are in UBS 16, UBKG 2.7.3, 2.7.5, 2.7.6, and 2.14.2.

4d) Non-numerical quantifiers (e.g. much in UBKG 2.7.3 in Appendix A2 and severe in HIDSM 2.1.8 in Appendix A1) are treated as operators.

There were several exceptions to the commensurability of quantity operators in each data set that offset each other (2 in A1 and 3 in A2). They were as follows:

In Appendix A1, the expressions for age in HIDSM 2.1.21 and HIDSM

2.1.23 (the age of the patient's siblings) were not treated as operators.

In Appendix A2, numerical expressions were not treated as operators in UBKG 4.4.3 (tenfold in tenfold increase) and in UBKG 4.6.7 (50 percent in 50 percent reduction, and remarkable 75% in remarkable 75% reduction). The numerical quantities here are associated with already-present quantity operators and also have a different character from those in the other examples (e.g. tenfold vs. 100 grams). Because of these features they were not counted as separate operators. Further analysis of different types of quantity expressions is being pursued in another study.

5. The same connectives are treated as operators in the decompositions of the two types of texts (with one exception noted below) even though their representation in the decompositions is different, and required the mapping described above. Thus, in the discharge summary decompositions the relative clause connective, the operator REL-CLAUSE in Appendix A1, is represented by * in the lipid text decompositions of Appendix A2. [See Figure 1.7 above for the counting equivalence.] Similarly, subordinate conjunctions in Appendix A1 give rise to the node SUB-CONJ whose value is the subordinate conjunction word (e.g. when in HIDSM 1.1.1), while a similar operator in Appendix 2 is represented by the word itself (UBKG 2.8.3). Connective verbs are named under the node RELATION in the decompositions in Appendix A1 (e.g. showed in EXDSM 2.1.4) whereas they stand for themselves in operator position in the decompositions in Appendix A2 (UBS 13).

A difference in the two representations occurs with regard to the treatment of comparatives. In the lipid text decompositions, the comparative word other was treated as an operator although than was not present (UBKG 2.7.2), whereas in the discharge summary texts this was not the case. There was one such occurrence, other in no other focal signs in HIDSM 1.1.13.

6. Verbs are counted as operators in both sets of decompositions. In Appendix A2, nominalized verbs are transformed to the infinitive form of the verb. In Appendix A1 the occurrence of ASSERTION or FRAGMENT implies the presence of an operator. Most nominalized forms of verbs in this sublanguage occur under a paraphrastic verb operator (had 2 transfusions) or as a noun FRAGMENT derived from such a construction. Counting ASSERTIONs and FRAGMENTs gives the same count as if the nominalized verbs were decomposed as far as main clauses of sentences are concerned. In the case of adjuncts that contain a nominalized verb (on transfer to University Hospital, STDSM 2.1.1) the operator status of the verb is made explicit in the decomposition.

Results

Tabulations were obtained for the following variables:

WORDS--Number of words in sentence

OPS----Total number of operators

W/O----Ratio of words to operators

ADJ----Number of operators in "adjuncts" (local modifiers)

NAJ----Number of operators not in "adjuncts"

DMAX---Maximum depth of nesting including "adjuncts"

DNAJ---Maximum depth of nesting excluding "adjuncts"

O*D----Product of DMAX and OPS, the "overall complexity"

The data for the two types of textual material are shown in Table 1 (for the discharge summary) and Table 2 (for the theoretical texts). The averages over all sentences of the two types are shown in Table 3, which also shows the ratios between the measures for the two data sets. As we hypothesized, the measures for the theoretical material are significantly larger than those for the discharge summaries.

We included the number of words in the sentence as an intuitive indicator of the amount and complexity of the information in a sentence. We were interested to include this measure (which is the only measure obtainable in many installations) in order to compare it to our more theoretically-motivated measures.

The number of operators is one measure of the amount of information in a sentence. It should be noted that sentences in the theoretical texts, in a number of cases, refer to structures in previous sentences. This referencing mechanism can be thought of as introducing prior information into the later sentence. As we have not counted these referred operators as part of the referring sentence, the operator count measures the new information contributed to the ongoing

Table 1: Complexity Measures for FACT Sentences

SENT	SENT WDS	OPS	W/O	ADJ	NAJ	DMAX	DNAJ	O*D
		OPERATORS				DEPTH		
HIDSM1.1.1	17	3	5.66	0	3	2	2	6
HIDSM1.1.2	16	4	4.00	1	3	3	3	12
HIDSM1.1.3	20	6	3.33	0	6	4	4	24
HIDSM1.1.4	5	2	2.50	0	2	2	2	4
HIDSM1.1.8	11	3	3.66	2	1	3	2	9
HIDSM1.1.11	5	2	2.50	0	2	2	2	4
HIDSM1.1.12	9	3	3.00	0	3	2	2	6
HIDSM1.1.13	17	21	0.80	12	9	4	3	84
HIDSM1.1.16	6	1	6.00	0	1	1	1	1
HIDSM1.1.18	6	1	6.00	0	1	1	1	1
HIDSM1.1.19	5	2	2.50	0	2	2	2	4
HIDSM1.1.22	4	2	2.00	0	2	2	2	4
HIDSM2.1.2	14	2	7.00	1	1	2	2	4
HIDSM2.1.5	13	3	4.33	0	3	2	2	6
HIDSM2.1.7	10	2	5.00	1	1	2	2	4
HIDSM2.1.8	16	6	2.66	2	4	4	3	24
HIDSM2.1.13	23	6	3.83	0	6	4	4	24
HIDSM2.1.14	10	1	10.00	0	1	1	1	1
HIDSM2.1.15	4	2	2.00	0	2	2	2	4
HIDSM2.1.16	3	1	3.00	0	1	1	1	1
HIDSM2.1.17	14	4	3.50	0	4	2	2	8
HIDSM2.1.18	3	2	1.50	0	2	2	2	4
HIDSM2.1.21	10	2	5.00	1	1	2	2	4
HIDSM2.1.23	8	1	8.00	0	1	1	1	1
HIDSM3.1.2	12	3	4.00	0	3	2	2	6
HIDSM3.1.3	10	4	2.50	0	4	2	2	8
HIDSM3.1.4	7	1	7.00	0	1	1	1	1
HIDSM3.1.6	8	3	2.66	0	3	2	2	6
HIDSM3.1.11	3	2	1.50	0	2	2	2	4
HIDSM3.1.12	3	2	1.50	0	2	2	2	4
HIDSM3.1.13	9	3	3.00	0	3	2	2	6
HIDSM3.1.14	7	9	0.77	0	9	3	3	27
HIDSM3.1.17	5	2	2.50	0	2	2	2	4
EXDSM1.1.1	6	4	1.50	3	1	3	2	12
EXDSM1.1.3	2	1	2.00	0	1	1	1	1
EXDSM1.1.5	4	1	4.00	0	1	1	1	1
EXDSM1.1.6	4	2	2.00	0	2	2	2	4
EXDSM1.1.7	4	1	4.00	0	1	1	1	1
EXDSM1.1.9	2	1	2.00	0	1	1	1	1
EXDSM1.1.11	5	4	1.25	0	4	3	3	12
EXDSM1.1.13	3	1	3.00	0	1	1	1	1
EXDSM1.1.15	5	3	1.66	0	3	3	3	9
EXDSM1.1.16	5	1	5.00	0	1	1	1	1
EXDSM1.1.17	2	1	2.00	0	1	1	1	1
EXDSM1.1.19	10	4	2.50	3	1	3	2	12
EXDSM1.1.20	3	1	3.00	0	1	1	1	1
EXDSM1.1.21	5	2	2.50	0	2	2	2	4
EXDSM1.1.22	3	1	3.00	0	1	1	1	1
EXDSM1.1.23	4	2	2.00	1	1	2	2	4

Table 1, cont.: Complexity Measures for FACT Sentences

SENT	SENT WDS	OPERATORS OPS	W/O	ADJ	NAJ	DEPTH DMAX	DNAJ	O*D
EXDSM2.1.4	12	5	2.40	2	3	4	3	20
EXDSM2.1.5	5	1	5.00	0	1	1	1	1
EXDSM2.1.6	7	1	7.00	0	1	1	1	1
EXDSM2.1.7	11	4	2.75	3	1	4	2	16
EXDSM2.1.8	1	1	1.00	0	1	1	1	1
EXDSM2.1.10	7	4	1.75	0	4	3	3	12
EXDSM2.1.14	11	6	1.83	0	6	4	4	24
EXDSM2.1.15	6	1	6.00	0	1	1	1	1
EXDSM2.1.20	4	2	2.00	0	2	2	2	4
EXDSM2.1.22	4	1	4.00	0	1	1	1	1
EXDSM3.1.2	3	1	3.00	0	1	1	1	1
EXDSM3.1.3	4	1	4.00	0	1	1	1	1
EXDSM3.1.4	2	1	2.00	0	1	1	1	1
EXDSM3.1.5	2	1	2.00	0	1	1	1	1
EXDSM3.1.8	7	5	1.40	0	5	3	3	15
EXDSM3.1.9	5	5	1.00	0	5	3	3	15
EXDSM3.1.13	5	1	5.00	0	1	1	1	1
EXDSM3.1.16	5	2	2.50	0	2	2	2	4
EXDSM3.1.17	12	11	1.09	0	11	3	3	33
EXDSM3.1.19	4	2	2.00	0	2	2	2	4
EXDSM3.1.23	3	2	1.50	0	2	2	2	4
EXDSM3.1.24	4	3	1.33	0	3	2	2	6
LDDSM1.1.1	11	9	1.22	0	9	4	4	36
LDDSM1.1.2	19	8	2.37	0	8	2	2	16
LDDSM2.1.4	7	2	3.50	0	2	2	2	4
LDDSM2.1.5	3	2	1.50	0	2	2	2	4
LDDSM2.1.6	8	2	4.00	0	2	2	2	4
LDDSM2.1.7	4	2	2.00	0	2	2	2	4
LDDSM2.1.11	2	2	1.00	0	2	2	2	4
LDDSM2.1.12	2	2	1.00	0	2	2	2	4
LDDSM2.1.13	2	2	1.00	0	2	2	2	4
LDDSM2.1.14	7	3	2.33	0	3	2	2	6
LDDSM3.1.2	3	1	3.00	0	1	1	1	1
LDDSM3.1.3	3	1	3.00	0	1	1	1	1
LDDSM3.1.4	2	1	2.00	0	1	1	1	1
LDDSM3.1.5	2	1	2.00	0	1	1	1	1
LDDSM3.1.6	10	4	2.50	0	4	2	2	8
LDDSM3.1.8	16	7	2.28	0	7	2	2	14
LDDSM3.1.9	7	1	7.00	0	1	1	1	1
LDDSM3.1.10	2	1	2.00	0	1	1	1	1
LDDSM3.1.11	2	1	2.00	0	1	1	1	1
LDDSM3.1.12	20	10	2.00	0	10	3	3	30
LDDSM3.1.15	10	5	2.00	0	5	3	3	15
LDDSM3.1.18	9	4	2.25	0	4	3	3	12
LDDSM3.1.19	2	1	2.00	0	1	1	1	1
LDDSM3.1.20	5	1	5.00	0	1	1	1	1
LDDSM3.1.22	9	3	3.00	0	3	2	2	6
IMDSM2.1.1	2	2	1.00	0	2	2	2	4
CODSM1.1.3	11	1	11.00	0	1	1	1	1

Table 1, cont.: Complexity Measures for FACT Sentences

SENT	SENT WDS	OPERATORS OPS	W/O	ADJ	NAJ	DEPTH DMAX	DNAJ	O*D
CODSM1.1.5	9	6	1.50	5	1	3	2	18
CODSM1.1.6	9	3	3.00	1	2	3	3	9
CODSM2.1.8	9	3	3.00	2	1	3	2	9
STDSM1.1.1	15	1	15.00	0	1	1	1	1
STDSM1.1.2	4	1	4.00	0	1	1	1	1
STDSM2.1.1	11	2	5.50	1	1	2	2	4
STDSM3.1.1	3	1	3.00	0	1	1	1	1
PLDSM1.1.1	2	2	1.00	0	2	2	2	4
PLDSM2.1.1	2	2	1.00	0	2	2	2	4
RXDSM2.1.1	2	1	2.00	0	1	1	1	1
RXDSM3.1.1	2	2	1.00	0	2	2	2	4
RADSM2.1.1	2	2	1.00	0	2	2	2	4
RADSM3.1.1	2	2	1.00	0	2	2	2	4
DODSM2.1.1	2	1	2.00	0	1	1	1	1
DODSM3.1.1	3	1	3.00	0	1	1	1	1

Table 2: Complexity Measures for Theoretical Sentences

SENT	SENT WDS	OPERATORS OPS	W/O	ADJ	NAJ	DEPTH DMAX	DNAJ	O*D
UBS1	12	6	2.00	4	2	5	2	30
UBS2	9	3	3.00	1	2	3	2	9
UBS3	22	9	2.44	2	7	5	4	45
UBS4	13	11	1.18	0	11	5	5	55
UBS5	10	3	3.33	1	2	3	2	9
UBS6	19	6	3.16	0	6	4	4	24
UBS7	14	5	2.80	1	4	3	3	15
UBS8	24	2	12.00	0	2	2	2	4
UBS9	17	5	3.40	2	3	5	3	25
UBS10	18	5	3.60	2	3	3	2	15
UBS11	32	12	2.66	1	11	7	7	84
UBS12	20	3	6.66	2	1	3	1	9
UBS13	39	12	3.25	1	11	6	6	72
UBS14	22	6	3.66	2	4	4	3	24
UBS15	15	5	3.00	2	3	4	3	20
UBS16	19	5	3.80	0	5	3	3	15
UBKG2.4.7	9	3	3.00	0	3	2	2	6
UBKG2.4.8	18	5	3.60	4	1	3	1	15
UBKG2.4.9	11	1	11.00	0	1	1	1	1
UBKG2.5.1	13	4	3.25	1	3	3	2	12
UBKG2.5.2	23	6	3.83	1	5	3	3	18
UBKG2.5.3	23	4	5.75	3	1	3	1	12
UBKG2.5.4	12	6	2.00	1	5	4	3	24
UBKG2.6.1	22	5	4.40	4	1	4	1	20
UBKG2.6.2	20	5	4.00	0	5	3	3	15
UBKG2.6.3	18	6	3.00	0	6	4	4	24
UBKG2.7.1	29	6	4.83	5	1	3	1	18
UBKG2.7.2	12	4	3.00	3	1	4	1	16
UBKG2.7.3	29	17	1.70	2	15	6	5	102
UBKG2.7.4	19	11	1.72	0	11	4	4	44
UBKG2.7.5	22	4	5.50	2	2	3	2	12
UBKG2.7.6	25	7	3.57	2	5	3	3	21
UBKG2.8.1	19	9	2.11	6	3	4	2	36
UBKG2.8.2	12	3	4.00	2	1	3	1	9
UBKG2.8.3	23	9	2.55	3	6	6	4	54
UBKG2.8.4	26	8	3.25	1	7	5	4	40
UBKG2.9.1	14	5	2.80	0	5	3	3	15
UBKG2.9.2	14	4	3.50	0	4	3	3	12
UBKG2.9.3	18	7	2.57	0	7	4	4	28
UBKG2.9.4	31	5	6.20	0	5	3	3	15
UBKG2.9.5	28	9	3.11	0	9	5	5	45
UBKG2.9.6	12	3	4.00	2	1	3	1	9
UBKG2.9.7	31	5	6.20	4	1	4	1	20
UBKG2.9.8	25	5	5.00	1	4	3	3	15
UBKG2.9.9	24	4	6.00	1	3	4	3	16
UBKG2.12.3	15	6	2.50	1	5	4	3	24
UBKG2.14.2	28	7	4.00	1	6	4	3	28
UBKG3.1.7	23	11	2.09	7	4	5	4	55
UBKG3.3.5	25	13	1.92	0	13	6	6	78

Table 2, cont.: Complexity Measures for Theoretical Sentences

SENT	SENT WDS	OPS	OPERATORS W/O	ADJ	NAJ	DEPTH DMAX	DNAJ	O*D
UBKG4.3.3	20	6	3.33	2	4	4	3	24
UBKG4.4.3	22	5	4.40	2	3	4	3	20
UBKG4.6.1	35	7	5.00	0	7	5	5	35
UBKG4.6.3	26	9	2.88	0	9	5	5	45
UBKG4.6.4	31	9	3.44	1	8	4	4	36
UBKG4.6.6	22	5	4.40	1	4	5	4	25
UBKG4.6.7	37	11	3.36	1	10	5	5	55
UBKG4.7.1	29	6	4.83	0	6	5	5	30
UBKG4.7.3	27	11	2.45	0	11	7	7	77
UBKG7.2.3	19	7	2.71	1	6	4	4	28
UBKG7.2.4	39	6	6.50	0	6	4	4	24
UBKG8.4.3	19	8	2.37	1	7	5	5	40
UBKGFIG.16.2	36	10	3.60	0	10	6	6	60

Table 3: Comparison of Measures for Contrasting Types of Text

1. FACT-Report Texts (T1).

*****TOTALS

Total Number of Sentences	113
Average Number of Words	6.73
Average Number of Operators	2.75
Average Number of Words per Operator	3.06
Average Number of Operators in Adjuncts	0.38
Average Number of Operators not in Adjuncts	2.37
Average Depth of Nesting	1.92
Average Depth of Nesting excluding Adjuncts	1.86
Average operator-nesting product	7.02

2. Theoretical Texts (T2).

*****TOTALS

Total Number of Sentences	62
Average Number of Words	21.61
Average Number of Operators	6.53
Average Number of Words per Operator	3.79
Average Number of Operators in Adjuncts	1.40
Average Number of Operators not in Adjuncts	5.12
Average Depth of Nesting	4.03
Average Depth of Nesting excluding Adjuncts	3.25
Average operator-nesting product	29.24

3. T2/T1 Ratios.

*****TOTALS

	THEORY TEXT	FACT TEXT	RATIO
Total Number of Sentences	62	113	-
Average Number of Words	21.61	6.73	3.21
Average Number of Operators	6.53	2.75	2.37
Average Number of Words per Operator	3.79	3.06	1.24
Average Number of Operators in Adjuncts	1.40	0.38	3.68
Average Number of Operators not in Adjuncts	5.12	2.37	2.16
Average Depth of Nesting	4.03	1.92	2.10
Average Depth of Nesting excluding Adjuncts	3.25	1.86	1.75
Average operator-nesting product	29.24	7.02	4.17

discourse by the referring sentence. By including these referred operators we would obtain a measure of the total information included in the sentence. For the total operator counts, the ratio between the theoretical texts and the factual report texts is 2.37.

The number of words per operator is a rough inverse measure of the density of information in a sentence. The two types of texts were roughly comparable in this variable (the ratio is 1.24), despite the great difference in types of textual material. Part of this similarity is an artifact of the processing of the theoretical texts, where long metatextual material (e.g., "It seems clear from the studies of BG..." in UBS 8) and multi-word phrases (e.g., "...integrated transport system..." in UBS 12) were analyzed as a single operator. On the other hand, there must exist a limit to the amount of packing and reduction that can be performed on information in the process of encoding it into sentences, and material at or near this limit would be expected to have the same density regardless of the complexity of the information. Greater complexity could only arise in such a case from a longer sentence.

The number of operators in adjuncts is a measure of how complicated the local modifiers are in a sentence. The number of operators not in adjuncts is a measure of how complicated the logical structure of the sentence as a whole is. The ratios in both variables are similar (3.68 for the operators in adjuncts, 2.16 for the operators not in adjuncts), indicating that the theoretical material is more complex both in the local modifiers and in the overall logical structure.

One feature of complexity is depth of nesting. For our sentences, we measured both the maximum depth of nesting, and the maximum depth excluding local modifiers. In the factual material, the two measures are essentially the same (1.92 vs. 1.86). This accurately reflects the shallow structure of local modifiers in the factual material. In the theoretical texts, on the other hand, there is a more pronounced difference (4.03 vs. 3.31), indicating greater complexity of modifier structures. In fact, the ratio (not shown in the Table) between the depths of nesting in adjuncts alone would be 13.0. The ratios for the given depth-of-nesting indices are 2.10 and 1.75.

As expected, the product of the number of operators and depth of nesting enhanced the contrast between sentences with regard to complexity of information. The ratio of the averages for the two types of text is 4.17.

Looking at Table 3 as a whole, it is clear that the differences in information complexity between two radically different types of textual material can be expressed quantitatively based on the structural properties of the sentences. Indeed, with the exception of the inverse density, the various measures are roughly in the same ratio, about 2.75.

Having established the validity of several measures on contrasting types of text, it was now possible to study the performance of these measures on sentences within a single type of textual material.

Table 4 shows the high and low extremes of sentence groups from the theoretical articles based on values of the overall complexity (O*D). In order that the reader may verify the discriminatory effect of grouping by complexity values, Table 5 shows the original text of all sentences from the high groups, and Table 6 shows the text from the low groups, each group in turn divided into subgroups.

Discussion

With regard to the theoretical materials in Table 5, two anomalous sentences should be noted.

1) In Table 5, group A2, the presence of UBS 4 requires some explanation, since the other sentences in this group seem manifestly more complicated. The reason for UBS 4's presence is the existence of the logical connector "In either case..." within the sentence. This connector requires expansion as noted in Appendix A2, i.e., "in either case , Z --> in case X, then Z, or in case Y, then Z." Here X and Y are references to sections of the previous sentence. In UBS 4, Z contains four operators, which must be repeated for the two cases, giving a higher operator count and a correspondingly higher overall complexity.

Table 4: High and Low Complexity Groups of T2 Sentences.

A. High O*D

A1. O*D > 60

Sentence #	O*D	OPS
UBKG 2.7.3	102	17
UBS 11	84	12
UBKG 3.35	78	13
UBKG 4.7.3	77	11
UBS 13	72	12
Fig. 16.2	60	10

A2. 50 < O*D < 59

UBS 4	55	11
UBKG 3.1.7	55	11
UBKG 4.6.7	55	11
UBKG 2.8.3	54	9

A3. 40 < O*D < 49

UBS 3	45	9
UBKG 2.7.4	44	11
UBKG 2.9.5	45	9
UBKG 4.6.3	45	9
UBKG 2.8.4	40	8
UBKG 8.4.3	40	8

B. Low O*D

B1. O*D = 12

Sentence #	O*D	OPS
UBKG 2.5.1	12	4
UBKG 2.5.3	12	4
UBKG 2.7.5	12	4
UBKG 2.9.2	12	4

B2. O*D = 9

UBS 2	9	3
UBS 5	9	3
UBS 12	9	3
UBKG2.8.2	9	3
UBKG 2.9.6	9	3

B3. O*D < 6

UBKG 2.4.7	6	3
UBS 8	4	2
UBKG 2.4.9	1	1

Table 5: High Complexity Groups

A1. O*D ≥ 60

U BKG 2.7.3: Much of the cholesterol and bile acid secreted by the liver is reabsorbed in the intestine and again delivered to the liver for excretion, thus forming an enterohepatic circulation.

UBS 11: The receptors bind circulating lipoproteins that transport cholesterol in the bloodstream, thereby initiating a process by which the lipoproteins are taken up and degraded by cells yielding their cholesterol for cellular use.

U BKG 3.3.5: The mutant FH cells grow in tissue culture because they adapt to the unavailability of LDL-cholesterol by increasing HMG CoA reductase activity and cholesterol synthesis.

U BKG 4.7.3: These particles accumulate in plasma because their removal rate fails to increase in proportion to the increased production of chylomicrons and VLDL that occurs during cholesterol feeding.

UBS 13: Recent experiments show that the number of lipoprotein receptors in the liver, and hence the rate of removal of cholesterol from plasma, is under regulation and that the number of receptors can be increased by certain cholesterol lowering drugs.

U BKG FIG.15.2: This pharmacologic manipulation takes advantage of the knowledge that the receptors are normally under feedback regulation and that the number of receptors in the liver can be increased when the liver's demand for cholesterol is increased.

Table 5, cont.: High Complexity Groups.

A2. $50 \leq O*D \leq 59$

UBS 4: In either case, with VLDL breakdown, the cholesterol ester would arrive at LDL.

U BKG 3.1.7: The number of receptors in fibroblasts can be increased by certain hormones that stimulate cell growth, including insulin, thyroxine, and platelet-derived growth factor.

U BKG 4.5.7: This increased efficiency of LDL clearance, plus a 50 percent reduction in the synthetic rate of LDL produced by mevinolin, contributed to a remarkable 75 percent drop in the plasma level of LDL-cholesterol in the treated dogs.

U BKG 2.3.3: When dietary cholesterol is available, the liver uses that source of sterol, derived from the receptor-mediated uptake of chylomicron remnants, for lipoprotein synthesis.

Table 5, cont.: High Complexity Groups.

A3. $40 \leq O*D \leq 49$

UBS 3: The cholesterol ester formed in HDL may be transferred either to VLDL in exchange for triglyceride or, in part, directly to LDL.

U BKG 2.7.4: During each cycle a portion of the cholesterol and bile acid escapes reabsorption and is lost in the feces.

U BKG 2.9.5: As the size of the VLDL particle diminishes owing to its interaction with lipoprotein lipase, its density increases, and the particles are converted to intermediate density lipoproteins (IDL).

U BKG 4.5.3: By trapping bile acids in the intestine, colestipol causes the liver to convert more cholesterol to bile acids and hence increases the heptic demand for cholesterol.

U BKG 2.3.4: When dietary cholesterol is insufficient, the liver synthesizes its own cholesterol by increasing the activity of a rate-controlling enzyme, 3-hydroxy-3-methlglutarylcoenzyme A reductase (HMG CoA reductase).

U BKG 8.4.3: Epidemiologic evidence suggests that excessive dietary intake of fat, cholesterol, and calories is responsible for the increased blood cholesterol.

Table 6: Low Complexity Groups

B1. O*D = 12

U BKG 2.5.1: The depleted chylomicron is released from the capillary wall and reenters the circulation.

U BKG 2.5.3: The remnant (diameter, 300 to 800 A) is carried to the liver where it binds to receptors on the surface of hepatic cells.

U BKG 2.7.5: With the typical American diet, which is high in cholesterol, about 1100 mg of sterol is lost from the body each day.

U BKG 2.9.2: The VLDL particles interact with lipoprotein lipase in capillaries, releasing most of their triglycerides.

B2. O*D = 9

UBS 2: HDL are thought to accumulate cholesterol from peripheral tissues.

UBS 5: Alternatively, VLDL may be the exclusive source of LDL cholesterol.

UBS 12: The lipoprotein receptors are components of an integrated transport system that shuttles cholesterol continuously among intestine, liver, and extrahepatic tissues.

U BKG 2.8.2: These lipoproteins also contain cholesterol, which will be delivered to extrahepatic cells.

U BKG 2.9.5: The excess surface materials, mostly phospholipids and cholesterol, are transferred to HDL.

B3. O*D = 6

U BKG 2.4.7: As the triglyceride core is depleted, the chylomicron shrinks.

UBS 8: It seems clear from the studies of B & G that LDL is also the major vehicle for transport of cholesterol to peripheral tissues.

U BKG 2.4.9: A similar transfer reaction occurs in the endogenous lipoprotein transport system.

2) In Table 5, group A3, UBKG 2.7.4 is interesting because it has relatively a large number of operators (11) but a complexity value (44) that is less than a number of others with fewer operators (UBS 3 45/9, UBKG 2.8.3 54/9, UBKG 2.9.5 45/9, UBKG 4.6.3 45/9). The reason appears to be the number of coordinate conjunctions that lead to expansion. This sentence has two nouns that are under a quantifier and a coordinate conjunction, which are then the arguments of two verbs under a coordinate conjunction. Expansion of all coordinate conjunctions results in a total operator count of 11 but a relatively low depth of nesting (4), yielding a complexity value of 44. The fact that the operators are of the "zeroing" type is reflected in the low inverse density value (1.72). One might contrast this sentence with ones that have a lower operator count and a higher complexity value (UBKG 2.8.3 and UBKG FIG. 16.2). These two sentences have complexity/operator rations 54/9 and 60/10 respectively. Neither of these sentences has an expanded coordinate conjunction.

Although the factual material was too simple to meaningfully divide into groups, there are a number of sentences with relatively high overall complexities compared with others in the set, that merit discussion. Table 7 lists all sentences from the factual texts with complexities of 6 or more. Of these thirty-nine sentences, thirty-one contained either coordinate conjunctions, semicolons, or both. (They are marked in the CONJ column of Table 7 with an asterisk.) Although these repeated conjunctions were not themselves counted as operators, the operator count for these sentences was higher because of the

Table 7: High Complexity Sentences From FACT Texts

SENT	SENT WDS	O*D	CONJ
HIDSM1.1.1	17	6	*
HIDSM1.1.2	16	12	*
HIDSM1.1.3	20	24	*
HIDSM1.1.8	11	9	
HIDSM1.1.12	9	6	
HIDSM1.1.13	17	84	*
HIDSM2.1.5	13	6	*
HIDSM2.1.8	16	24	*
HIDSM2.1.13	23	24	*
HIDSM2.1.17	14	8	*
HIDSM3.1.2	12	6	*
HIDSM3.1.3	10	8	*
HIDSM3.1.6	8	6	*
HIDSM3.1.13	9	6	*
HIDSM3.1.14	7	27	*
EXDSM1.1.1	6	12	*
EXDSM1.1.11	5	12	
EXDSM1.1.15	5	9	*
EXDSM1.1.19	10	12	*
EXDSM2.1.4	12	20	
EXDSM2.1.7	11	16	
EXDSM2.1.10	7	12	
EXDSM2.1.14	11	24	*
EXDSM3.1.8	7	15	*
EXDSM3.1.9	5	15	*
EXDSM3.1.17	12	33	*
EXDSM3.1.24	4	6	*
LDDSM1.1.1	11	36	*
LDDSM1.1.2	19	16	*
LDDSM2.1.14	7	6	*
LDDSM3.1.6	10	8	*
LDDSM3.1.8	16	14	*
LDDSM3.1.12	20	30	*
LDDSM3.1.15	10	15	*
LDDSM3.1.18	9	12	*
LDDSM3.1.22	9	6	*
CODSM1.1.5	9	18	*
CODSM1.1.6	9	9	
CODSM2.1.8	9	9	

expansion of zeroed material. In these sentences, the inverse density is quite low, reflecting the presence of the zeroed elements.

While use of the operator-argument representation provides a uniform basis for measuring structural properties, the measures based on counting the operators, without further distinction as to their type, is admittedly crude. Future refinements of the measures should take account of (at least) the differences between

(a) a sublanguage verb; this operator carries a whole proposition in the "object language" and should have more weight than (b) or (c);

(b) a sentence operator or "meta" verb (i.e. a verb like know, believe, that occurs with an embedded subject or object); this operator should carry perhaps half the weight of a full proposition;

(c) a conjunction or conjunctional verb (e.g. indicates, leads to); these operators provide relational information among propositions. Their weight should be investigated.

It should be noted that the design of the investigation was modified in the course of the study. We had initially hypothesized that the relevant data for developing measures would be the semantic representation of texts (in LSP terms, the sublanguage information-formats [18], of texts). After some study it was clear that only qualitative comparisons could be made with regard to the properties of information formats for different types of texts within the same sublanguage; quantitative results required the development of a uniform informational representation that was independent of the type of textual material, and preferably also of the semantic domain of the texts. Thus, at a certain time, well into the project period, we refocused the inquiry to deal exclusively with the regularized English (as opposed to sublanguage) analysis of the texts.

The most interesting of the qualitative results from the early period of the investigation was the verification of the hypothesis that the semantic categories appearing as distinguished nodes (or "column heads") in the information formats of fact-level texts in a given sublanguage were present again in the information formats of generalized (survey-type) fact reports, along with new categories, most of which were relational categories, i.e. contained the words that carried specific relations among the old categories. The new categories were (not implausibly) for the most part relations of subsetting, frequency, and correlation.

Conclusions

In this study, we have demonstrated that a linguistically-based analysis of the structural properties of text can lead to quantitative measures of the amount and complexity of the information stored in the text. Types of texts which clearly differ in these properties turn out to have correspondingly contrastive measures. Furthermore, within a single type of text, groups of sentences graded according to one measure (the overall complexity) appear to reflect perceived differences in complexity of information.

These results are a first step toward defining and measuring important properties of textual material that have hitherto been only qualitatively perceived and described. Many practical and theoretical questions have to be answered before we can implement a quantitative system of some generality.

The first issue is finding a common representation that can be obtained by computer from a wide variety of texts. In this study, we worked out the correspondences between different representations so that valid comparisons could be made. For future work, it would be desirable to map all representations to a "lowest-common-denominator" informational form. This is not an elementary task, but it is within reach of current techniques.

Further study is also required to refine the measures with regard to different types of operators that have been counted, to give different weight to the amount of information contributed by each type of operator. For example, more weight could be given to operators that bring new propositions into the discourse, as opposed to operators that connect existing propositions.

In addition, it would be desirable to account for the role of reference in the total information content of an individual sentence, a project that was beyond the scope of the present investigation. With such an accounting for the weight of total information in each sentence of a discourse, it would be possible to map the informational complexity of a discourse as a function of sentence-position in a paragraph, or progression through a paragraph or section.

With regard to applications, a measurement of information in a discourse could be of benefit to systems for indexing or abstracting from free text. Selection of terms for indexing could take account of the complexities of the sentences into which that term entered. Retrievals of text passages in a text-retrieval system could be weighted by the measures of amount, density and complexity of the information in each candidate passage in a list of possible retrievals.

More generally, situations arise where it is desired to grade text as to level of difficulty. Examples are training manuals and graded reading materials for use in schools. Current applications use *ad hoc* measures whose relation to actual complexity of information are poorly defined.

In another area, the designers of text-derived database systems need information about the maximum depth of nesting and overall complexity to be expected in texts that will be mapped into the database.

Finally, continued research in this area can help to define more precisely the properties that distinguish textual information from purely quantitative data, and contribute to a precise definition of information that takes account of properties beyond those studied in the mathematical theory of information.

References

1. Shannon, C. (1949), The Mathematical Theory of Communication, University of Illinois Press.

2. Harris, Z. (1957), Co-occurrence and Transformation in Linguistic Structure, Presidential address, Linguistic Society of America, 1955, Language 33:3, 283-340.

3. _____ (1964), The Elementary Transformations, Transformations and Discourse Analysis Papers (T.D.A.P.) 54, University of Pennsylvania. Reprinted in part in Z.S. Harris, Papers in Structural and Transformational Linguistics, D. Reidel, Dordrecht, Holland, 1970.

4. _____ (1969), The Two Systems of Grammar: Report and Paraphrase, T.D.A.P. 79. Reprinted in Papers in Structural and Transformational Linguistics, D. Reidel, Dordrecht, Holland, 1970.

5. _____ (1976), A Theory of Language Structure, American Philosophical Quarterly 13, 237-255.

6. _____ (1982), A Grammar of English on Mathematical Principles, Wiley Interscience, New York.

7. Sager, N. (1976), Evaluation of Automated Natural Language Processing in the Further Development of Science Information Retrieval, String Program Reports 10, Linguistic String Project, New

York University. (Final Report to the Division of Information Science and Technology, National Science Foundation.)

8. Salton, G., and McGill, M.J. (1983), Introduction to Information Retrieval, McGraw Hill.

9. Winograd, T. (1983), Language as a Cognitive Process, Addison-Wesley, Reading, Mass.

10. Kittredge, R., and Lehrberger, J. (1982), Sublanguage: Studies of Language in Restricted Semantic Domains, Walter de Gruyter, Berlin.

11. Sager, N. (1981), Natural Language Information Processing, Addison-Wesley, Reading, Mass.

12. Joshi, A., Webber, B.L., and Sag, I. (1981), Elements of Discourse Understanding, Cambridge University Press.

13. Walter, D.E. (1978), Understanding Spoken Language, Elsevier/North-Holland.

14. Plath, W. (1976), REQUEST: A Natural Language Question-Answering System, IBM Journal of Research and Development 20:4, 326-335.

15. Petrick, S. (1973), Transformational Analysis, in Natural Language Understanding (R. Rustin, ed.), Algorithmics Press.

16. _____ (1981), Field-testing the Transformational Question-Answering (TQA) System, Proc 19th Ann ACL.

17. Hobbs, J., and Grishman, R. (1976), The Automatic Transformational Analysis of English Sentences: An Implementation, International Journal of Computer Mathematics, Section A, Vol. 5, pp. 267-283.

18. Sager, N. (1978), Natural Language Information Formatting: The Automatic Conversion of Texts to a Structured Data Base, in Advances in Computers 17 (M.C. Yovits, ed.), 89-162, Academic Press, New York.

19. Grishman, R., and Hirschman, L. (1978), Question Answering from Natural Language Medical Data Bases, Artificial Intelligence 11, 25-43.

20. Marsh, E. (1983), Utilizing Domain-Specific Information for Processing Compact Text. To appear in Proceedings of the Conference on Applied Natural Language Processing, Santa Monica, CA, February 1983.

The Implementation of English Transformations

Elaine Marsh, Lynette Hirschman, and Carol Friedman

The English decomposition transformations are executed on the output trees from the parsing component. Their effect is to reduce, paraphrastically, the variety of syntactic structures found in the parse tree; i.e. they regularize the syntax of the sentence without changing its informational content. The transformations which regularize the parses fall into several classes:

(1) conjunction expansion

(2) creation of assertion from verb-containing string

(3) reduction to canonical forms

(4) rearrangement transformations

(5) relative clause and adjunct expansion transformations

(6) regularization to LXR form.

Conjunction expansion transformations include those transformations that expand embedded conjunctions into conjoined structures at a higher level. For example, a conjoined noun phrase in a subject will yield two conjoined assertions after expansion.

Transformations which create assertions from verb-containing strings convert reduced embedded sentences (in our system, strings such as NTOVO, THATS, NSVINGO, etc.) into full assertions. A node-attribute 'TFORM-ATT' is created pointing from the newly created assertion to a list which contains an attribute identified with the original reduced form of the assertion. Missing elements are filled out in the verb-containing strings, for example subjects are reconstructed. (See Figure 1.)

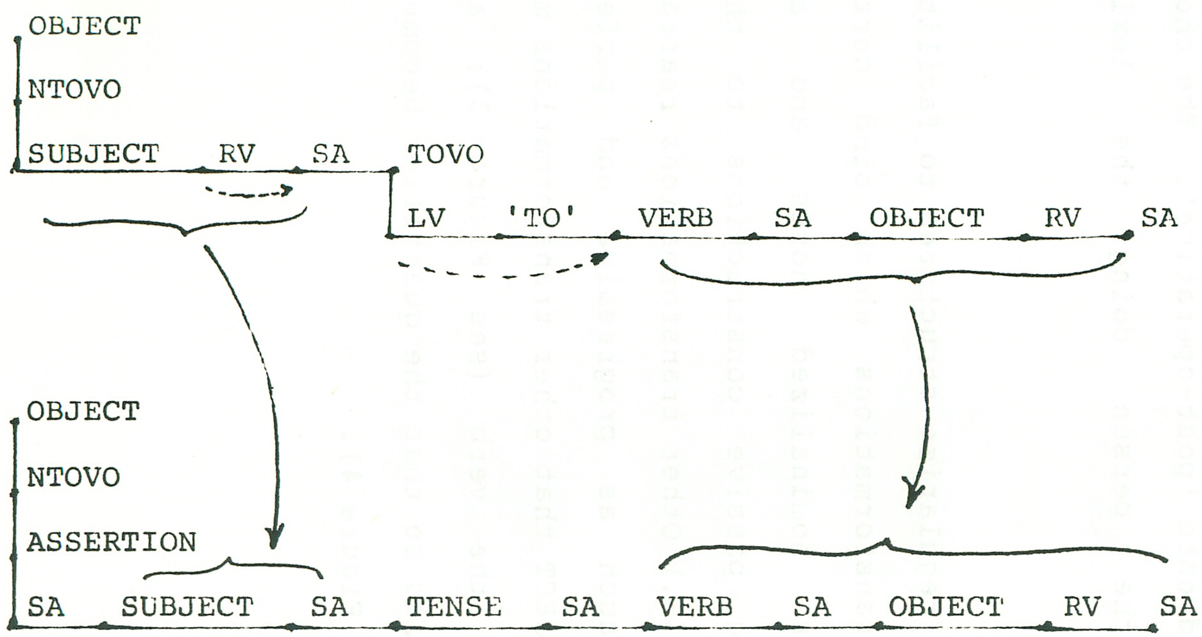
Transformations for reduction to canonical form involve:

(A) stripping affixes off words

(B) restoring abbreviations to full forms

(C) mapping varients of a form onto standard or canonical form.

Affixes are stripped off words and the words are reduced to their base forms. The information contained in the affixes, for example whether the word is a past tense of a verb or the plural of a noun, is identified by either setting a node-attribute pointing from the base form of the word, an atom, to a list containing attributes identifying the information, or by building a marker to the right of the core with that information contained in it. The former course is taken with tense information of verbs. For example, reducing 'went' to 'go' and setting a node attribute 'TENSE-ATT' pointing to a list containing the attribute 'PAST'. On the other hand, a marker 'PLURAL' is built for plural forms of nouns, e.g. hands --> hand plural. In either case, a word is replaced by its canonical form and the information contained in



RV is "added" to SA; LV is added to LV under VERB.

These elements are then used to make up an ASSERTION

Figure 1 . T-NTOVO Transformation

the affix is retained in a form appropriate for further use by the processing program.

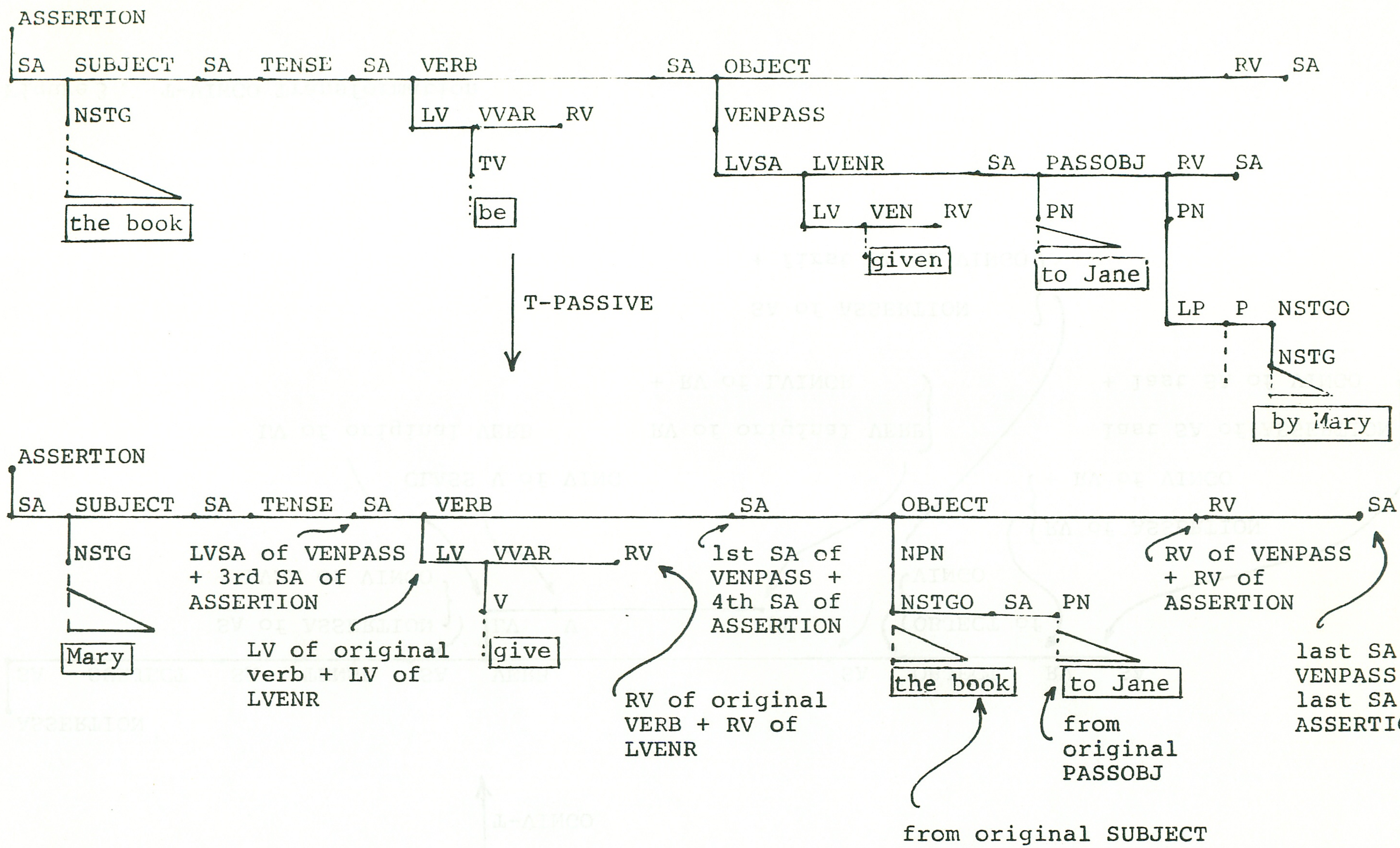
Abbreviations are replaced by their full forms using the class operation.

e.g., 'RR' ⇛ 'respiration rate'

Abbreviations are identified with the full forms by pointers among lexical entries in the dictionary.

Finally, variants are mapped onto their standard form. For example, 'postoperative' is mapped onto 'post-operative'. The choice of standard form is made by the person doing the lexical classification.

Rearrangement transformations regularize structures to facilitate further processing. They include transformations which bind certain prepositional phrases to their nominalized hosts and also transformations such as regularizing passive constructions to their active counterparts. (See Figure 2.) Other transformations rearrange strings containing compound tenses such as progressive and perfect, both regularizing the type of OBJECT that other transformations must deal with and removing affixes from the verb (see Figure 3); also quantifier expressions are regularized so that the quantifier becomes a modifier of the real head noun (see Figure 4).



The NEXT-ADJUNCT routine searches all possible adjunct slots (SA, RV) following the verb (given) for a potential agent PN (by Mary). If it is found, the NSTG is copied under SUBJECT and is removed from the RV or SA before these are mapped into the new ASSERTION.

Figure 2 . T-PASSIVE Transformation

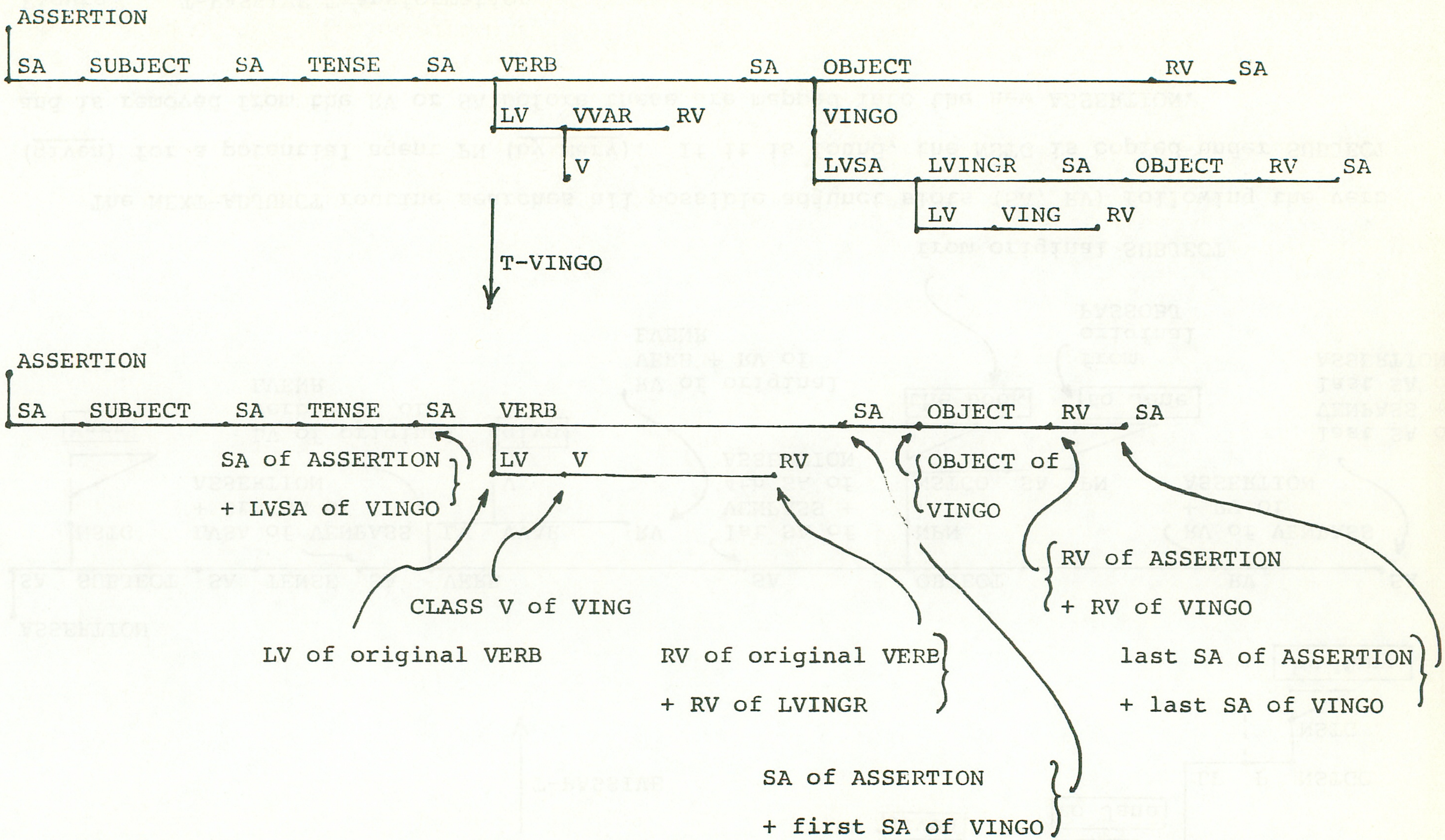


Figure 3 . T-VINGO Transformation

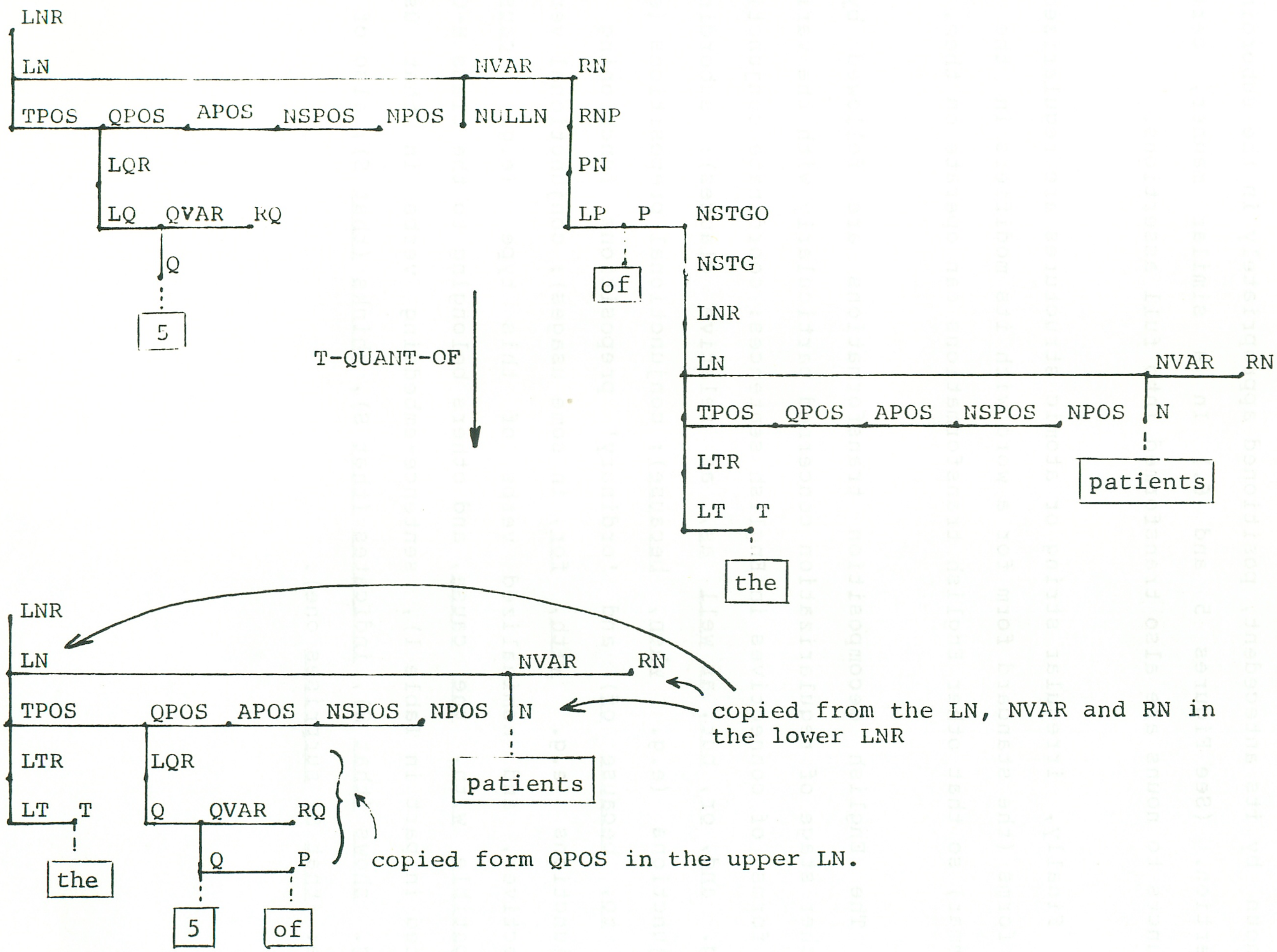


Figure 4. T-QUANT-OF Transformation

Relative clause and adjunct expansion transformations expand certain types of adjoined material into separate full subordinate assertions. Relative clauses are expanded by replacing the relative pronoun by its antecedent, positioned appropriately in the subordinate assertion. (See Figures 5 and 6.) In a similar manner, certain adjuncts to nouns are also transformed into full assertions.

Finally, irregular string or atomic structures are regularized to LXR forms (the standard form for a word with its modifiers in the LSP grammar) so that other English transformations can operate on them.

The English decomposition transformations are followed by a further stage of regularization concerned particularly with the variety of forms of connectives in English sentences: coordinate conjunctions (e.g. and, or, but, as well as, as relative clauses); subordinate conjunctions (e.g. when, because); conjunctional prepositions (e.g. due to, because of) and 'ordinary' prepositions functioning as conjunctions (e.g. with, for, in some usages); conjunctional verbs, adjectives, and nominalized verbs of this type (e.g. causes, compatible with, the cause, and others belonging to the class H-CONN listed in part in Table 1), sentence-embedding verbs in that usage (e.g. shows (that S), indicates (that S), thinks (that S); also of the type: (that S) surprises one).

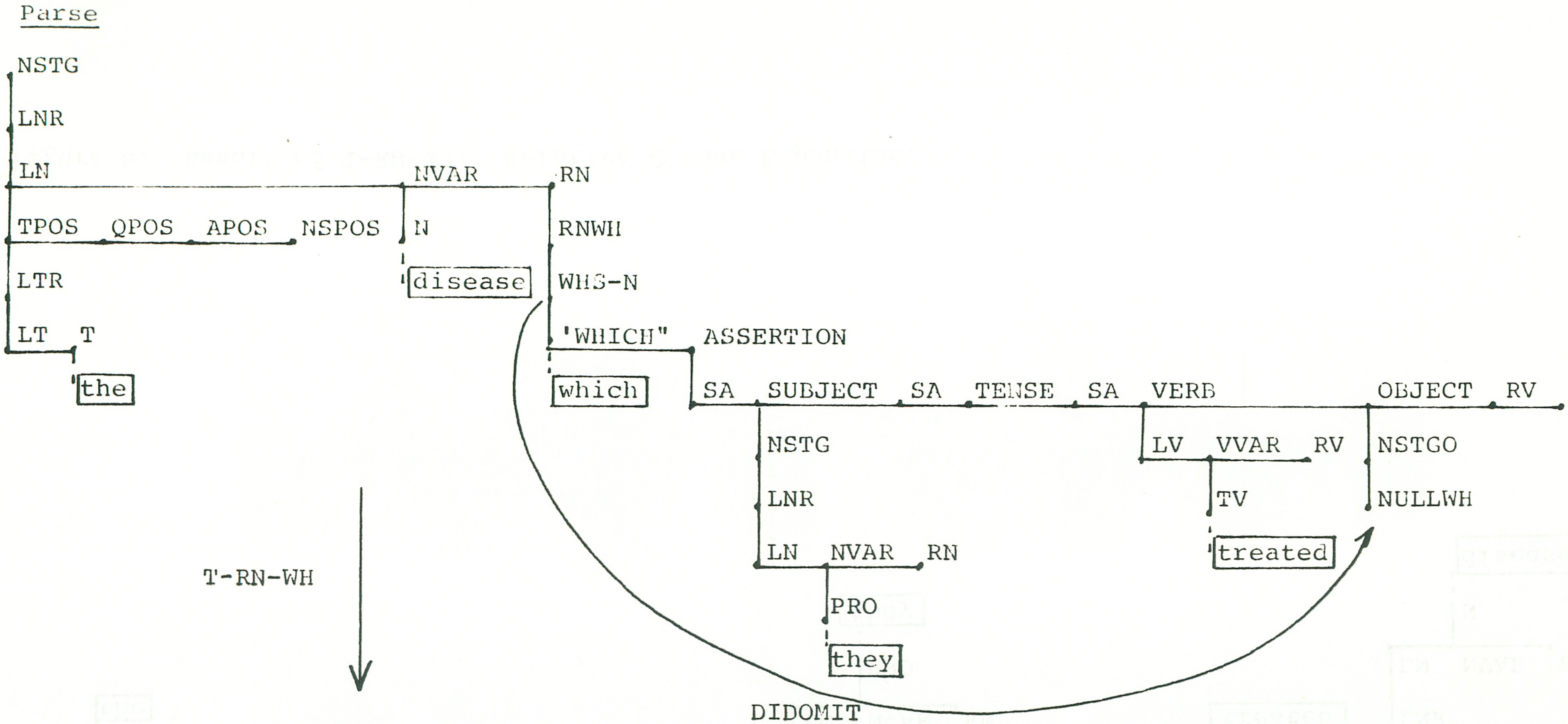


Figure 5. Relative Clause - Parse Tree

Transformation

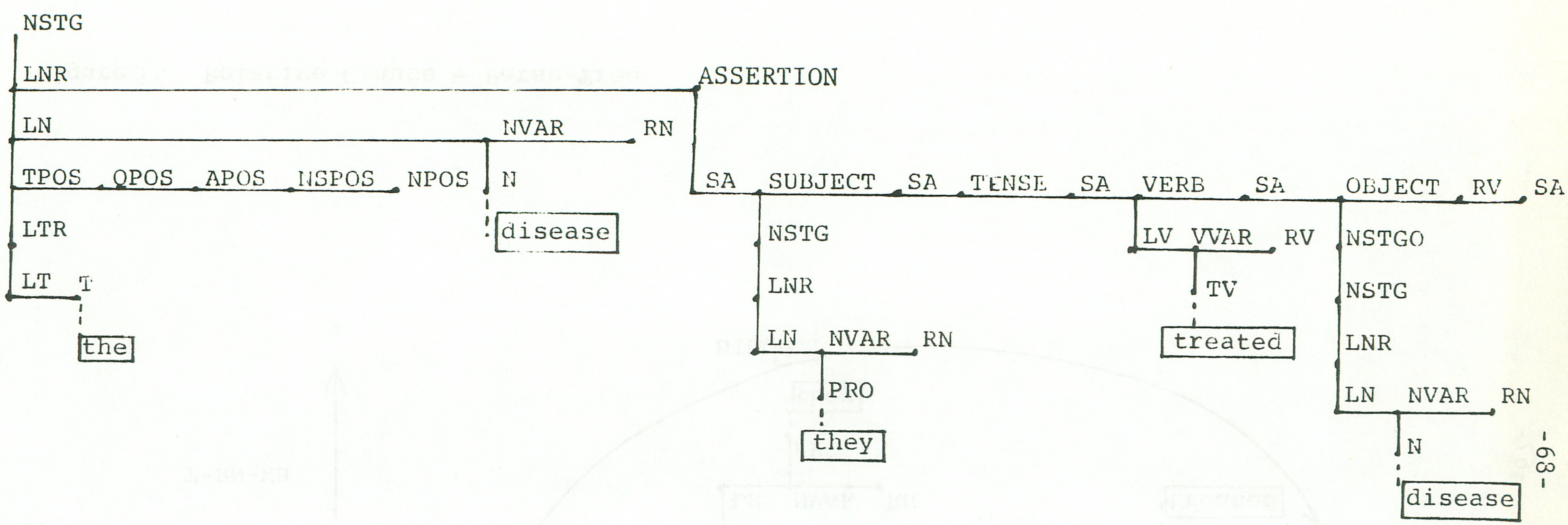


Figure 6. Result of T-RN-WH: Relative Clause Expansion.

TABLE 1

****

SUBCLASS H-CONN

****

V	ACCOMPANY
V	ADD
V	ASSOCIATE
N	ASSOCIATION
ADJ	ATYPICAL (OF)
V	AVOID
P	BECAUSE_OF
N	CAUSE
V	CAUSE
ADJ	COMMON
V	COMPARE
N	COMPARISON
ADJ	COMPATIBLE (WITH)
V	COMPLICATE
V	CONFIRM
N	CONNECTION
N	CONSEQUENCE
ADJ	CONSISTENT (WITH)
ADJ	CONTRIBUTORY
V	CONVERT
N	CORRELATE
V	CORRELATE
N	CORRELATION (WITH)
V	DAMAGE
V	DEMONSTRATE
V	DEVELOP
N	DEVELOPMENT
V	DICTATE
V	DOCUMENT
V	DOMINATE
P	DUE_TO
N	EFFECT
V	EFFECT
ADJ	ETIOLOGIC
N	ETIOLOGY
N	EXACERBATION
N	EXCEPTION
P	EXCEPT_FOR

V	EXPLAIN
P	FOR
V	FORCE
V	FOSTER
V	GUIDE
V	INDUCE
V	INVOLVE
N	INVOLVEMENT
P	IN_CONTRAST_TO
CS1	IN_THAT
V	LEAD
V	MEAN
N	OBJECTIVE
N	ORDER
N	ORIGIN
V	POINT
V	PRECEDE
V	PRECIPITATE
V	PRODUCE
V	PROVOKE
P	RE
V	REDUCE
N	REDUCTION
V	RELATE
N	RELATIONSHIP
ADJ	RELATIVE (TO)
V	RENDER

V	RESULT
V	REVEAL
P	SECONDARY_TO
V	SHOW
N	SOURCE
V	SUGGEST
ADJ	SUGGESTIVE
V	SUPERIMPOSE
V	SUPPORT
P	SUSPICIOUS_OF
N	SWITCH
V	SWITCH
N	TRANSITION
N	TRIGGER
V	TRIGGER
ADJ	TYPICAL (OF)
ADJ	UNDOCUMENTED
V	YIELD

What follows below is a detailed description of the transformations that carry out the transformational decomposition and further syntactic regularization in the LSP system.

```
* ********** ENGLISH DECOMPOSITION TRANSFORMATIONS **********
*
*
* T-EXPAND-TENSE
*       FIRST EXPANDS TENSE AND THEN THE VERB.
*       EXPANDS ANY CONJOINED TENSE OR VERB IN ASSERTION IN ORDER TO
*       SIMPLIFY VERB DE-TENSING TRANSFORMATIONS.
* T-COMPOUND-ADJ
*       TRANSFORMS REPEATED ADJECTIVES [WITHOUT CONJUNCTIONS] INTO RELATIVE
*       CLAUSES AND INSERTS THE REL-CL BEFORE THE VALUE OF RN, LEAVING
*       THE ADJECTIVE CLOSEST TO THE HOST AS LN ON ITS HOST.
*       E.G. 'AN ALERT, IRRITABLE CHILD' ==>
*             'AN IRRITABLE CHILD WHO BE ALERT'.
*       IF THE ADJECTIVE IN LAR1 IS INVOLVED IN A COMPUTED-ATTRIBUTE
*       CONSTRUCTION, E.G. 'TONIC' IN A 'CLONIC, TONIC TYPE OF CONVULSION',
*       IT WILL NOT BE EXPANDED INTO A RELATIVE CLAUSE.
*       T-COMPOUND-ADJ IS HOUSED IN ADJADJ AND WILL EXECUTE WHEN LN
*       IS SEQUENCED.  NEWLY CREATED RELATIVE CLAUSES WILL BE
*       TRANSFORMED WHEN RN IS TRANSFORMED.
* T-NO-X-OR-Y
*       CONVERTS AN 'OR' IN THE SCOPE OF A NEGATIVE INTO A COPY OF THE
*       NEGATIVE AND AN 'AND'.  THE STRUCTURE CAN THEN BE EXPANDED
*       BY T-EXPAND-TO-ASSERT TO GIVE THE SEMANTICALLY
*       EQUIVALENT CONJOINED ASSERTIONS.
*            E.G. NO HEART DISEASE OR KIDNEY DISEASE ==>
*                 NO HEART DISEASE AND NO KIDNEY DISEASE.
*       NOTE THAT THIS TRANSFORMATION REARRANGES THE LOGICAL
*       OPERATORS IN A PARSE TREE JUST PRIOR TO THE EXPANSION OF CONJUNCTIONS.
*       SPECIFICALLY,  NO (X OR Y) ==> NO X AND NO Y.
*       IN THE CASE OF 'NO X OF Y OR Z', THE CONNECTIVE 'OR' NEED ONLY BE
*       CHANGED TO 'AND', GIVING AFTER EXPANSION 'NO X OF Y AND NO X OF Z'.
*       IN THIS CASE, THE TRANSFORMATION MUST RECURSIVELY CHECK FOR OTHER
*       ORSTGS IN THE ADJUNCTS.
* T-EXPAND-TO-ASSERT
*       EXPANDS PRESENT-ELEMENT- AND KEEPS EXPANDING UPWARDS UNTIL IT REACHES
*       ASSERTION OR CENTER EXCEPT WHEN THE PRESENT-ELEMENT- IS AN LNR
*       THAT IS A CONJOINED NUMERICAL EXPRESSION,
*       E.G. 'BETWEEN 2% AND 4%' IS NOT EXPANDED.
* T-EXPAND-OBJ
*       EXPANDS CONJOINED ELEMENTS AND VERBS UNDER OBJECT , OBJBE,
*       AND PASSOBJ.
*       THIS TRANSFORMATION OPERATES RECURSIVELY FOR OBJECTS WHICH
*       THEMSELVES CONTAIN COMPLEX OBJECT STRINGS.  AGAIN THIS
*       SIMPLIFIES VERB DE-TENSING TRANSFORMATIONS, PARTICULARLY
*       THOSE WHICH INVOLVE MANIPULATION OF OBJECT STRINGS,
*       SUCH AS T-VENO WHICH "UNDOES" THE PERFECT TENSE.
*            E.G. (1) WE FOUND THE TEXTBOOKS IN THE LIBRARY AND THE ATLAS
*                           ON MY SHELF.  ==>
                      WE FOUND THE TEXTBOOKS IN THE LIBRARY AND WE FOUND THE
*                           ATLAS ON MY SHELF.
```

```
*              (2) THEY HAD EATEN DINNER AND GONE TO THE MOVIES.  ==>
*                  THEY HAD EATEN DINNER AND THEY HAD GONE TO THE MOVIES.
* T-EXPAND-LQR
*      THIS TRANSFORMATION DOES A PARTIAL EXPANSION (UP TO LNR OR QN OR NQ)
*      IN CONJOINED QUANTITY EXPRESSIONS.  THE EXPANSION IS ONLY DONE
*      UP TO LNR, QN, OR NQ LEVEL BECAUSE THESE CONJUNCTIONS
*      GENERALLY REPRESENT RANGES, NOT TWO SEPARATE ASSERTIONS.
*          E.G. 'HEMATOCRIT WAS 25 TO 34' IS NOT EQUIVALENT TO
*               'HEMATOCRIT WAS 25 TO HEMATOCRIT WAS 34'.
*      THE PARTIAL EXPANSION IS DONE SO THAT EACH NUMBER CARRIES ITS
*      UNIT:
*            E.G.  25 TO 34 MG. ==>
*                  25 MG. TO 34 MG..
*      THIS MAKES THE QUANTITY EXPRESSION MORE REGULAR.
* T-RNSUBJ MOVES RNSUBJ TO RN OF HOST
*      1. IN OBJECT IF SA OCCURS TO THE RIGHT OF OBJECT/PASSOBJ
*         WHICH HAS AN NSTG
*              E.G. HE DESCRIBED A PROCEDURE YESTERDAY WHICH WAS INTERESTING.
*                           ==>
*                   HE DESCRIBED A PROCEDURE WHICH WAS INTERESTING YESTERDAY.
*      2, IN SUBJECT OTHERWISE,
*              E.G. A PROCEDURE WAS DESCRIBED WHICH WAS INTERESTING. ==>
*                   A PROCEDURE WHICH WAS INTERESTING WAS DESCRIBED.
* T-NTOVO
*     CONVERTS NTOVO INTO A FULL ASSERTION IN OBJECT AND SETS
*     A NODE ATTRIBUTE TFORM-ATT POINTING FROM ASSERTION TO
*     LIST CONTAINING THE ATTRIBUTE TNTOVO.  THE NODE ATTRIBUTE
*     MARKS THE SOURCE OF THE ASSERTION.
*          E.G. THEY PERMITTED HER TO GO. ==>
*               THEY PERMITTED (HER GO).
* T-NSVINGO
*     CONVERTS NSVINGO IN SUBJECT OR OBJECT POSITION INTO A FULL EMBEDDED
*     ASSERTION.  IT ALSO SETS A NODE ATTRIBUTE TFORM-ATT POINTING
*     FROM NEW ASSERTION TO LIST CONTAINING THE ATTRIBUTE TNSVINGO.
*          E.G. JOHN'S GOING TO BOSTON DISTURBED MARY.  ==>
*               (JOHN GO TO BOSTON) DISTURBED MARY.
* T-SVINGO
*     CONVERTS SVINGO INTO A FULL ASSERTION.  IT ALSO SETS NODE ATTRIBUTE
*     TFORM-ATT POINTING FROM NEWLY BUILT ASSERTION TO A LIST CONTAINING
*     THE ATTRIBUTE TSVINGO.
*          E.G. THEY KEPT JOHN WAITING.  ==>
*               THEY KEPT (JOHN WAIT).
* T-NPVINGO
*     CONVERTS NPVINGO OBJECT INTO A FULL ASSERTION AND SETS NODE
*     ATTRIBUTE TFORM-ATT POINTING FROM ASSERTION TO LIST CONTAINING
*     ATTRIBUTE TNPVINGO.
*          E.G.  'PREVENT JOHN FROM LEAVING' ==>
*                'PREVENT (JOHN LEAVE)'.
* T-SOBJBE
*     CONVERTS SOBJBE INTO A FULL ASSERTION AND SETS NODE ATTRIBUTE TFORM-ATT
*     TO LIST CONTAINING ATTRIBUTE TSOBJBE WHICH MARKS THE SOURCE OF THE
*     ASSERTION.
*        E.G.  FINDINGS SHOWED THE MASS SUSPICIOUS ==>
*              FINDINGS SHOWED (THE MASS BE SUSPICIOUS).
* T-SASOBJBE
```

```
*      CONVERTS SASOBJBE INTO A FULL ASSERTION WITH VERB = 'AS'% (VBE).
*      IT ALSO SETS A NODE ATTRIBUTE TFORM-ATT POINTING FROM THE NEWLY
*      BUILT ASSERTION TO LIST CONTAINING THE ATTRIBUTE TSASOBJBE
*      MARKING THE SOURCE OF THE ASSERTION.
*         E.G.  THEY TREATED> THE CHILD AS A SERVANT
*                             ==>
*               THEY TREATED (THE CHILD AS(=VBE) A SERVANT
* T-FORTOVO
*      CONVERTS FORTOVO INTO A COMPLETE ASSERTION AND SET A NODE ATTRIBUTE
*      TFORM-ATT POINTING FROM NEW ASSERTION TO LIST CONTAINING ATTRIBUTE
*      TFORTOVO MARKING THE SOURCE OF THE ASSERTION.
*           E.G. PATIENT WISHES FOR HER TO KNOW ==>
*                PATIENT WISHES (HER KNOW ).
* T-MODAL
*   (1) FOR SUFFIXED FORM, REPLACES W BY CLASS W  (TO REMOVE SUFFIX);
*   (2)PLACES NEGATIVE FROM SUFFIX 'NT' OR FROM RW AT BEG. OF 3RD SA;
*         (NOTE -NEGATED 'WILL' IS NOT CONSIDERED FUTURE TENSE)
*   (3)MOVES NODE TENSE (OTHER THAN FORMS OF DO) UNDER MODAL>
*         AFTER CORE OF VERB.
*           MODAL> WILL DOMINATE ELEMENTS SUCH AS 'CAN', 'MUST', ETC.
*   (4)REMOVES TENSE FROM STRING.
*   (5)SETS NODE ATTRIBUTE TENSE-ATT POINTING FROM CORE OF VERB
*      TO LIST CONTAINING TENSE ATTRIBUTES PRESNT, PAST, FUTURE.
*           E.G. HE DIDN'T GO TO THE LIBRARY.  ==>
*                HE NOT GO TO THE LIBRARY.  (TENSE-ATT = PAST)
*
* NOTE% THIS DOES NOT DEAL WITH SCOPE PROBLEMS- IN RELATION TO SA'S OR
*    TO NEGATIVES
*    (IN PARTICULAR IT DOES NOT HANDLE  'MAY NOT' VS. 'CAN NOT').
* T-TENSE
*      STRIPS SIMPLE PRESENT OR PAST TENSE FROM VERB NODE, REPLACES VERB
*      BY ITS CANONICAL FORM, AND SETS NODE ATTRIBUTE TENSE-ATT
*      POINTING FROM CORE- OF VERBAL TO LIST CONTAINING ATTRIBUTE
*      IDENTIFYING TENSE INFORMATION THAT WAS STRIPPED, I.E. ATTRIBUTE
*      PAST OR PRESNT.
*           E.G. HE WENT TO CALIFORNIA.  ==>
*                HE GO TO CALIFORNIA.  (TENSE-ATT = PAST)
*      IF THE VERB IS IN THE PRESENT TENSE, AND NOT 'HAVE' OR 'BE', IT
*      IS ALSO MARKED AS GENERIC TO MARK THE FACT THAT THE EVENT
*      TAKES PLACE EITHER REPEATEDLY OR OVER A TIME PERIOD.
*           E.G. HE LIVES IN CANADA.  ==>
*                HE LIVE 'GENERIC' IN CANADA.
* T-REMOVE-THERE
*      T-REMOVE-THERE REPLACES A 'THERE' SUBJECT OF ANN ASSERTION BY ITS
*     NSTG OBJECT AND THEN MOVES THE RN [VENPASS, VINGO, ADJINRN, OR PN]
*     OF THE NEW SUBJECT NSTG INTO THE OBJECT POSITION, LEAVING A NULL RN
*     IN THE SUBJECT.  THIS PREVENTS THE FORMATION OF EXTRA ASSERTIONS BY
*     T-RN-FILLIN.  IF THE RN IS NOT ONE OF THE CLASS SPECIFIED ABOVE,
*     THEN THE OBJECT NSTG IS MOVED INTO SUBJECT POSITION AND THE VERB VBE
*     IS REPLACED BY 'EXIST'.
*           1.   THERE WAS CARDIOMEGALY NOTED-->CARDIOMEGALY WAS NOTED.
*           2.   THERE WAS MENINGITIS DUE TO BACTERIAL INNFECTION -->
*                MENINGITIS WAS DUE TO BACTERIAL INFECTION.
*           3.   THERE WAS HISTORY OF PAIN --> HISTORY OF PAIN EXISTS.
*
```

```
* T-REG-FRAG EXPANDS NSTG FRAGMENTS WITH VENPASS OR VINGO IN RN INTO
*     FULL ASSERTIONS WITH BE.
*      EX.  NO RALES HEARD PREVIOUSLY==>NO RALES BE HEARD PREVIOUSLY.
* T-FRAG-TO-ASSRT
*     REGULARIZES FRAGMENTS, CHANGING THEM INTO ASSERTIONS%
*    1. SOBSBESHOW
*        REMOVE THE SOBJBESHOW NODE, ADD APPROPRIATE VERB BASED ON
*        SELECT-ATT ASSIGNED TO SOBJBESHOW, AND BUILD CORRECT OBJECT
*        NODES; ADD SA'S AND RV'S.
*              E.G. EXTREMITIES NORMAL. ==> EXTREMITIES BE NORMAL.
*                   SKIN NO ERUPTIONS.  ==> SKIN SHOW NO ERUPTIONS.
*
*    2. VFORM  = TOVO, VINGO, VENPASS%
*          ADD NULL SUBJECT, VERB = BE, BUILD CORRECT OBJECT, ADD SA'S
*          AND RV.
*             E.G. TOLD TO RETURN TO CLINIC. ==>
*                  () BE TOLD TO RETURN TO CLINIC.
*    3. ASTG OR PN FRAGMENT%
*          ADD NULL SUBJECT, VERB = BE, BUILD CORRECT OBJECT (=OBJBE),
*          ADD SA'S AND RV.
*             E.G. ON PROPHYLAXIS WITH BICILLIN. ==>
*                  () BE ON PROPHYLAXIS WITH BICILLIN.
*    4. TVO%
*          ADD NULL SUBJECT.
*             E.G. WAS SEEN BY LOCAL MD.  ==>
*                  () WAS SEEN BY LOCAL MD.
*
*    REGISTERS%
*       X1 = ORIGINAL FRAGMENT, UNDER WHICH ELEMENTS OF ASSERTION
*            ARE INSERTED
*       X2 = SECOND ELEMENT OF FRAG - DETERMINES TYPE OF FRAGMENT.
*       X3 = VERB BUILT FOR ASSERTION
*       X4 = TYPE OF OBJECT-TO-BE% OBJBE/VINGO/VENPASS
*       X5 = NEW ASSERTION, REPLACING FRAGMENT NODE.
*    STRATEGY%
*       FILL IN MISSING ELEMENTS OF ASSERTION [UNDER EXISTING FRAG NODE]
*       FROM LEFT TO RIGHT.
* T-IMPERATIVE
*     CONVERTS AN IMPERATIVE INTO AN ASSERTION WITH NULL (= YOU) SUBJECT;
*     THE OPTIONAL ELEMENT 'PLEASE' IS PLACED IN THE POST-SUBJECT SA;
*     OPTIONAL 'DO' IS REMOVED ; A NEGATIVE ON 'DO' IS PLACED IN POST-
*     TENSE SA;  THE LVR IS REPLACED BY VERB> WITH THE NODE ATTRIBUTE
*     TENSE-ATT POINTING FROM THE CORE- OF THE VERB TO A LIST CONTAINING
*     ATTRIBUTE IMPERTVE.
* T-VENO
*     TRANSFORMS PERFECT TENSE CONSTRUCTIONS BY REPLACING THE AUX HAVE (IN VERB)
*     WITH THE UNTENSED FORM OF THE VERB.  IT ALSO MARKS THE CORE- OF THE VERB
*     AS PERFECT BY SETTING NODE ATTRIBUTE TENSE-ATT TO POINT TO LIST
*     CONTAINING ELEMENT PERF.  NOTE THAT TENSE-ATT LIST MAY CONTAIN
*     MORE THAT ONE ELEMENT (E.G. PAST PERF).
*     THE OBJECT OF VENO IS MOVED INTO THE OBJECT OF ASSERTION
*     AND THE SA'S AND RV'S OF VENO ARE RELOCATED IN ASSERTION.
*     THIS TRANSFORMATION NOT ONLY REMOVES AFFIXES BUT ALSO
*     REGULARIZES THE TYPE OF OBJECT OTHER TRANSFORMATIONS MUST DEAL
*     WITH.  IT IS CRUCIALLY ORDERED BEFORE T-VINGO AND AFTER T-TENSE.
```

```
*      THE SENTENCE 'HE HAD FINISHED THE ASSIGNMENT', AFTER BEING OPERATED
*      ON PREVIOUSLY BY T-TENSE TO GET THE NODE ATTRIBUTE TENSE-ATT = PRESNT
*      IS OPERATED ON BY T-VENO.
*           E.G. HE HAVE (TENSE-ATT = PRESNT) FINISHED THE ASSIGNMENT. ==>
*                HE FINISH (TENSE-ATT = PRESNT PERF) THE ASSIGNMENT.
*      NOTE%  THIS TRANSFORMATION WILL NOT OPERATE ON PARTICIPIAL
*      CONSTRUCTIONS, E.G. 'AFTER HAVING TAKEN PENICILLIN...', SINCE
*      HERE T-VENO WOULD HAVE TO BE EXECUTED AFTER T-VINGO.
* T-VINGO
*      TRANSFORMS ASSERTIONS WITH PROGRESSIVE VERBS INTO ASSERTIONS WITHOUT THE
*      PROGRESSIVE STRUCTURE AND THE TENSE INFORMATION 'PROG' ADDED TO THE SIMP-
*      LIFIED VERB.  IT SUBSTITUTES THE UNTENSED V OF VING FOR THE FORM OF BE AND
*      THEN MOVES ANY NON-EMPTY STRUCTURE IN THE ORIGINAL OBJECT UP TO THE ASSER-
*      TION LEVEL, INCLUDING RELOCATING THE SA'S AND RV'S OF VINGO IN THE ASSER-
*      TION SA AND RV SLOTS.
*     IN T-VINGO, X30 IS VINGO, AND X35  IS THE VERB IN ASSERTION.
* T-VERBAL-OBJECT
*      TRANSFORMS ASSERTIONS WITH VINGO OR TOVO OBJECTS INTO ASSERTIONS
*     WITH VERB FROM TOVO OR VINGO AS MAIN VERB OF THE ASSERTION. THE
*     FIRST THREE SA'S FROM THE ORIGINAL ASSERTION ARE COPIED INTO THE
*     LV OF VERBAL; THE 4TH SA, RV, 5TH SA ARE COPIED INTO RV OF
*     VERBAL. THEN THE ORIGINAL VERB OF ASSERTION IS MOVED INTO THE
*     3RD SA [SA AFTER TENSE] OF THE ASSERTION. THE ORIGINAL VERB
*     CARRIES ALONG ITS TENSE INFORMATION AS A NODE-ATTRIBUTE TENSE-
*     ATT. TENSE-ATT WAS PREVIOUSLY ASSIGNED, BEFORE T-VERBAL-OBJECT IS
*     EXECUTED. THEN MOVES ANY NON-EMPTY STRUCTURE IN THE ORIGINAL
*     OBJECT UP TO THE ASSERTION LEVEL, INCLUDING RELOCATING THE SA'S
*     AND RV'S OF VERBAL IN OBJECT INTO THE ASSERTION SA AND RV SLOTS.
*          'HE BEGAN TO WALK' --> 'HE BEGAN WALK'
*          'HE BEGAN WALKING' --> 'HE BEGAN WALK'.
* T-TOVO-IN-SA
*      TRANSFORMS A TOVO IN SA INTO A CSSTG WITH 'IN ORDER TO' AS THE
*      SUBORDINATING CONJUNCTION, FOLLOWED BY AN ASSERTION.
*      THE SUBJECT OF THE ASSERTION IS COPIED FROM THE ASSERTION
*      CONTAINING THE SA, WHILE THE OTHER ELEMENTS, I.E. VERB, SA'S,
*      OBJECT, RV, ARE COPIED FROM TOVO.  THE ENTIRE SUBJECT RATHER THAN
*      JUST THE CORE OF THE NSTG IN SUBJECT WAS COPIED BECAUSE THE
*      SUBJECT CAN HAVE OTHER VALUES, E.G. NSVINGO, VINGOFN, ETC.
*      FOR WHICH THE TRANSFORMATION IS ALSO VALID.
*           E.G.  THE PATIENT WAS TRANSFERRED TO MEDICINE TO BEGIN
*                        THERAPY     ==>
*                 THE PATIENT WAS TRANSFERRED TO MEDICINE
*                     IN ORDER TO PATIENT BEGIN THERAPY.
* T-TOVO-FUTURE
*      CHANGES 'IS TO' TOVO PHRASES (E.G. IS TO BE FOLLOWED) INTO AN ASSERTION
*      IF THE SUBJECT INFORMATION IS PRESENT OR INTO TVO WHEN IT IS NOT.
*      THE VERB OF THE ASSERTION IS COPIED FROM THE VERB OF THE TOVO
*      AND A NODE ATTRIBUTE 'TENSE-ATT' IS SET POINTING FROM THE CORE
*      OF THE VERB TO A LIST CONTAINING THE ATTRIBUTE 'FUTURE'.
*      OBJECT OF TOVO IS MOVED UP TO OBJECT OF ASSERTION OR TVO, AND
*      RV'S AND SA'S OF TOVO ARE RELOCATED IN THE HIGHER STRING.
*      NOTICE THAT IT IS POSIBLE TO GET A TENSE 'FUTURE PAST PERFECT' FROM
*      A SENTENCE LIKE 'HE WAS TO HAVE BEEN FOLLOWED'.
*           E.G. HE IS TO BE FOLLOWED IN CLINIC. ==>
*                HE IS (HE BE (TENSE-ATT = FUTURE) FOLLOWED IN CLINIC).
```

```
* T-SA-VFORM
*   TRANSFORMS VENPASS OR VINGO IN SA OF ASSERTION-LIKE STRING INTO
*   SUB2 [VENPASS] WITH CS = 'IN-STATE' OR SUB3 [VINGO] WITH CS =
*   'WHILE' IN CSSTG UNDER THE SA. T-CSSTG WILL OPERATE ON THESE
*   STRINGS TO FILL THEM OUT INTO COMPLETE ASSERTIONS.
*         E.G. [VENPASS]
*               ADMITTED TO HOSPITAL, PATIENT COMPLAINED OF HEADACHE.-->
*               'IN-STATE' ADMITTED TO HOSPITAL, PATIENT COMPLAINED....
*              [VINGO]
*               HAVING A SEVERE HEADACHE, PATIENT WAS ADMITTED.
*               'WHILE' HAVING A SEVERE HEADACHE.
*   NOTE THAT T-SA-VFORM DOES NOT PICK UP THE VERBAL SA AND MOVE IT
*   INTO RN OF SUBJECT OR OBJECT OF THE ASSERTION DUE TO PROBLEM
*   WITH VINGO CONTAINING 'HAVE' -
*      HAVING A SEVERE HEADACHE, PATIENT WAS ADMITTED TO HOSPITAL
*      --> *PATIENT SUCH THAT PATIENT BE HAVING A HEADACHE
*               WAS ADMITTED TO HOSPITAL.
* T-CSSTG OPERATES ON 3 CASES OF CSSTG%
*     1) SUBO%  OBJBE + SA.
*            E.G. WHILE IN THE ARMY, HE WAS DECORATED. ==>
*                 WHILE HE BE IN THE ARMY, HE WAS DECORATED.
*     2) SUB2%  VENPASS.
*            E.G. AS REQUESTED, HE FILLED THE ORDER. ==>
*                 AS HE BE REQUESTED, HE FILLED THE ORDER.
*     3) SUB3%  VINGO.
*           E.G. AFTER SLEEPING FOR THREE HOURS, HE AWOKE. ==>
*                AFTER HE BE SLEEPING FOR THREE HOURS, HE AWOKE.
*       IN CASES 1- 3, IN ORDER TO CONVERT THESE STRINGS
*     INTO AN ASSERTION, THE SUBJECT AND VERB = 'BE' MUST BE BUILT.
*     IF THE SUBJECT OF THE HOST STRING IS A PERMISSIBLE ONE, IT IS
*     COPIED AS SUBJECT OF THE NEWLY-BUILT ASSERTION UNDER CSSTG;
*     OTHERWISE SUBJECT IS LEFT BLANK.
*     THE REMAINING CASES (ASSERTION, NSVINGO, SVINGO, SOBJBE, SVEN)
*     ARE HANDLED BY OTHER TRANSFORMATIONS.
* T-IT-REPLACEMENT
*      REPLACES 'IT' IN SUBJECT BY SN WHEN THE SN IS IN RV OF VENPASS
*     IN THE OBJECT OF AN ASSERTION.
*         E.G. IT WAS DECIDED TO CONTINUE TREATMENT.  ==>
*              TO CONTINUE TREATMENT WAS DECIDED.
*      NOTE% T-PASSIVE WILL OPERATE ON THE TRANSFORMED ASSERTION.
* T-PASSIVE
*     TURNS CERTAIN PASSIVE ASSERTIONS INTO ACTIVE ASSERTIONS.
*     T-PASSIVE CAN BE CONSIDERED A SPECIAL TYPE OF CANONICAL FORM
*     TRANSFORMATION WHICH CONVERTS A PASSIVE INTO AN ACTIVE.
*     IT MUST ALSO REARRANGE THE ASSERTION:
*     THE SUBJECT IS COMBINED WITH THE PASSOBJ FOUND IN VENPASS
*     TO FORM A NEW OBJECT; THE TRANSFORMATION SEARCHES FOR AN AGENT IN
*     FORM OF PN, WHERE P = 'BY', WHICH BECOMES THE NEW SUBJECT (IF NONE IS
*     FOUND, THE SUBJECT = NSTG (NULL));  THE ADJUNCTS FROM VENPASS
*     STRING IN OBJECT ARE MAPPED INTO ADJUNCT SLOTS ON THE NEW VERB
*     OR UNDER THE MAIN ASSERTION.
*   $CORR-OBJECT SUBSTATEMENT IDENTIFIES TYPE OF PASSIVE OBJECT AND
*     COMBINES IT WITH SURFACE SUBJECT TO FORM ACTIVE OBJECT.
*    -$401 - NULLOBJ
*        - $4011 - NSTG SUBJ  ==> NSTGO
```

```
*           - $4012 - SNSUBJ ==> THATS, TOVO OR SNWH (VALUE)
*           - $4013 - VINGSTG SUBJ  ==> VINGSTG
*       - $402 - P1 OBJ
*           - $4021 - NSTG SUBJ  ==> PN
*           - $4022 - VINGSTG SUBJ  ==> PVINGSTG
*           - $4023 - SNWH SUBJ  ==> PSNWH
*       - $403 - MISCELLANEOUS
*           - $4031 - VINGO OBJ  ==> SVINGO  (THEY WERE SEEN LAUGHING ==>
*                                             NULL SAW THEY LAUGHING)
*           - $4032 - TOVO OBJ  ==> NTOVO (THEY WERE KNOWN TO KILL  ==>
*                                          NULL KNEW THEY TO KILL.
*           - $4033 - TOBE OBJ  ==> NTOBE
*       - $404 - NSTGO PASSOBJ (WE ARE CALLED DOGS BY THEM )
*           - $4041 - NSTG SUBJ  ==> NN
*           - $4042 - THAT SUBJ  ==> NTHATS (THAT X WE WARN THEM)
*           - $4043 - SNWHSUBJ ==> NSNWH
*       - $405 - IT SUBJ - THAT OBJ  ==> NULL OBJ/THAT SUBJ (IT WAS KNOWN THAT X ==
>
*                                                      NULL KNEW THAT X)
*       - $406 - PN OBJ
*           - $NSTG - NSTG SUBJ NPN
*           - $VINGSTG - VINGSTG SUBJ  ==> VINGSTGPN
*           - $SNWH - SNWH SUBJ  ==> PNSNWH
*           - $THATS - THAT SUBJ  ==> PNTHATS (THATS) --(TENSED VERB)
*           - $C1SHOULD - THATS SUBJ  ==> PNTHATSVO (C1SHOULD) - (UNTENSED VERB)
*       -$407 - DP OBJECTS
*           - $4071 - NSTG/VINGSTG SUBJECT, DP1 OBJECT  ==> DP2
*           - $4072 - NSTG/VINGSTG SUBJECT, DP1PN OBJECT  ==> DP2PN
*           - $4073 - NSTG/VINGSTG SUBJECT, DP1P OBJECT  ==> DP1PN
*        - $408 - NSTG SUBJECT- ASSERTION/THATS OBJ  ==> NTHATS
*                   (MOTHER WAS TOLD SHE HAD A SORE THROAT , S
*                    NULL TOLD MOTHER SHE HAD A SORE THROAT.)
*        - $409 - ASOBJBE OBJECT  ==> SASOBJBE (THE CHILD WAS TREATED
*                                  AS A PARTIALLY TREATED MENINGITIS  ==>
*                                  () TREATED THE CHILD AS A PARTIALLY
*                                   TREATED MENINGITIS.
* T-DP
*     INSERTS THE DP FROM A DP OBJECT AFTER THE CORE OF ITS VERBAL
*     ELEMENT AND REPLACES THE VALUE OF THE OBJECT BY AN APPROPRIATE
*     OBJECT STRING CONTAINING THE REMAINING ELEMENTS OF THE DP OBJECT,
*        E.G. TREATMENT WAS CARRIED OUT AT ANOTHER HOSPITAL (DP = OUT).
* T-TIME-PREFIX
*     CHECKS FOR TIME-EXPRESSION ADJUNCTS        (E.G. 'NEONATAL'
*     OR 'POST-OPERATIVELY'); WHEN IT FINDS ONE, IT LOOKS %
*        1.  FOR DSTG --FOR IMMEDIATE RV OR SA WHICH CAN ACCOMODATE A PN
*        2.  FOR LAR1 --FOR RN OF HOST, WHICH CAN ACCOMODATE A PN
*        3.  FOR LAR IN ADJINRN --FOR IMMEDIATE RN
*        4.  FOR LAR IN ASTG --FOR IMMEDIATE OBJBE WHICH CAN HAVE A PN
*      IT THEN ADDS TO THE RN OR OBJBE OR RV OR SA (STORED IN X2)
*      A PN WHERE PREP = PREFIX (MARKED AS H-PREPTIME FOR FORMATTING)
*      AND N = CLASS NOUN OF THE ADJ OR D; THE ORIGINAL PREFIXED WORD
*      IS THEN DELETED.
*          E.G. NEONATAL ==>   (P = NEO) BIRTH
*               POST-OPERATIVE ==>   (P = POST) OPERATION.
* T-NOUN
*     STRIPS OFF PLURAL INTO A MARKER APPENDED AFTER THE N;  IT ALSO
```

```
*       REPLACES THE NOUN BY ITS CANONICAL FORM, SO THAT E.G. 'ADM' WILL
*       BE REPLACED BY 'ADMISSION'.
*
* T-ADJ
*       REPLACES ADJECTIVE BY ITS CANONICAL FORM.
*
* T-D
*       REPLACES D (ADVERB) BY ITS CANONICAL FORM
*            E.G. '2ND' FOR 'SECOND'.
*
* T-Q
*       REPLACES Q BY ITS CANONICAL FORM,
*            E.G. THE DIGIT '2' REPLACES 'TWO'.
*
* T-WITHOUT
*       CHANGES 'WITHOUT' PREPOSITIONAL PHRASES TO 'WITH' + NEGATIVE:
*         (1) IF THE ARTICLE ON THE OBJECT OF THE PREPOSITION IS EMPTY
*               OR INDEFINITE
*            THEN TRANSFORMATIONS FILLS IN THE ARTICLE WITH 'NO'
*               E.G. WITHOUT COMPLICATIONS ==> WITH NO COMPLICATIONS.
*         (2) IF THE ARTICLE IS NOT EMPTY OR INDEFINITE, THEN 'NOT' IS
*            INSERTED AS ADJUNCT TO THE PREPOSITION 'WITH',
*               E.G. WITHOUT THE COMPLICATIONS ==> NOT WITH THE COMPLICATIONS.
*            IF 'NOT' IS ALREADY IN THE MODIFIER TO THE PREPOSITION, THEN
*            IT IS REPLACED  BY NULL.
*               E.G. NOT WITHOUT THE COMPLICATIONS ==> WITH THE COMPLICATIONS.
* T-NOM-PREP-ARG
*           IDENTIFIES PREPOSITIONS WHICH "BELONG" TO A NOMINALIZED
*           VERB, E.G. 'ADMISSION', MARKING PREP ('TO') WITH NODE ATTRIBUTE
*           PVAL-ATT POINTING TO HOST (CONSISTENT WITH $PVAL IN WSEL-P-N IN
*           SELECTION). THESE PREPOSITIONS ARE STATED IN THE
*           DICTIONARY ENTRY OF THE VERB WHICH IS RELATED TO THE
*           NOMINALIZATION.
*              E.G. 'FIRST ADMISSION TO HOSPITAL.'  ==>
*                   (PVAL-ATT POINTING FROM 'TO' TO 'ADMISSION').
* T-VERB-PREP-ARG
*        MARKS PREPOSITIONAL ARGUMENT OF VERB (E.G. 'ADMIT') WITH PVAL-ATT
*        POINTING FROM P TO VERB.
*             E.G. THE PATIENT WAS ADMITTED TO HOSPITAL.  ==>
*                 (PVAL-ATT POINTING FROM 'TO' TO 'ADMIT').
* T-WHATS-N
*        TRANSFORMS THE 'WHAT' OF NSTG INTO A 'THAT WHICH', WHICH CAN
*           BE HANDLED BY THE T-RN-WH TFORM.
*              E.G. I SAW WHAT HE HAD BOUGHT. ==>
*                   I SAW THAT WHICH HE HAD BOUGHT.
* T-QUANT-OF
*       TRANSFORM STRUCTURES OF THE TYPE 'MANY OF THE PATIENTS',
*        'SOME OF THE NEW PEOPLE', ETC. REARRANGING STRUCTURE SO THAT
*        THE QUANTIFIER BECOMES A MODIFIER IN LN OF THE REAL HEAD N
*        (E.G. 'PATIENTS', 'PEOPLE').
*        1. IF THE NOUN IS NULLN, IT SEARCHES FOR 'SOME' OR Q IN LN AND
*           RN = OF + N;
*      2. $CHECK-EMPTY CHECKS THAT ALL OTHER POSITIONS IN LN ARE EMPTY;
*                ASSUMES THERE IS NO INFORMATION IN UPPER LN OTHER THAN
*                QUANTIFIER INFORMATION (E.G. *SEVERAL NEW OF THE PATIENTS).
```

```
*        IT WILL FAIL IF THIS IS NOT TRUE--WITH THE EXCEPTION OF 'A' IN
*        IN TPOS WHERE Q = 'FEW'; IN THIS CASE, 'A' IS INSERTED BEFORE 'FEW';
*     3. IT MOVES THE QUANTIFIER + OF FOUND IN STEP 1 TO QPOS OF THE LNR
*        IN RN;
*        IF QPOS IS FILLED [E.G. 'SOME OF THE 10 PEOPLE], $MOVE-QPOS MOVES
*        MATERIAL IN QPOS INTO A SEPARATE SENTENCE%
*        'SOME OF THE PEOPLE SUCH THAT PEOPLE NUMBER 10', WHERE OBJECT
*        IS ASTG(LQR).
*      4. IT PROMOTES THE LNR IN PN TO A HIGHER LEVEL TO REPLACE THE
*         ORIGINAL LNR = NULLN.
* T-VENAPOS
*     FLIPS VEN'S IN APOS INTO AN ADJINRN IN RN SO T-RN-FILLIN CAN EXPAND THE
*     HOST NOUN + VEN INTO AN ASSERTION.
*         E.G. HIS INJURED FOOT ==>
*              HIS FOOT INJURED.
*             LATER BY T-RN-FILLIN:
*               HIS FOOT INJURED  ==>  HIS FOOT (FOOT BE INJURED).
* T-RN-WH
*     FILLS IN NULLWH IN ASSERTION OF WHS-N, THATS-N, OR S-N;
*     IN THE CASE OF WHENS OR PWHS, IT COPIES A PN TO LAST SA OF
*     ASSERTION.  ASSERTION IS MOVED UP TO SAME LEVEL AS LNR, TO CREATE
*     STRUCTURE NSTG (LNR +  (ASSERTION)).  TO FILLIN NULLWH, CORE
*     OF LNR HOST IS COPIED ALONG WITH ANY ADJUNCT INVOLVED IN A
*      COMPUTED-ATT CONSTRUCTION (USING GLOBAL $TRIM-LN-RN).  THE
*     FILLED IN ASSERTION IS THEN TRANSFORMED.
*         E G   THE MAN WHO LEFT TOWN    ==>
*               THE MAN (THE MAN LEFT TOWN).
*     THE NODE ATTRIBUTE 'TFORM-ATT' IS SET POINTING FROM THE NEWLY
*     CREATED ASSERTION TO A LIST CONTAINING THE ATTRIBUTE 'TRNWH'.
* T-RN-FILLIN EXPANDS CERTAIN RNS INTO A FULL ASSERTION, WITH SUBJECT =
*    HOST; IT IS THEN TREATED SIMILARLY TO AN EXPANDED RELATIVE CLAUSE.
*    $CHECK-COOC LIMITS ACTION OF THIS TFORM TO CASES WHERE THE RN
*    CONTAINS AN H-CONN WORD E.G. 'ANEMIA CAUSED BY/ DUE TO/ SCD' WILL
*    BE EXPANDED, BUT 'ANEMIA NOTED BY MOTHER' WILL NOT BE.
*     A NODE ATTRIBUTE 'TFORM-ATT' IS SET POINTING FROM THE NEWLY CREATED
*     ASSERTION TO A LIST CONTAINING THE ELEMENT 'TRNFILLIN'.
*          E.G. ANEMIA DUE TO SCD ==>
*               ANEMIA (ANEMIA BE DUE TO SCD)...


* ***** *************************************************************
*
*
*                        CONNECTIVE TRANSFORMATIONS
*
*
* *******************************************************************
*   THIS COMPONENT IS EXECUTED USING OUTPUT TREES FROM THE ENGLISH
* DECOMPOSITION COMPONENT. EACH TRANSFORMATION IN THE CONNECTIVE
* COMPONENT (WITH THE EXCEPTION OF T-FIND-HOST) CREATES A
* CONNECTIVE --PARSE-CONN-- CONNECTING ONE ASSERTION/FRAGMENT
* TO ANOTHER ASSERTION/FRAGMENT. PARSE-CONN HAS THE FOLLOWING
* STRUCTURE%
*      PARSE-CONN  =  X  =  SA + LCONNR + SA.
*        WHERE X IS THE NAME OF THE TYPE OF CONNECTIVE SUCH AS
```

```
*        'CONJOINED', 'EMBEDDED', ETC.
*      LCONNR  =  LCONN + HEADCONN + RCONN.
* LCONN AND RCONN ARE THE LEFT AND RIGHT ADJUNCTS OF THE CONECTIVE
* HEADCONN. THE SUBSTRUCTURE OF HEADCONN DEPENDS ON THE TYPE OF
* CONNECTIVE. IT WILL BE DESCRIBED IN EACH TRANSFORMATION.
*   IN GENERAL, WHEN ONE OF THE TRANSFORMATIONS FINDS A RELEVANT
* SUBSTRUCTURE IN AN ASSERTION 'A', IT ATTACHES A CONNECTIVE
* 'PARSE-CONN' TO THE LEFT OF 'A' AND CREATES AN ASSERTION/FRAGMENT
* 'B' FROM 'A' AND ATTACHES IT TO THE RIGHT OF 'A'. 'A' MAY BE
* CHANGED TO 'A1' IN THE PROCESS. WHEN A SUCCESSFUL CONNECTIVE
* TRANSFORMATION IS COMPLETED IN ASSERTION 'A', THE STRUCTURE
*              PARSE-CONN + A1 + B + CONJ-NODE
* REPLACES 'A'. WHEN AN ASSERTION IN 'A' IS MOVED UP, ITS CONJUNCT (IF
* IT HAS ONE) IS ALSO MOVED UP ALONG WITH CONJ-NODE SO THAT WHEN THE
* TRANSFORMATION IS COMPLETED WE HAVE%
*      PARSE-CONN + A1 + B + CONJ-NODE
*        WHERE    CONJ-NODE  =  CONJ + CONJUNCT OF B.
* WHEN TRANSFORMING ASSERTION 'B', THE FIRST TRANSFORMATION IS
* T-CONJ-IN-ASSERTION WHICH WILL CREATE A CONNECTIVE 'PARSE-CONN'  =
* 'CONJOINED' TO THE LEFT OF 'B' AND THE CONJUNCT OF 'B' WILL BE ATTACHED
* TO THE RIGHT OF 'B'. THE ABOVE STRUCTURE WILL THEN BE%
*      PARSE-CONN + A1 + PARSE-CONN + B + CONJUNCT OF B.
*
* T-CONJ-IN-CENTER
*      OPERATES WHEN THE VALUE V1 OF OLD-SENTENCE/CENTER HAS A CONJUNCT
* V2. PARSE-CONN IS ATTACHED TO THE LEFT OF V1. V2 IS MOVED TO THE RIGHT
* OF V1 SO THAT WE HAVE%    PARSE-CONN + V1 + V2    .
*           PARSE-CONN  =   CONJOINED  =   SA + LCONNR + SA
*           HEADCONN  =   FIRST ELEMENT OF CONJ-NODE.
* IF SACONJ IS NOT EMPTY, ITS VALUE SA IS MOVED TO THE FIRST SA OF
* CONJOINED. IF 'NOT' IS IN CONJ-NODE,  DSTG = D = NOT  IS CREATED
* AND ALSO MOVED TO THE FIRST SA OF CONJOINED.
*
* T-CONJ-IN-ASSERTION IS SIMILAR TO T-CONJ-IN-CENTER. IT OPERATES WHEN
* ASSERTION HAS A CONJUNCT.
* IT HANDLES SITUATIONS WHERE CONJOINED ASSERTIONS ARE NOT UNDER
* CENTER.
*
* T-CSSTG
*      OPERATES WHEN AN SA IN ASSERTION A = LCS + CSSTG.
*      PARSE-CONN = SUB-CONJ IS ATTACHED TO THE LEFT OF A.
*      THE ASSERTION(S) OF CSSTG IS(ARE) MOVED TO THE RIGHT OF A.
*         HEADCONN  =  CS  (I.E., THE FIRST ELEMENT OF CSSTG).
*         LCONN  =  ELEMENTS OF LCS.
*         CSSTG IS REPLACED IN A BY NULL.
*
* T-SA-PNCONN
*   OPERATES WHEN SA IN ASSERTION A (OR SA IN OBJECT OF A = NN/NPN/
* PNN) = PN WHERE P = H-CONN.
*      PARSE-CONN = RELATION IS ATTACHED TO THE LEFT OF A.
*      HEADCONN = P OF PN.
*      LCONN = ELEMENTS OF LP OF PN.
*      FRAGMENT B = NSTG [WHICH IS COPIED FROM NSTG OF NSTGO OF PN]  IS
*        ATTACHED TO LEFT OF A.
*      PN IS REPLACED BY NULL IN A.
```

```
*
* T-FIND-CONN
*   OPERATES WHEN THERE IS A CONNECTIVE OR H-BECONN IN ASSERTION A.
* IT FINDS H-CONN/H-BECONN AND ITS TWO ARGUMENTS. A IS REPLACED BY
* PARSE-CONN + FRAGMENT1 + FRAGMENT2   WHERE%
*           PARSE-CONN = RELATION = SA. + LCONNR + SA.
*   AND     FRAGMENT1 = NSTG (THE FIRST ARGUMENT OF H-CONN/H-BECONN).
*   AND     FRAGMENT2 = NSTG (THE SECOND ARGUMENT OF H-CONN/H-BECONN).
*           HEADCONN = ATOM CORRESPONDING TO WORD WHICH IS H-CONN/H-BECONN.
* THE SEARCH FOR H-CONN/H-BECONN AND ITS ARGUMENTS IS AS FOLLOWS%
*      1) CORE-SELATT OF CORE OF VERB HAS MEMBER CONN-LIST.
*           FRAGMENT1 = NSTG OF SUBJECT.
*           FRAGMENT2 = NSTG OF OBJECT.
*         CONTENTS OF RV IN ASSERTION ARE MOVED TO RV OF VERB.
*           HEADCONN = V
*           LCONN = ELEMENTS OF LV
*           RCONN = ELEMENTS OF RV
*         ALL SA'S TO THE LEFT OF VERB ARE MOVED INTO THE FIRST SA OF
*         RELATION. ALL SA'S TO THE RIGHT OF VERB ARE MOVED INTO THE
*         LAST SA OF RELATION.
*        EX.% 'FEVER CAUSES HEADACHE.'
*              FRAGMENT1 = 'FEVER'
*              FRAGMENT2 = 'HEADACHE'
*              HEADCONN = V = 'CAUSE'
*      2) VERB IS VBE/BEREP AND OBJECT = ASTG WHERE%
*           CORE-SELATT OF ADJ = H-CONN/C-BECONN  AND
*           RA = PN
*           FRAGMENT1 = NSTG OF SUBJECT
*           FRAGMENT2 = NSTG OF NSTGO OF PN
*           HEADCONN = ADJ + LPR
*             WHERE  LPR = LP = ELEMENTS OF LP OF PN  +  P OF PN.
*         EX.% 'FEVER IS COMPATABLE WITH HEADACHE.'
*           FRAGMENT1 = 'FEVER'
*           FRAGMENT2 = 'HEADACHE'
*           HEADCONN = 'COMPATABLE WITH'
*      3) VERB IS VBE/BEREP  AND
*           OBJECT = NSTG  WHERE
*               CORE N = H-CONN/H-BECONN  AND
*               RN = PN
*           FRAGMENT1 = NSTG OF SUBJECT
*           FRAGMENT2 = NSTG OF PN
*           HEADCONN = N + LPR
*        EX.% 'FEVER IS CAUSE OF HEADACHE.'
*               FRAGMENT1 = 'FEVER'
*               FRAGMENT2 = 'HEADACHE'
*               HEADCONN = 'CAUSE OF'.
*      4) VERB IS VBE/BEREP  AND
*           OBJECT = PN  WHERE
*             P = H-CONN/H-BECONN
*           FRAGMENT1 = NSTG O  F SUBJECT
*           FRAGMENT2 = NSTG O  F PN
*           HEADCONN = P
*        EX.% 'HEADACHE IS DUE TO FEVER.'
*              FRAGMENT1 = 'HEADACHE'
*              FRAGMENT2 = 'FEVER'
```

```
*               HEADCONN = 'DUE % TO'.
*        5) VERB IS VBE/BEREP  AND
*             SUBJECT = NSTG  WHERE
*               N IS H-CONN/BECONN  AND
*               RN = PN
*             FRAGMENT1 = NSTG OF OBJECT
*             FRAGMENT2 = NSTG OF PN
*             HEADCONN = N + LPR
*         EX.% 'CAUSE OF HEADACHE IS FEVER.'
*               FRAGMENT1 = 'FEVER'
*               FRAGMENT2 = 'HEADACHE'
*               HEADCONN = 'CAUSE OF'
*
* T-MOVE-S-UP
*   OPERATES WHEN VALUE OF SUBJECT/OBJECT OF ASSERTION A IS AN
* ASSERTION B.
*      PARSE-CONN = EMBEDDED   IS ATTACHED TO LEFT OF A.
*      B IS MOVED TO RIGHT OF A.
*      IN A, ASSERTION B IS REPLACED BY NULL>, WHICH IS ASSIGNED NODE
*          ATTRIBUTE EMBED-SUBJ/EMBED-OBJ.
*
* T-REL-CLAUSE
*   OPERATES WHEN AN LNR HAS AN ASSERTION B TO ITS RIGHT. THE ASSERTION
* WAS CREATED BY AN ENGLISH TRANSFORMATION WHICH ASSIGNED TO B THE
* NODE ATTRIBUTE TFORM-ATT, POINTING TO A LIST.THE LIST CONTAINS THE
* NAME OF THE TRANSFORMATION CREATING THE ASSERTION - IN THIS CASE
* T-RN-WH OR TRN-FILLIN.
*      ASSERTION A IS FOUND BY GOING UP FROM B.
*      PARSE-CONN = REL-CLAUSE IS ATTACHED TO THE LEFT OF A.
*      B IS MOVED TO THE RIGHT OF A.
*      B IS DELETED FROM A.
*      HEADCONN = N = 'TRN-WH'/'TRN-FILLIN'.
*
*
* T-EXPAND-REFPT
*   OPERATES ON PN WHEN%
*      1) THERE IS A NODE ATTRIBUTE TIME-ADVERBIAL ON PN AND
*         N IS PN IS NOT NTIME1/NTIME2/H-TIMELOC/H-AGE
*  OR  2) HOST OF PN IS NTIME1/NTIME2/H-TIMELOC
* ASSERTION A IS FOUND BY GOING UP TO ASSERTION FROM PN.
*      PARSE-CONN = REL-CLAUSE IS ATTACHED TO THE LEFT OF A.
*      HEADCONN = 'T-EXPAND-REFPT'.
*      FRAGMENT B = NSTG OF PN IS ATTACHED TO THE RIGHT OF A. IF N IS
*                   H-HOSP THERE IS A FURTHER CHECK - SEE $CHECK-H-HOSP.
*
*
* T-WITH-CONJ
*   OPERATES ON PN  WHERE
*      P IS 'WITH'  AND
*      P HAS NODE ATTRIBUTE SELECT-ATT WITH VALUE CONJ-LIKE.
*      PARSE-CONN = PREP-CONN
*      HEADCONN = 'WITH'
* ASSERTION A IS FOUND BY GOING UP FROM PN.
* FRAGMENT B = NSTG OF PN  IS ATTACHED TO THE RIGHT OF A.
```

APPENDIX A1

Transformational Decomposition of a Fact-Report Text

Elaine Marsh, Catherine Christenson, and Carolyn White

```
* HIDSM    1. 1. 1
* PRESENT ILLNESS - THE PATIENT WAS WELL UNTIL
* 2 WEEKS PTA WHEN SHE DEVELOPED BILATERAL
* OTITIS MEDIA .
(INTRODUCER='PRESENT' 'ILLNESS' '-' )

 (
    (SUB-CONJ='WHEN' )
    (ASSERTION=
       (SUBJECT='THE' 'PATIENT' )
       (VERB='BE' 'PAST' )
       (OBJECT='WELL' 'UNTIL' '2' 'WEEK' 'PLURAL' 'PTA' )
    )
    (ASSERTION=
       (SUBJECT='SHE' )
       (VERB='DEVELOP' 'PAST' )
       (OBJECT='BILATERAL' 'OTITIS' 'MEDIA' )
    )
 )


* HIDSM    1. 1. 2
* WAS SEEN BY LOCAL MD AND WAS TREATED WITH
* ERYTHROMYCIN , WHICH SHE TOOK FOR 5 DAYS
* .

 (
    (-CONJOINED='AND' )
    (ASSERTION=
       (SUBJECT='LOCAL' 'DOCTOR' )
       (VERB='SEE' 'PAST' )
    )
    (
       (REL-CLAUSE='TRN-WH' )
       (ASSERTION=
          (VERB='TREAT' 'PAST' )
          (OBJECT='WITH' 'ERYTHROMYCIN' ',' )
       )
       (ASSERTION=
          (SUBJECT='SHE' )
          (VERB='TAKE' 'PAST' 'FOR' '5' 'DAY' 'PLURAL' )
          (OBJECT='ERYTHROMYCIN' ',' )
       )
    )
 )
```

```
* HIDSM     1. 1. 3
* THE MEDICINE WAS THEN DISCONTINUED BECAUSE
* IT HAD BEEN KEPT OUT OF THE REFRIGERATOR
* AND WAS FELT TO BE STALE .

 (
    (SUB-CONJ='BECAUSE' )
    (ASSERTION=
       (VERB='DISCONTINUE' 'PAST' )
       (SA='THEN' )
       (OBJECT='THE' 'MEDICINE' )
    )
    (
       (-CONJOINED='AND' )
       (ASSERTION=
          (VERB='KEEP' 'OUT' 'PAST' 'PERF' )
          (OBJECT='IT' )
          (RV='OF' 'THE' 'REFRIGERATOR' )
       )
       (
          (EMBEDDED='EMBEDDED-OBJ' )
          (ASSERTION=
             (VERB='FEEL' 'PAST' )
          )
          (ASSERTION=
             (SUBJECT='IT' )
             (VERB='BE' )
             (OBJECT='STALE' )
          )
       )
    )
 )


* HIDSM     1. 1. 4
* THE CHILD WAS THEN ASYMPTOMATIC .

 (ASSERTION=
    (SUBJECT='THE' 'CHILD' )
    (VERB='BE' 'PAST' )
    (OBJECT='THEN'
       (NEG-PREFIX='SYMPTOMATIC' )
    )
 )


* HIDSM     1. 1. 8
* SHE WAS BROUGHT TO OUR ER WITH TEMPERATURE
* OF 103 DEGREES .

 (
    (REL-CLAUSE='TRN-FILLIN' )
    (ASSERTION=
       (VERB='BRING' 'PAST' )
```

```
      (OBJECT='SHE' 'TO' 'OUR' 'EMERGENCY' 'ROOM' )
   )
   (ASSERTION=
      (SUBJECT='SHE' )
      (VERB='HAVE' )
      (OBJECT=
         (QUANT='TEMPERATURE' 'OF' '103' 'DEGREE' 'PLURAL' )
      )
   )
)
```

```
* HIDSM    1. 1.11
* THE GRAM STAIN WAS NEGATIVE .

(ASSERTION=
   (SUBJECT='THE' 'GRAM' 'STAIN' )
   (VERB='BE' 'PAST' )
   (OBJECT=
     (NEG-FIND='NEGATIVE' )
   )
)
```

```
* HIDSM    1. 1.12
* THE CHILD WAS HENCEFORTH ADMITTED FOR TREATMENT
* OF MENINGITIS .

(
   (RELATION='FOR')
   (ASSERTION=
       (VERB='ADMIT' 'PAST' )
       (SA='HENCEFORTH' )
       (OBJECT='THE' 'CHILD' )
   )
   (FRAGMENT='TREATMENT' 'OF' 'MENINGITIS' )
)
```

```
* HIDSM    1. 1.13
* THERE WAS NO HX OF STIFF NECK , ABNORMAL
* POSTURING , DIARRHEA OR OTHER FOCAL SIGNS
* COMPATIBLE WITH MENINGITIS .

(
   (CONJOINED=',' )
   (
      (REL-CLAUSE='TRN-FILLIN' )
      (ASSERTION=
         (SUBJECT=
            (NEG-FIND='NO' 'HISTORY' 'PAST' 'OF' 'STIFF' 'NECK' )
         )
```

```
        (VERB='EXIST' 'PAST' )
     )
     (
        (RELATION='COMPATIBLE' 'WITH' )
        (FRAGMENT=
           (NSTG='STIFF' 'NECK' )
        )
        (FRAGMENT=
           (NSTG='MENINGITIS' )
        )
     )
  )
  (CONJOINED=',' )
  (
     (REL-CLAUSE='TRN-FILLIN' )
     (ASSERTION=
        (SUBJECT=
           (NEG-FIND='NO' 'HISTORY' 'PAST' 'OF' 'ABNORMAL' 'POSTURING' )
        )
        (VERB='EXIST' 'PAST' )
     )
     (
        (RELATION='COMPATIBLE' 'WITH' )
        (FRAGMENT=
           (NSTG='ABNORMAL' 'POSTURING' )
        )
        (FRAGMENT=
           (NSTG='MENINGITIS' )
        )
     )
  )
  (-CONJOINED='AND' )
  (
     (REL-CLAUSE='TRN-FILLIN' )
     (ASSERTION=
        (SUBJECT=
           (NEG-FIND='NO' 'HISTORY' 'PAST' 'OF' 'DIARRHEA' )
        )
        (VERB='EXIST' 'PAST' )
     )
     (
        (RELATION='COMPATIBLE' 'WITH' )
        (FRAGMENT=
           (NSTG='DIARRHEA' )
        )
        (FRAGMENT=
           (NSTG='MENINGITIS' )
        )
     )
  )
  (
     (REL-CLAUSE='TRN-FILLIN' )
     (ASSERTION=
        (SUBJECT=
           (NEG-FIND='NO' 'HISTORY' 'PAST' 'OF' 'OTHER' 'FOCAL' 'SIGN' 'PLURAL
' )
```

```
            )
            (VERB='EXIST' 'PAST' )
         )
         (
            (RELATION='COMPATIBLE' 'WITH' )
            (FRAGMENT=
               (NSTG='FOCAL' 'SIGN' 'PLURAL' )
            )
            (FRAGMENT=
               (NSTG='MENINGITIS' )
            )
         )
      )
   )
```

```
* HIDSM    1. 1.16
* CHILD WAS 5 LBS. 1 OZ. .

 (ASSERTION=
    (SUBJECT='CHILD' )
    (VERB='BE' 'PAST' )
    (OBJECT='5' 'POUND' 'PLURAL' '1' 'OUNCE' )
 )
```

```
* HIDSM    1. 1.18
* GROWTH AND DEVELOPMENT W.N.L. .

 (ASSERTION=
    (SUBJECT='GROWTH' 'AND' 'DEVELOPMENT' )
    (VERB='BE' )
    (OBJECT='WITHIN' 'NORMAL' 'LIMITS' )
 )
```

```
* HIDSM    1. 1.19
* PATIENT HAD NO PREVIOUS HOSPITALIZATIONS
* .

 (ASSERTION=
    (SUBJECT='PATIENT' )
    (VERB='HAVE' 'PAST' )
    (OBJECT=
       (NEG-FIND='NO' 'PREVIOUS' 'HOSPITALIZATION' 'PLURAL' )
    )
 )
```

```
* HIDSM    1. 1.22
* NO HX OF BLEEDING .
```

```
(FRAGMENT=
   (NSTG=
      (NEG-FIND='NO' 'HISTORY' 'PAST' 'OF' 'BLEEDING' )
   )
)
```

```
* HIDSM    2. 1. 2
* SHORTLY AFTER THE ONSET OF THE HEADACHE ,
* VOMITING ENSUED WITH 8 EPISODES BY HISTORY
* .

(
   (REL-CLAUSE='T-EXPAND-REFPT' )
   (ASSERTION=
      (SA='SHORTLY' 'AFTER' 'ONSET' )
      (SUBJECT='VOMITING' )
      (VERB='ENSUE' 'PAST' )
      (SA='WITH' '8' 'EPISODE' 'PLURAL' 'BY' 'HISTORY' 'PAST' )
   )
   (FRAGMENT=
      (NSTG='THE' 'ONSET' 'OF' 'THE' 'HEADACHE' ',' )
   )
)
```

```
* HIDSM    2. 1. 5
* HE NOTED THE ONSET OF A STIFF NECK AND A
* RASH THAT DAY .

(
   (-CONJOINED='AND' )
   (ASSERTION=
      (SUBJECT='HE' )
      (VERB='NOTE' 'PAST' )
      (OBJECT='THE' 'ONSET' 'OF' 'A' 'STIFF' 'NECK' 'NSTGT-TIME' 'THAT' 'DAY' )
   )
   (ASSERTION=
      (SUBJECT='HE' )
      (VERB='NOTE' 'PAST' )
      (OBJECT='THE' 'ONSET' 'OF' 'A' 'RASH' 'NSTGT-TIME' 'THAT' 'DAY' )
   )
)
```

```
* HIDSM    2. 1. 7
* THE PATIENT APPARENTLY HAD A URI LAST WEEK
* WHICH RESOLVED .
```

```
 (
    (REL-CLAUSE='TRN-WH' )
    (ASSERTION=
       (SUBJECT='THE' 'PATIENT' )
       (SA='APPARENTLY' )
       (VERB='HAVE' 'PAST' )
       (OBJECT='A' 'URI' 'NSTGT-TIME' 'LAST' 'WEEK' )
    )
    (ASSERTION=
       (SUBJECT='URI' )
       (VERB='RESOLVE' 'PAST' )
    )
 )


* HIDSM    2. 1. 8
* ON TRANSFER TO THE FLOOR , HE WAS LETHARGIC
* AND COMPLAINING OF A VERY SEVERE STIFF NECK
* .

 (
    (-CONJOINED='AND' )
    (
       (REL-CLAUSE='T-EXPAND-REFPT' )
       (ASSERTION=
          (SA='ON' 'TRANSFER' )
          (SUBJECT='HE' )
          (VERB='BE' 'PAST' )
          (OBJECT='LETHARGIC' )
       )
       (FRAGMENT=
          (NSTG='TRANSFER' 'TO' 'THE' 'FLOOR' ',' )
       )
    )
    (
       (REL-CLAUSE='T-EXPAND-REFPT' )
       (ASSERTION=
          (SA='ON' 'TRANSFER' )
          (SUBJECT='HE' )
          (VERB='COMPLAIN' 'PAST' 'PROG' )
          (OBJECT='OF'
            (QUANT='A' 'VERY' 'SEVERE' 'STIFF' 'NECK' )
          )
       )
       (FRAGMENT=
          (NSTG='TRANSFER' 'TO' 'THE' 'FLOOR' ',' )
       )
    )
 )


* HIDSM    2. 1.13
* HE WAS ADMITTED TO ORANGE COUNTY HOSPITAL
```

```
* IN ORLANDO , FLORIDA AND DIAGNOSED BY SPINAL
* TAP TO HAVE A MENINGITIS ; ETIOLOGY NOT QUITE
* CLEAR .

 (
    (
       (-CONJOINED='AND' )
       (ASSERTION=
          (VERB='ADMIT' 'PAST' )
          (OBJECT='HE' 'TO' 'ORANGE' 'COUNTY' 'HOSPITAL' 'IN' 'ORLANDO' ',' 'FLO
RIDA' )
       )
       (
          (EMBEDDED='EMBEDDED-OBJ' )
          (ASSERTION=
             (VERB='DIAGNOSE' 'PAST' 'BY' 'SPINAL' 'TAP' )
          )
          (ASSERTION=
             (SUBJECT='HE' )
             (VERB='HAVE' )
             (OBJECT='A' 'MENINGITIS' )
          )
       )
    )
    (ENDMARK=';' )
    (ASSERTION=
       (SUBJECT='ETIOLOGY' )
       (VERB='BE' )
       (OBJECT=
          (NEG-FIND='NOT' 'QUITE' 'CLEAR' )
       )
    )
 )


* HIDSM    2. 1.14
* THE PATIENT WAS TREATED WITH 10 DAYS OF IV
* ANTIBIOTICS .

 (ASSERTION=
    (VERB='TREAT' 'PAST' )
    (OBJECT='THE' 'PATIENT' 'WITH' '10' 'DAY' 'PLURAL' 'OF' 'INTRAVENOUS' 'ANTIB
IOTIC' 'PLURAL' )
 )


* HIDSM    2. 1.15
* ALLERGIES - NONE KNOWN .
(INTRODUCER='ALLERGY' 'PLURAL' '-' )

 (ASSERTION=
    (VERB='KNOW' )
    (OBJECT=
       (NEG-FIND='NONE' )
    )
```

```
 )


* HIDSM    2. 1.16
* IMMUNIZATIONS - COMPLETE .
(INTRODUCER='IMMUNIZATION' 'PLURAL' '-' )

 (ASSERTION=
    (VERB='BE' )
    (OBJECT='COMPLETE' )
 )


* HIDSM    2. 1.17
* ACUTE INFECTIOUS DISEASE - THE PATIENT HAS
* HAD THE CHICKENPOX , MEASLES AND THE MUMPS
* .
(INTRODUCER='ACUTE' 'INFECTIOUS' 'DISEASE' '-' )

 (
    (CONJOINED=',' )
    (ASSERTION=
       (SUBJECT='THE' 'PATIENT' )
       (VERB='HAVE' 'PRESENT' 'PERF' )
       (OBJECT='THE' 'CHICKENPOX' )
    )
    (-CONJOINED='AND' )
    (ASSERTION=
       (SUBJECT='THE' 'PATIENT' )
       (VERB='HAVE' 'PRESENT' 'PERF' )
       (OBJECT='THE' 'MEASLES' )
    )
    (ASSERTION=
       (SUBJECT='THE' 'PATIENT' )
       (VERB='HAVE' 'PRESENT' 'PERF' )
       (OBJECT='THE' 'MUMPS' )
    )
 )


* HIDSM    2. 1.18
* ROS - NONCONTRIBUTORY .
(INTRODUCER='ROS' '-' )

 (ASSERTION=
    (VERB='BE' )
    (OBJECT=
       (NEG-PREFIX='CONTRIBUTORY' )
    )
 )
```

```
* HIDSM     2. 1.21
* THE PATIENT HAS A 28 YR. OLD SISTER WITH
* RHD .

 (
    (REL-CLAUSE='TRN-FILLIN' )
    (ASSERTION=
       (SUBJECT='THE' 'PATIENT' )
       (VERB='HAVE' 'PRESENT' )
       (OBJECT='A' '28' 'YEAR' 'OLD' 'SISTER' )
    )
    (ASSERTION=
       (SUBJECT='SISTER' )
       (VERB='HAVE' )
       (OBJECT='RHD' )
    )
 )


* HIDSM     2. 1.23
* BROTHER 18 ALSO HAS RHD , ON CARDIAC MEDS
* .

 (ASSERTION=
    (SUBJECT='BROTHER' '18' )
    (SA='ALSO' )
    (VERB='HAVE' 'PRESENT' )
    (OBJECT='RHD' ',' )
    (SA='ON' 'CARDIAC' 'MEDICATION' 'PLURAL' )
 )


* HIDSM     3. 1. 2
* PT. ALSO NOTED PAIN AND BURNING IN HER EYES
* 1 DAY PTA .

 (
    (-CONJOINED='AND' )
    (ASSERTION=
       (SUBJECT='PATIENT' )
       (SA='ALSO' )
       (VERB='NOTE' 'PAST' )
       (OBJECT='PAIN' 'IN' 'HER' 'EYE' 'PLURAL' )
       (SA='NSTGT-TIME' '1' 'DAY' 'PTA' )
    )
    (ASSERTION=
       (SUBJECT='PATIENT' )
       (SA='ALSO' )
       (VERB='NOTE' 'PAST' )
       (OBJECT='BURNING' 'IN' 'HER' 'EYE' 'PLURAL' )
```

A1

```
       (SA='NSTGT-TIME' '1' 'DAY' 'PTA' )
    )
 )


* HIDSM    3. 1. 3
* SHE NOTED ONSET OF FEVER , SEVERE HEADACHE
* AND NECK STIFFNESS .

 (
    (CONJOINED=',' )
    (ASSERTION=
       (SUBJECT='SHE' )
       (VERB='NOTE' 'PAST' )
       (OBJECT='ONSET' 'OF' 'FEVER' )
    )
    (-CONJOINED='AND' )
    (ASSERTION=
       (SUBJECT='SHE' )
       (VERB='NOTE' 'PAST' )
       (OBJECT='ONSET' 'OF' 'SEVERE' 'HEADACHE' )
    )
    (ASSERTION=
       (SUBJECT='SHE' )
       (VERB='NOTE' 'PAST' )
       (OBJECT='ONSET' 'OF' 'NECK' 'STIFFNESS' )
    )
 )


* HIDSM    3. 1. 4
* SOON THEREAFTER SHE HAD TEETH CHATTERING
* CHILLS .

 (ASSERTION=
    (SA='SOON' 'THEREAFTER' )
    (SUBJECT='SHE' )
    (VERB='HAVE' 'PAST' )
    (OBJECT='TOOTH' 'PLURAL' 'CHATTERING' 'CHILL' 'PLURAL' )
 )


* HIDSM    3. 1. 6
* PT. ALSO NOTED PAIN IN VAGINA AND DYSURIA
* .

 (
    (-CONJOINED='AND' )
    (ASSERTION=
       (SUBJECT='PATIENT' )
       (SA='ALSO' )
       (VERB='NOTE' 'PAST' )
```

A1

```
    (OBJECT='PAIN' 'IN' 'VAGINA' )
  )
  (ASSERTION=
    (SUBJECT='PATIENT' )
    (SA='ALSO' )
    (VERB='NOTE' 'PAST' )
    (OBJECT='DYSURIA' )
  )
)


* HIDSM    3. 1.11
* MEDICINES % NONE .
(INTRODUCER='MEDICINE' 'PLURAL' ' )

 (FRAGMENT=
   (NSTG=
     (NEG-FIND='NONE' )
   )
 )


* HIDSM    3. 1.12
* ALLERGIES % NONE .
(INTRODUCER='ALLERGY' 'PLURAL' ' )

 (FRAGMENT=
   (NSTG=
     (NEG-FIND='NONE' )
   )
 )


* HIDSM    3. 1.13
* FAMILY HISTORY - MOTHER AND FATHER ALIVE
* AND WELL .
(INTRODUCER='FAMILY' 'HISTORY' '-' )

 (
   (-CONJOINED='AND' )
   (ASSERTION=
     (SUBJECT='MOTHER' )
     (VERB='BE' )
     (OBJECT='ALIVE' 'AND''WELL' )
   )
   (ASSERTION=
     (SUBJECT='FATHER' )
     (VERB='BE' )
     (OBJECT='ALIVE' 'AND' 'WELL' )
   )
 )
```

```
* HIDSM    3. 1.14
* NO HYPERTENSION , DIABETES , HEART OR KIDNEY
* DISEASE .

 (
    (CONJOINED=',' )
    (FRAGMENT=
       (NSTG=
          (NEG-FIND='NO' 'HYPERTENSION' )
       )
    )
    (CONJOINED=',' )
    (FRAGMENT=
       (NSTG=
          (NEG-FIND='NO' 'DIABETES' )
       )
    )
    (-CONJOINED='AND' )
    (FRAGMENT=
       (NSTG=
          (NEG-FIND='NO' 'HEART' 'DISEASE' )
       )
    )
    (FRAGMENT=
       (NSTG=
          (NEG-FIND='NO' 'KIDNEY' 'DISEASE' )
       )
    )
 )


* HIDSM    3. 1.17
* REVIEW OF SYSTEMS - NONCONTRIBUTORY .
(INTRODUCER='REVIEW' 'OF' 'SYSTEMS' '-' )

 (ASSERTION=
    (VERB='BE' )
    (OBJECT=
       (NEG-PREFIX='CONTRIBUTORY' )
    )
 )


* EXDSM    1. 1. 1
* AN IRRITABLE CHILD , FAT AND CRYING .

 (
    (REL-CLAUSE='TRN-FILLIN' )
    (FRAGMENT=
```

```
      (NSTG='AN' 'IRRITABLE' 'CHILD' ',' )
    )
    (
      (-CONJOINED='AND' )
      (ASSERTION=
        (SUBJECT='CHILD' ',' )
        (VERB='BE' )
        (OBJECT='FAT' )
      )
      (ASSERTION=
        (SUBJECT='CHILD' ',' )
        (VERB='BE' )
        (OBJECT='CRYING' )
      )
    )
 )


* EXDSM     1. 1. 3
* R 25 .

 (FRAGMENT=
    (NSTG='RESPIRATION''RATE' '25' )
 )


* EXDSM     1. 1. 5
* HEAD - NORMAL CIRCUMFERENCE .
(INTRODUCER='HEAD' '-' )

 (FRAGMENT=
    (NSTG='NORMAL' 'CIRCUMFERENCE' )
 )


* EXDSM     1. 1. 6
* NO EVIDENCE OF TRAUMA .

 (FRAGMENT=
    (NSTG=
       (NEG-FIND='NO' 'EVIDENCE' 'OF' 'TRAUMA' )
    )
 )


* EXDSM     1. 1. 7
* ANTERIOR FONTANELLE WAS CLOSED .

 (ASSERTION=
    (VERB='CLOSE' 'PAST' )
    (OBJECT='ANTERIOR' 'FONTANEL' )
```

A1

```
 )


* EXDSM    1. 1. 9
* CONJUNCTIVAE CLEAR .

 (ASSERTION=
    (SUBJECT='CONJUNCTIVA' 'PLURAL' )
    (VERB='BE' )
    (OBJECT='CLEAR' )
 )


* EXDSM    1. 1.11
* MOUTH - DRY WITHOUT EXUDATE .
(INTRODUCER='MOUTH' '-' )

 (
    (PREP-CONN='WITH' )
    (ASSERTION=
       (VERB='BE' )
       (OBJECT='DRY' )
    )
    (FRAGMENT=
       (NSTG=
          (NEG-FIND='NO' 'EXUDATE' )
       )
    )
 )


* EXDSM    1. 1.13
* CHEST - CLEAR .
(INTRODUCER='CHEST' '-' )

 (ASSERTION=
    (VERB='BE' )
    (OBJECT='CLEAR' )
 )


* EXDSM    1. 1.15
* ABDOMEN - SOFT ; NO MASSES .
(INTRODUCER='ABDOMEN' '-' )

 (
    (ASSERTION=
       (VERB='BE' )
       (OBJECT='SOFT' )
    )
    (ENDMARK=';' )
```

```
    (FRAGMENT=
       (NSTG=
          (NEG-FIND='NO' 'MASS' 'PLURAL' )
       )
    )
 )


* EXDSM     1. 1.16
* PULSES IN EXTREMITIES WERE 2+ .

 (ASSERTION=
    (SUBJECT='PULSE' 'PLURAL' 'IN' 'EXTREMITY' 'PLURAL' )
    (VERB='BE' 'PAST' )
    (OBJECT='2+' )
 )


* EXDSM     1. 1.17
* EXTREMITIES NORMAL .

 (ASSERTION=
    (SUBJECT='EXTREMITY' 'PLURAL' )
    (VERB='BE' )
    (OBJECT='NORMAL' )
 )


* EXDSM     1. 1.19
* CNS - AN ALERT , IRRITABLE CHILD WITH A STRONG
* GRASP .
(INTRODUCER='CNS' '-' )

 (
    (REL-CLAUSE='TRN-WH' )
    (FRAGMENT=
       (NSTG='AN' 'ALERT' ',' 'CHILD' )
    )
    (
       (-CONJOINED='AND' )
       (ASSERTION=
          (SUBJECT='CHILD' )
          (VERB='BE' )
          (OBJECT='IRRITABLE' )
       )
       (ASSERTION=
          (SUBJECT='CHILD' )
          (VERB='HAVE' )
          (OBJECT='A' 'STRONG' 'GRASP' )
       )
    )
 )
```

A1

```
* EXDSM     1. 1.20
* MOTOR FUNCTIONS INTACT .

 (ASSERTION=
    (SUBJECT='MOTOR' 'FUNCTION' 'PLURAL' )
    (VERB='BE' )
    (OBJECT='INTACT' )
 )


* EXDSM     1. 1.21
* NO LOCALIZING SIGNS WERE NOTED .

 (ASSERTION=
    (VERB='NOTE' 'PAST' )
    (OBJECT=
       (NEG-FIND='NO' 'LOCALIZING' 'SIGN' 'PLURAL' )
    )
 )


* EXDSM     1. 1.22
* CRANIAL NERVES INTACT .

 (ASSERTION=
    (SUBJECT='CRANIAL' 'NERVE' 'PLURAL' )
    (VERB='BE' )
    (OBJECT='INTACT' )
 )


* EXDSM     1. 1.23
* REFLEXES PRESENT RATED 2+ .

 (
    (REL-CLAUSE='TRN-FILLIN' )
    (ASSERTION=
       (SUBJECT='REFLEX' 'PLURAL' )
       (VERB='RATE' 'PAST' )
       (OBJECT='2+' )
    )
    (ASSERTION=
       (SUBJECT='REFLEX' 'PLURAL' )
       (VERB='BE' )
       (OBJECT='PRESENT' )
    )
 )
```

```
* EXDSM    2. 1. 4
* HEENT SHOWED A HEAD WHICH HAD RESIDUAL FACIAL
* SCARS AROUND BOTH EYES .

 (
    (RELATION='SHOW' 'PAST' )
    (FRAGMENT=
       (NSTG='HEENT' )
    )
    (
       (REL-CLAUSE='TRN-WH' )
       (FRAGMENT=
          (NSTG='A' 'HEAD' )
       )
       (
          (REL-CLAUSE='TRN-WH' )
          (ASSERTION=
             (SUBJECT='HEAD' )
             (VERB='HAVE' 'PAST' )
             (OBJECT='RESIDUAL' 'SCAR' 'PLURAL' 'AROUND' 'BOTH' 'EYE' 'PLURAL' )
          )
          (ASSERTION=
             (SUBJECT='SCAR' 'PLURAL' )
             (VERB='BE' )
             (OBJECT='FACIAL' )
          )
       )
    )
 )


* EXDSM    2. 1. 5
* EARS % TM'S WERE CLEAR .
(INTRODUCER='EAR' 'PLURAL' ' )

 (ASSERTION=
    (SUBJECT='TYMPANIC''MEMBRANE' 'PLURAL' )
    (VERB='BE' 'PAST' )
    (OBJECT='CLEAR' )
 )


* EXDSM    2. 1. 6
* EYES % C AND S WERE WNL .
(INTRODUCER='EYE' 'PLURAL' ' )

 (ASSERTION=
    (SUBJECT='C' 'AND' 'S' 'PLURAL' )
    (VERB='BE' 'PAST' )
    (OBJECT='WITHIN' 'NORMAL' 'LIMITS' )
```

```
 )


* EXDSM    2. 1. 7
* HE DID HAVE WHAT APPEARED TO BE A MILD BILATERAL
* PTOSIS .

 (
    (REL-CLAUSE='TRN-WH' )
    (ASSERTION=
       (SUBJECT='HE' )
       (VERB='HAVE' 'PAST' )
       (OBJECT='THAT' )
    )
    (
       (MODAL='APPEAR' 'PAST')
       (
           (REL-CLAUSE='TRN-WH' )
           (ASSERTION=
              (SUBJECT='THAT' )
              (VERB='BE' )
              (OBJECT='A' 'MILD' 'PTOSIS' )
           )
           (ASSERTION=
              (SUBJECT='PTOSIS' )
              (VERB='BE' )
              (OBJECT='BILATERAL' )
           )
        )
    )
 )


* EXDSM    2. 1. 8
* PERRLA .

 (FRAGMENT=
    (NSTG='PERRLA' )
 )


* EXDSM    2. 1.10
* NOSE % SEPTUM WAS MIDLINE , WITHOUT PERFORATIONS
* .
(INTRODUCER='NOSE' ' )

 (
    (PREP-CONN='WITH' )
    (ASSERTION=
       (SUBJECT='SEPTUM' )
       (VERB='BE' 'PAST' )
       (OBJECT='MIDLINE' ',' )
```

A1

```
    )
    (FRAGMENT=
       (NSTG=
          (NEG-FIND='NO' 'PERFORATION' 'PLURAL' )
       )
    )
 )
```

```
* EXDSM    2. 1.14
* THE PATIENT HAD NO JVD AND NO INCREASE IN
* THYROID SIZE .

 (
    (-CONJOINED='AND' )
    (ASSERTION=
       (SUBJECT='THE' 'PATIENT' )
       (VERB='HAVE' 'PAST' )
       (OBJECT=
          (NEG-FIND='NO' 'JVD' )
       )
    )
    (ASSERTION=
       (SUBJECT='THE' 'PATIENT' )
       (VERB='HAVE' 'PAST' )
       (OBJECT=
          (NEG-FIND='NO'
             (QUANT='INCREASE' 'IN' 'THYROID' 'SIZE' )
          )
       )
    )
 )
```

```
* EXDSM    2. 1.15
* CHEST CLEAR TO P AND A .

 (ASSERTION=
    (SUBJECT='CHEST' )
    (VERB='BE' )
    (OBJECT='CLEAR' 'TO' 'P' 'AND' 'A' )
 )
```

```
* EXDSM    2. 1.20
* EXTREMITIES WERE WITHOUT CCE .

 (ASSERTION=
    (SUBJECT='EXTREMITY' 'PLURAL' )
    (VERB='BE' 'PAST' )
    (OBJECT='WITH'
       (NEG-FIND='NO' 'CCE' )
```

A1

```
    )
 )

* EXDSM    2. 1.22
* MOTOR WAS GROSSLY WNL .

 (ASSERTION=
    (SUBJECT='MOTOR' )
    (VERB='BE' 'PAST' )
    (OBJECT='GROSSLY' 'WITHIN' 'NORMAL''LIMITS' )
 )

* EXDSM    3. 1. 2
* TEMP. 103.2 .

 (FRAGMENT=
    (NSTG='TEMPERATURE' '103.2' )
 )

* EXDSM    3. 1. 3
* BP 130 / 80 .

 (FRAGMENT=
    (NSTG='BLOOD''PRESSURE' '130' '/' '80' )
 )

* EXDSM    3. 1. 4
* PULSE 100 .

 (FRAGMENT=
    (NSTG='PULSE' '100' )
 )

* EXDSM    3. 1. 5
* RESP. 14 .

 (FRAGMENT=
    (NSTG='RESPIRATION' '14' )
 )

* EXDSM    3. 1. 8
```

```
* NOSE % NO RHINORRHEA OR CSF DISCHARGE .
(INTRODUCER='NOSE' ' )

 (
    (-CONJOINED='AND' )
    (FRAGMENT=
       (NSTG=
          (NEG-FIND='NO' 'RHINORRHEA' )
       )
    )
    (FRAGMENT=
       (NSTG=
          (NEG-FIND='NO' 'CEREBROSPINAL' 'FLUID' 'DISCHARGE' )
       )
    )
 )


* EXDSM    3. 1. 9
* THROAT NO INJECTION OR PETECHIAE .

 (
    (-CONJOINED='AND' )
    (ASSERTION=
       (SUBJECT='THROAT' )
       (VERB='SHOW' )
       (OBJECT=
          (NEG-FIND='NO' 'INJECTION' )
       )
    )
    (ASSERTION=
       (SUBJECT='THROAT' )
       (VERB='SHOW' )
       (OBJECT=
          (NEG-FIND='NO' 'PETECHIAE' 'PLURAL' )
       )
    )
 )


* EXDSM    3. 1.13
* CHEST % CLEAR TO PMI .
(INTRODUCER='CHEST' ' )

 (ASSERTION=
    (VERB='BE' )
    (OBJECT='CLEAR' 'TO' 'PMI' )
 )


* EXDSM    3. 1.16
* BACK % NO CVA TENDERNESS .
```

```
(INTRODUCER='BACK' ' )

 (FRAGMENT=
    (NSTG=
       (NEG-FIND='NO' 'COSTOVERTEBRAL' 'ANGLE' 'TENDERNESS' )
    )
 )


* EXDSM    3. 1.17
* EXTREMITIES % NO CLUBBING , CYANOSIS OR EDEMA
* , NO CALF TENDERNESS , NO PETECHIAE .
(INTRODUCER='EXTREMITY' 'PLURAL' ' )

 (
    (CONJOINED=',' )
    (FRAGMENT=
       (NSTG=
          (NEG-FIND='NO' 'CLUBBING' )
       )
    )
    (CONJOINED='AND' )
    (FRAGMENT=
       (NSTG=
          (NEG-FIND='NO' 'CYANOSIS' )
       )
    )
    (CONJOINED=',' )
    (FRAGMENT=
       (NSTG=
          (NEG-FIND='NO' 'EDEMA' )
       )
    )
    (-CONJOINED=',' )
    (FRAGMENT=
       (NSTG=
          (NEG-FIND='NO' 'CALF' 'TENDERNESS' )
       )
    )
    (FRAGMENT=
       (NSTG=
          (NEG-FIND='NO' 'PETECHIAE' 'PLURAL' )
       )
    )
 )


* EXDSM    3. 1.19
* BREASTS % NO MASSES .
(INTRODUCER='BREAST' 'PLURAL' ' )

 (FRAGMENT=
    (NSTG=
```

A1

```
        (NEG-FIND='NO' 'MASS' 'PLURAL' )
     )
  )


* EXDSM    3. 1.23
* NO CEREBELLAR SIGNS .

 (FRAGMENT=
     (NSTG=
        (NEG-FIND='NO' 'CEREBELLAR' 'SIGN' 'PLURAL' )
     )
 )


* EXDSM    3. 1.24
* DTR'S BRISK AND EQUAL .

 (
     (-CONJOINED='AND' )
     (ASSERTION=
        (SUBJECT='DTRS' 'PLURAL' )
        (VERB='BE' )
        (OBJECT='BRISK' )
     )
     (ASSERTION=
        (SUBJECT='DTRS' 'PLURAL' )
        (VERB='BE' )
        (OBJECT='EQUAL' )
     )
 )


* LDDSM    1. 1. 1
* SPINAL TAP SHOWED 860 RBC'S , 230 WBC'S ;
* PROTEIN 25 ; GLUCOSE 72 .

 (
     (
        (-CONJOINED=',' )
        (
           (RELATION='SHOW' 'PAST' )
           (FRAGMENT=
              (NSTG='SPINAL' 'TAP' )
           )
           (FRAGMENT=
              (NSTG='860' 'RED' 'BLOOD' 'CELL' 'PLURAL' )
           )
        )
        (
```

```
          (RELATION='SHOW' 'PAST' )
          (FRAGMENT=
             (NSTG='SPINAL' 'TAP' )
          )
          (FRAGMENT=
             (NSTG='230' 'WHITE' 'BLOOD' 'CELL' 'PLURAL' )
          )
       )
    )
    (ENDMARK=';' )
    (FRAGMENT=
       (NSTG='PROTEIN' '25' )
    )
    (ENDMARK=';' )
    (FRAGMENT=
       (NSTG='GLUCOSE' '72' )
    )
 )


* LDDSM    1. 1. 2
* ELECTROLYTES - SODIUM 138 , POTASSIUM 4.5
* , CHLORIDE 102 , BUN 10 , CREAT. .8 , GLUCOSE
* OF 100 , CALCIUM 10.6 .
(INTRODUCER='ELECTROLYTE' 'PLURAL' '-' )

 (
    (CONJOINED=',' )
    (FRAGMENT=
       (NSTG='SODIUM' '138' )
    )
    (CONJOINED=',' )
    (FRAGMENT=
       (NSTG='POTASSIUM' '4.5' )
    )
    (CONJOINED=',' )
    (FRAGMENT=
       (NSTG='CHLORIDE' '102' )
    )
    (CONJOINED=',' )
    (FRAGMENT=
       (NSTG='BUN' '10' )
    )
    (CONJOINED=',' )
    (FRAGMENT=
       (NSTG='CREATININE' '.8' )
    )
    (-CONJOINED=',' )
    (FRAGMENT=
       (NSTG='GLUCOSE' 'OF' '100' )
    )
    (FRAGMENT=
       (NSTG='CALCIUM' '10.6' )
    )
```

```
 )


* LDDSM     2. 1. 4
* CSF CULTURES X 4 SHOWED NO MICROORGANISMS
* .

 (ASSERTION=
    (SUBJECT='CEREBROSPINAL' 'FLUID' 'CULTURE' 'PLURAL' )
    (SA='X4' )
    (VERB='SHOW' 'PAST' )
    (OBJECT=
       (NEG-FIND='NO' 'MICROORGANISM' 'PLURAL' )
    )
 )


* LDDSM     2. 1. 5
* GRAM STAIN NEG .

 (ASSERTION=
    (SUBJECT='GRAM' 'STAIN' )
    (VERB='BE' )
    (OBJECT=
       (NEG-FIND='NEGATIVE' )
    )
 )


* LDDSM     2. 1. 6
* BLOOD CULTURES X 4 WERE ALL NO GROWTH .

 (ASSERTION=
    (SUBJECT='BLOOD' 'CULTURE' 'PLURAL' )
    (SA='X4' )
    (VERB='BE' 'PAST' )
    (SA='ALL' )
    (OBJECT=
       (NEG-FIND='NO' 'GROWTH' )
    )
 )


* LDDSM     2. 1. 7
* THROAT CULTURE WAS NEG .

 (ASSERTION=
    (SUBJECT='THROAT' 'CULTURE' )
    (VERB='BE' 'PAST' )
    (OBJECT=
       (NEG-FIND='NEGATIVE' )
```

```
    )
 )


* LDDSM    2. 1.11
* BILE NEG .

 (ASSERTION=
    (SUBJECT='BILE' )
    (VERB='BE' )
    (OBJECT=
       (NEG-FIND='NEGATIVE' )
    )
 )


* LDDSM    2. 1.12
* UROBILI NEG .

 (ASSERTION=
    (SUBJECT='UROBILIN' )
    (VERB='BE' )
    (OBJECT=
       (NEG-FIND='NEGATIVE' )
    )
 )


* LDDSM    2. 1.13
* PROTEIN NEG .

 (ASSERTION=
    (SUBJECT='PROTEIN' )
    (VERB='BE' )
    (OBJECT=
       (NEG-FIND='NEGATIVE' )
    )
 )


* LDDSM    2. 1.14
* 0 RBC'S AND 3 - 4 WBC'S .

 (
    (-CONJOINED='AND' )
    (FRAGMENT=
       (NSTG='0' 'RED' 'BLOOD' 'CELL' 'PLURAL' )
    )
    (FRAGMENT=
       (NSTG='3' 'WHITE''BLOOD' 'CELL' 'PLURAL' '-' '4' 'WHITE' 'BLOOD' 'CELL' '
PLURAL' )
    )
```

```
 )



* LDDSM    3. 1. 2
* ABDOMEN % NORMAL .
(INTRODUCER='ABDOMEN' ' )

 (FRAGMENT=
    (NSTG='NORMAL' )
 )


* LDDSM    3. 1. 3
* HCT. 39.3 .

 (FRAGMENT=
    (NSTG='HEMATOCRIT' '39.3' )
 )


* LDDSM    3. 1. 4
* MCV 79 .

 (FRAGMENT=
    (NSTG='MCV' '79' )
 )


* LDDSM    3. 1. 5
* ESR 35 .

 (FRAGMENT=
    (NSTG='ESR' '35' )
 )


* LDDSM    3. 1. 6
* WBC 12.3 - 75 SEGS , 5 BANDS , 12 LYMPHS
* .

 (
    (CONJOINED=',' )
    (FRAGMENT=
       (NSTG='WHITE' 'BLOOD' 'CELL' '12.3' 'SEGMENT' 'PLURAL' '-' '75' 'SEGMENT'
 'PLURAL' )
    )
    (-CONJOINED=',' )
    (FRAGMENT=
       (NSTG='5' 'BAND' 'PLURAL' )
```

```
    )
    (FRAGMENT=
       (NSTG='12' 'LYMPHOCYTE' 'PLURAL' )
    )
 )


* LDDSM    3. 1. 8
* ELECTROLYTES % NA 144 , K 3.5 , CO2 19 ,
* CL 107 , BUN 13 , CREATININE 1.2 .
(INTRODUCER='ELECTROLYTE' 'PLURAL' ' )

 (
    (CONJOINED=',' )
    (FRAGMENT=
       (NSTG='SODIUM' '144' )
    )
    (CONJOINED=',' )
    (FRAGMENT=
       (NSTG='POTASSIUM' '3.5' )
    )
    (CONJOINED=',' )
    (FRAGMENT=
       (NSTG='CO2' '19' )
    )
    (CONJOINED=',' )
    (FRAGMENT=
       (NSTG='CHLORIDE' '107' )
    )
    (-CONJOINED=',' )
    (FRAGMENT=
       (NSTG='BUN' '13' )
    )
    (FRAGMENT=
       (NSTG='CREATININE' '1.2' )
    )
 )


* LDDSM    3. 1. 9
* PRO TIME 10.5 / 11.5 .

 (FRAGMENT=
    (NSTG='PROTHROMBIN''TIME' '10.5' '/' '11.5' )
 )


* LDDSM    3. 1.10
* GLUCOSE 120 .

 (FRAGMENT=
    (NSTG='GLUCOSE' '120' )
```

```
 )


* LDDSM    3. 1.11
* AMYLASE 105 .

 (FRAGMENT=
    (NSTG='AMYLASE' '105' )
 )


* LDDSM    3. 1.12
* CSF % CLOUDY ; NO XANTHOCHROMIA NOTED ; 1450
* WBC'S , 78 POLYS , 20 TO 30 MONOS ; 0 RBC'S
* ; PROTEIN 134 ; GLUCOSE 58 .
(INTRODUCER='CEREBROSPINAL' 'FLUID' ' )

 (
    (ASSERTION=
       (VERB='BE' )
       (OBJECT='CLOUDY' )
    )
    (ENDMARK=';' )
    (ASSERTION=
       (VERB='NOTE' )
       (OBJECT=
          (NEG-FIND='NO' 'XANTHOCHROMIA' )
       )
    )
    (ENDMARK=';' )
    (
       (CONJOINED=',' )
       (FRAGMENT=
          (NSTG='1450' 'WHITE''BLOOD' 'CELL' 'PLURAL' )
       )
       (-CONJOINED=',' )
       (FRAGMENT=
          (NSTG='78' 'POLY' 'PLURAL' )
       )
       (FRAGMENT=
          (NSTG='20' 'MONOCYTE' 'PLURAL' 'TO' '30' 'MONOCYTE' 'PLURAL' )
       )
    )
    (ENDMARK=';' )
    (FRAGMENT=
       (NSTG='0' 'RED' 'BLOOD' 'CELL' 'PLURAL' )
    )
    (ENDMARK=';' )
    (FRAGMENT=
       (NSTG='PROTEIN' '134' )
    )
    (ENDMARK=';' )
    (FRAGMENT=
```

```
        (NSTG='GLUCOSE' '58' )
     )
  )


* LDDSM    3. 1.15
* NOTHING EVER GREW OUT ON CULTURES OF CSF
* OR BLOOD .

  (
     (-CONJOINED='OR' )
     (ASSERTION=
        (SUBJECT=
           (NEG-PREFIX='THING' )
        )
        (VERB='EVER' 'GROW' 'OUT' 'PAST' )
        (SA='ON' 'CULTURE' 'PLURAL' 'OF' 'CEREBROSPINAL' 'FLUID' )
     )
     (ASSERTION=
        (SUBJECT=
           (NEG-PREFIX='THING' )
        )
        (VERB='EVER' 'GROW' 'OUT' 'PAST' )
        (SA='ON' 'CULTURE' 'PLURAL' 'OF' 'BLOOD' )
     )
  )


* LDDSM    3. 1.18
* WBC WENT UP TO 1850 ; 80 POLYS , 20 LYMPHS
* .

  (
     (ASSERTION=
        (SUBJECT='WHITE' 'BLOOD' 'CELL' )
        (VERB='GO' 'PAST' )
        (OBJECT='UP' 'TO' '1850' )
     )
     (ENDMARK=';' )
     (
        (-CONJOINED=',' )
        (FRAGMENT=
           (NSTG='80' 'POLY' 'PLURAL' )
        )
        (FRAGMENT=
           (NSTG='20' 'LYMPHOCYTE' 'PLURAL' )
        )
     )
  )


* LDDSM    3. 1.19
```

```
* PROTEIN 110 .

 (FRAGMENT=
    (NSTG='PROTEIN' '110' )
 )



* LDDSM    3. 1.20
* GLUCOSE 48 AT THAT TIME .

 (ASSERTION=
    (SUBJECT='GLUCOSE' )
    (VERB='BE' )
    (OBJECT='48' 'AT' 'THAT' 'TIME' )
 )


* LDDSM    3. 1.22
* PROTEIN WAS 14 AT THAT TIME , GLUCOSE WAS
* 75 .

 (
    (-CONJOINED=',' )
    (ASSERTION=
       (SUBJECT='PROTEIN' )
       (VERB='BE' 'PAST' )
       (OBJECT='14' 'AT' 'THAT' 'TIME' )
    )
    (ASSERTION=
       (SUBJECT='GLUCOSE' )
       (VERB='BE' 'PAST' )
       (OBJECT='75' )
    )
 )



* IMDSM    2. 1. 1
* NONE GIVEN .

 (ASSERTION=
    (VERB='GIVE' )
    (OBJECT=
       (NEG-FIND='NONE' )
    )
 )



* CODSM    1. 1. 3
* THE CHILD WAS ON CHLORAMPHENICOL 100 MGM.
```

```
* PER KILOGRAM P.O. .

 (ASSERTION=
    (SUBJECT='THE' 'CHILD' )
    (VERB='BE' 'PAST' )
    (OBJECT='ON' 'CHLORAMPHENICOL' '100' 'MILLIGRAM' 'PER' 'KILOGRAM' 'PO' )
 )


* CODSM    1. 1. 5
* ALL CULTURES INCLUDING URINE , BLOOD , THROAT
* , CSF WERE NORMAL .

 (
    (REL-CLAUSE='TRN-FILLIN' )
    (ASSERTION=
       (SUBJECT='ALL' 'CULTURE' 'PLURAL' )
       (VERB='BE' 'PAST' )
       (OBJECT='NORMAL' )
    )
    (
       (CONJOINED=',' )
       (ASSERTION=
          (SUBJECT='CULTURE' 'PLURAL' )
          (VERB='INCLUDE' 'PROG' )
          (OBJECT='URINE' )
       )
       (CONJOINED=',' )
       (ASSERTION=
          (SUBJECT='CULTURE' 'PLURAL' )
          (VERB='INCLUDE' 'PROG' )
          (OBJECT='BLOOD' )
       )
       (-CONJOINED=',' )
       (ASSERTION=
          (SUBJECT='CULTURE' 'PLURAL' )
          (VERB='INCLUDE' 'PROG' )
          (OBJECT='THROAT' )
       )
       (ASSERTION=
          (SUBJECT='CULTURE' 'PLURAL' )
          (VERB='INCLUDE' 'PROG' )
          (OBJECT='CEREBROSPINAL' 'FLUID' )
       )
    )
 )


* CODSM    1. 1. 6
* THE CHILD WAS TREATED AS A PARTIALLY TREATED
* MENINGITIS .

 (
```

A1

```
    (EMBEDDED='EMBEDDED-OBJ' )
    (ASSERTION=
       (VERB='TREAT' 'PAST' )
    )
    (
       (REL-CLAUSE='TRN-FILLIN' )
       (ASSERTION=
          (SUBJECT='THE' 'CHILD' )
          (VERB='AS' )
          (OBJECT='A' 'MENINGITIS' )
       )
       (ASSERTION=
          (VERB='PARTIALLY' 'TREAT' )
          (OBJECT='MENINGITIS' )
       )
    )
 )


* CODSM    2. 1. 8
* THE PATIENT HAD A BRAIN SCAN WHICH WAS NEG
* .

 (
    (REL-CLAUSE='TRN-WH' )
    (ASSERTION=
       (SUBJECT='THE' 'PATIENT' )
       (VERB='HAVE' 'PAST' )
       (OBJECT='A' 'BRAIN' 'SCAN' )
    )
    (ASSERTION=
       (SUBJECT='SCAN' )
       (VERB='BE' 'PAST' )
       (OBJECT=
          (NEG-FIND='NEGATIVE' )
       )
    )
 )


* STDSM    1. 1. 1
* THE CHILD WAS DISCHARGED ON 1 / 31 / 00 IN CARE
* OF THE PARENTS .

 (ASSERTION=
    (VERB='DISCHARGE' 'PAST' 'ON' '1' '/' '31' '/' '00' )
    (OBJECT='THE' 'CHILD' )
    (RV='IN' 'CARE' 'OF' 'THE' 'PARENT' 'PLURAL' )
 )
```

```
* STDSM    1. 1. 2
* THE CHLOROMYCETIN WAS DISCONTINUED .

 (ASSERTION=
    (VERB='DISCONTINUE' 'PAST' )
    (OBJECT='THE' 'CHLOROMYCETIN' )
 )


* STDSM    2. 1. 1
* ON TRANSFER TO UNIVERSITY HOSPITAL THE PATIENT
* WAS IN GOOD CONDITION .

 (
    (REL-CLAUSE='T-EXPAND-REFPT' )
    (ASSERTION=
       (SA='ON' 'TRANSFER' )
       (SUBJECT='THE' 'PATIENT' )
       (VERB='BE' 'PAST' )
       (OBJECT='IN' 'GOOD' 'CONDITION' )
    )
    (FRAGMENT=
       (NSTG='TRANSFER' 'TO' 'UNIVERSITY''HOSPITAL' )
    )
 )


* STDSM    3. 1. 1
* PT. WAS WELL .

 (ASSERTION=
    (SUBJECT='PATIENT' )
    (VERB='BE' 'PAST' )
    (OBJECT='WELL' )
 )


* PLDSM    1. 1. 1
* NONE GIVEN .

 (ASSERTION=
    (VERB='GIVE' )
    (OBJECT=
       (NEG-FIND='NONE' )
    )
 )


* PLDSM    2. 1. 1
* NONE GIVEN .
```

A1

```
(ASSERTION=
   (VERB='GIVE' )
   (OBJECT=
      (NEG-FIND='NONE' )
   )
)
```

```
* RXDSM    2. 1. 1
* RECURRENT MENINGITIS .

(FRAGMENT=
   (NSTG='RECURRENT' 'MENINGITIS' )
)
```

```
* RXDSM    3. 1. 1
* NONE GIVEN .

(ASSERTION=
   (VERB='GIVE' )
   (OBJECT=
      (NEG-FIND='NONE' )
   )
)
```

```
* RADSM    2. 1. 1
* NONE GIVEN .

(ASSERTION=
   (VERB='GIVE' )
   (OBJECT=
      (NEG-FIND='NONE' )
   )
)
```

```
* RADSM    3. 1. 1
* NONE GIVEN .

(ASSERTION=
   (VERB='GIVE' )
   (OBJECT=
      (NEG-FIND='NONE' )
   )
)
```

```
* DODSM    2. 1. 1
* DR. TAYLOR .

 (FRAGMENT=
    (NSTG='DR.' 'TAYLOR' )
 )


* DODSM    3. 1. 1
* DR. PETER LEFKOW .

 (FRAGMENT=
    (NSTG='DR.' 'PETER' 'LEFKOW' )
 )
```

APPENDIX A2

Transformational Decomposition of Theoretical Texts

Michiko Kosaka

These sentences are drawn from the experimental literature on lipoprotein kinetics mainly from BKG (Brown, M.S., Kovanen, P.T., Goldstein, J.L. "Regulation of Plasma Cholesterol by Lipoprotein Receptors". 1981, SCIENCE vol. 212) and some from STM (Sniderman, A., Thomas, D., Marpole, D., Teng, B. "Low Density Lipoprotein: A Metabolic Pathway for Return of Cholesterol to the Splanchnic Bed". 1977, J. Clin. Invest.).

The decompositions were manually done, using standard English transformations as much as possible. The following notational conventions are used in the representation.

1. Operators with their arguments are enclosed in parens; operators precede their ordered arguments (Polish notation).

2. Operators are delimited by slashes (but the slash can be eliminated if another delimiter - parenthesis or the asterisk, as in 3., below -- is present and no ambiguity results).

EX. Cholesterol ester is formed in HDL.

(be formed in/ CE/ HDL)

carry the main logical structure of the sentence are distinguished from local modifiers or "adjuncts!"

3. We distinguish operators, that are not local modifiers from adjuncts. These operators make up the logical structure of the sentence. Adjuncts are, on the whole, associated properties, not part of the main hierarchy of operators. An adjunct is preceded by a *. When the adjunct itself has operator argument structure, it is placed in parens (with the same internal structure as any operator) with the * preceding the parens. Adjuncts of the noun are placed next to the noun; adjuncts of the verb (or whole sentence) are placed after the verb.

EX. The remnants are immediately internalized...

(be internalized *immediately
/remnants)...

EX. The cholesterol ester formed in HDL...

...CE *(be formed in/ CE/ HDL)

4.Multiple references to a given entity (whether noun or sentence part) are indexed with the same subscript n, represented on the keyboard by $n. This is followed by the sentence identification number of the first occurrence of the entity unless the referencing is within the same sentence. If the reference is to a stretch of one or more whole sentences, the sentence numbers can be strung out with + signs between them.

EX. Plasma HDL are thought to accumulate cholesterol from peripheral tissue.

(are thought(INV-accumulate/CH$1 *(be from/ CH$1/ PT)
/HDL$1 *(be in/ HDL$1/ PLASMA)))

Note 1: "INV-" means that arguments are to be read in reverse order; e.g. INV-accumulate/ CH/ HDL says that HDL accumulates CH.

Note 2: plasma HDL → HDL in plasma → HDL be in plasma.

The following abbreviations are used:

NOUN REDUCTIONS
****

APO	apoproteins	[BKG 2.10.2]
APOB	apoprotein-B	[UBS 15]
APOE	apoprotein-E	[BKG 2.5.2]
AT	adipose tissues	[BKG 2.8.1]
BA	bile acids	[BKG 2.7.1]
CAP	capillaries	[BKG 2.9.2]
CAPW	capillary wall	[BKG 2.5.1]
CBH	carbohydrates	[BKG 2.8.1]
CE	cholesterol ester	[UBS 3]
CH	cholesterol	[UBS 2]
CIR	circulation	[BKG 2.9.4]
CY	chylomicrons	[UBS 15]
DIET	diet	[BKG]
EXHT	extrahepatic tissues	[UBS 10]
FA	fatty acids	[BKG 2.8.1]
FC	unesterified cholesterol (free cholesterol)	[BKG 2.7.1]
HDL	high density lipoprotein	[UBS 2]
IDL	intermediate density lipoprotein	[BKG 2.9.5]
INSLN	insulin	
INT	intestine	[UBS 12]
LI	lipids	[UBS 14]
LCAT	lecithin cholesterol acyltransferase	[BKG 2.9.6]
LDL	low density lipoprotein	[UBS 3]
LE	lecithin	[BKG 2.9.7]

LIVER	liver	[UBS 10]
LP	lipoprotein	[UBS 11]
LPL	lipoprotein lipase	[BKG 2.9.2]
PH	phospholipids	[BKG 2.4.8]
PL	plasma	[UBS 9]
PT	peripheral tissues	[UBS 2]
RC	receptors	[UBS 10]
RC-LDL	LDL receptors	
RC-LP	lipoprotein receptors	[UBS 12]
REM	(chylomicron) remnants	[BKG 2.5.2]
TG	triglyceride	[UBS 3]
TSYSTEM	transport system	[BKG 2.14.2]
TSYSTEM-LP	lipoprotein transport system	[BKG 2.4.9]
VLDL	very low density lipoprotein	[UBS 3]
VLDL-B	very low density lipoprotein-B	[UBS 1]

In the analyses below, unstructured input sentences (UBS's) are followed by their transformationally structured (SBS's) representation.

BEGIN DATA

UBS 1: Insulin increases VLDL-B in Type III patients on a high triglyceride diet.

```
SBS 1:1  (increase *(in type III patients$1 *(be on /patients$1
                                             /diet$1 *(be/diet$1
                                                          (high/TG))))
               /INSLN
               (QAMT/VLDL-B))
```

Note: For QAMT see 4b in section on Commensurability.

UBS 2: HDL are thought to accumulate cholesterol from peripheral tissues.

```
SBS 2:1  (are thought(INV-accumulate/CH$1 *(be from/CH$1/PT)
                                   /HDL))
```

UBS 3: The cholesterol ester formed in HDL may be transferred either to VLDL in exchange for triglyceride or, in part, directly to LDL.

```
SBS 3:1  (either or(may(in(be transferred to$1
                                   /CE$1 *(be formed in/CE$1/HDL)
                                   /VLDL)
                        (exchange for/CE$1
                                    /TG)))
                 (may(be transferred to$2 *in part *directly
                                   /CE$1 *(be formed in/CE$1/HDL)
                                   /LDL)))
```

UBS 4: In either case, with VLDL breakdown, the cholesterol ester would arrive at LDL.

[Note: in either case = in case X or in case Y

```
            in either case, Z = in case X, then Z or
                                         in case Y, then Z.]

SBS 4:1 (OR(in case be transferred to 2$1 SBS3 (with(breakdown/VLDL)
                                               (would(arrive at/CE/LDL))))
           (in case be transferred to 2$2 SBS3 (with(breakdown/VLDL)
                                               (would(arrive at/CE/LDL)))))

UBS 5: Alternatively, VLDL may be the exclusive source of LDL cholesterol.

SBS 5:1  (may *alternatively (INV-be source *exclusive
                                            /CH$1 *(be in/CH$1/LDL)
                                            /VLDL))

UBS 6: Cholesterol ester is removed from LDL in the splanchnic bed and LDL
       particle stability maintained by addition of triglyceride.

SBS 6:1  (and(be removed from *in the splanchnic bed
                              /CE
                              /LDL)
             (by(maintained (stability/LDL))
                (addition of/TG)))

UBS 7: The cholesterol derived from LDL is reused in the synthesis of VLDL
       and HDL.

SBS 7:1  (be reused in/CH$1 *(be derived from/CH$1/LDL)
                       (and(synthesis of/VLDL)
                           (synthesis of/HDL)))

UBS 8: It seems clear from the studies of B & G that LDL is also the
       major vehicle for transport of cholesterol to peripheral tissues.

SBS 8:1  (It seems clear from the studies of B & G that
                   (be the vehicle for transport of to *major *also
                                                      /LDL
                                                      /CH
                                                      /PT))

UBS 9: It is now apparent that normal humans possess efficient mechanisms
       for the removal of cholesterol from plasma.

      [Note 9A: We propose a reduction mechanism that will reduce the
                string "normal humans possess efficient mechanisms for
                the removal of" to "the removal of ... by efficient
                mechanisms." The process involves
                TSUB-MECHANISM:  Mechanisms for V = V by mechanisms
                TFM-POSSESS: N possess X = X P N
                     in special case where X is physiologically or
                     anatomically in N, then P=in.]

SBS 9:1  (It is now apparent that(by(removal of from
```

*(in humans *(be normal/humans))
/CH
/PL)
/mechanisms))

UBS 10: This disposal process depends on receptors located on the surface of cells in the liver and extrahepatic tissues.

SBS 10:1 (and(depends on/this disposal process
/receptors$1 *(located on *in the liver
/receptors$1
/surface of cells))
(depends on/this disposal process
/receptors$2 *(located on *extrahepatic tissues
/receptors$2
/surface of cells)))

UBS 11: The receptors bind circulating lipoproteins that transport cholesterol in the bloodstream, thereby initiating a process by which the lipoproteins are taken up and degraded by cells yielding their cholesterol for cellular use.

SBS 11:1 (thereby(bind/receptors/LP$1 *circulating
*(transport *in the bloodstream
/LP$1
/CH
/))
(by(initiate(for(ing(and(take up/cells/LP)
(degrade/cells/LP))
(yield/LP/CH))
(use/cell/CH))
/process)))

UBS 12: The lipoprotein receptors are components of an integrated transport system that shuttles cholesterol continuously among intestine, liver, and extrahepatic tissues.

SBS 12:1 (are components of/RC-LP
/integrated transport system$1
*(shuttles among *continuously
/integrated transport system$1
/CH
(and/intestine
/liver
/extrahepatic tissues)))

UBS 13: Recent experiments show that the number of lipoprotein receptors in the liver, and hence the rate of removal of cholesterol from plasma, is under regulation and that the number of receptors can be increased by certain cholesterol lowering drugs.

```
SBS 13:1  (show that/recent experiments
                     (and(and hence(be under regulation *in the liver
                                           (number/LP receptors))
                                (be under regulation
                                           (rate
                                             (removal of from/CH
                                                           /plasma))))
                          (can(increase/drugs$1 *(lower/drugs$1/CH)
                                       (number/receptors)))))
```

UBS 14: The lipoprotein transport pathway can be divided conceptually into exogenous and endogenous systems that transport lipids of dietary and hepatic origin, respectively.

```
SBS 14:1  (The LP trannsport pathway can be divided conceptually into
                   (and(TSYSTEM$1 *exogenous
                                  *(transport/TSYSTEM$1
                                             /lipid *dietary origin
                                             /     ))
                       (TSYSTEM$2 *endogenous
                                  *(transport/TSYSTEM$2
                                             /lipid *hepatic origin
                                             /     ))))
```

UBS 15: Recent studies indicate that the APO-B of chylomicrons is not identical to that of VLDL.

```
SBS 15:1  (Recent studies indicate that
                         (not(be identical to/APOB$1 *(of/APOB$1
                                                        /chylomicrons)
                                             /APOB$2 *(of/APOB$2
                                                        /VLDL))))
```

UBS 16: A typical American adult absorbs about 100 grams of triglyceride and 250 milligrams of cholesterol from the diet daily.

```
SBS 16:1  (and(absorb from *daily
                         /adult *American
                         (100g/TG)
                         /diet)
              (absorb from *daily
                         /adult *American
                         (250mg/CH)
                         /diet))
```

U BKG 2.4.7: As the triglyceride core is depleted, the chylomicron shrinks.

```
S BKG 2.4.7:1
              (as(shrink/chylomicron)
                 (be depleted/TG core))
```

U BKG 2.4.8: The excess surface material, primarily phospholipids and free cholesterol, is transferred to another plasma lipoprotein, high-density lipoprotein (HDL).

```
S BKG 2.4.8:1
               (transferred to/excess surf material$1 *(be *primarily
                                                       /excess surf material$1
                                                       (and/PH/FC))
                         /another LP$1 *(be in/LP$1 *(be/LP$1/HDL)
                                                  /plasma))
```

U BKG 2.4.9: A similar transfer reaction occurs in the endogenous lipoprotein transport system.

```
S BKG 2.4.9:1
               (occur in/similar transfer reaction
                      /lipoprotein transport system *endogenous)
```

U BKG 2.5.1: The depleted chylomicron is released from the capillary wall and reenters the circulation.

```
S BKG 2.5.1:1
               (and(be released from/CY$1 *(be depleted/CY$1)
                                    /capillary walls)
                   (reenter/CY$1
                           /circulation))
```

U BKG 2.5.2: The particle, now known as a chylomicron remnant (or simply, remnant), retains its cholesteryl ester and apoB along with another important apoprotein, apoE.

```
S BKG 2.5.2:1
               (along with(and(retain/REM
                                     /CE)
                              (retain/REM
                                     /APOB))
                          (retain/REM
                                 /APOE$1 *(be important/APOE$1)))
```

U BKG 2.5.3: The remnant (diameter, 300 to 800 A) is carried to the liver where it binds to receptors on the surface of hepatic cells.

```
S BKG 2.5.3:1
               (be carried to/REM$1 *(be/diameter of REM$1/300-800A)
                             /liver$1 *(bind to *in liver$1
                                                /REM$1
                                                /RC$1 *(be on/RC$1
                                                            /surface of hepatic cells)))
```

U BKG 2.5.4: The remnants are immediately internalized by receptor-mediated endocytosis and degraded in lysosomes.

```
S BKG 2.5.4:1
               (and(by(internalize *immediately
                                   /hepatic cells
                                   /REM)
                      (endocytosis$1 *(mediate/RC
                                              /endocytosis$1)))
                   (be degraded *in the lysosomes
```

```
                        /REM))

U BKG 2.6.1: The two-step pathway of chylomicron metabolism (triglyceride
             removal in extrahepatic tissues followed by cholesteryl ester
             uptake in the liver) is quite efficient.

S BKG 2.6.1:1
             (be efficient/2-step pathway$1 *(equals/2-step pathway$1
                                              (followed by(removal of in/TG
                                                                       /EXHT)
                                                          (uptake of in/CE
                                                                     /liver))))

U BKG 2.6.2: In man, the half-time for the clearance of chylomicrons and their
             remnants from the plasma is 4 to 5 minutes.

S BKG 2.6.2:1
             (and(half-time of 4-5 mins(clearance of from *in man
                                                           /CY
                                                           /plasma))
                 (half-time of 4-5 mins(clearance of from *in man
                                                           /REM
                                                           /plasma)))

U BKG 2.6.3: Thus, the plasma level of cholesterol rises very little, if at all,
             after a single high choelsterol meal.

S BKG 2.6.3:1
             (very little, if at all(after(rise *plasma
                                                 (level/CH))
                                           (have/( )
                                                /single high CH meal)))

U BKG 2.7.1: The liver, which rapidly takes up dietary cholesterol in the form
             of chylomicron remnants, disposes of the sterol in the bile,
             either as unesterified cholesterol or as bile acids.

S BKG 2.7.1:1
             (dispose of in/liver$1 *(take up *rapidly
                                              /liver$1
                                              /dietary CH$1 *(be in the form of
                                                             /dietary CH$1
                                                             /REM))
                         /sterol$1 *(either$1(as/sterol$1
                                                 /FC)
                                             (as/sterol$1
                                                 /bile acids))
                         /bile)

U BKG 2.7.2: No other quantitatively significant breakdown products of
             cholesterol are formed in man.

S BKG 2.7.2:1
            (INV-breakdown to *in man
                             /products$1 *(no(other than/products$1
```

 A2

(OR$1BKG2.7.1)))
/CH)

[Note: form breakdown products = break down into products]

U BKG 2.7.3: Much of the cholesterol and bile acid secreted by the liver is reabsorbed in the intestine and again delivered to the liver for excretion, thus forming an enterohepatic circulation.

S BKG 2.7.3:1

```
(thus(and$1(and(reabsorb/intestine
                         (much/CH$1 *(secrete/liver/CH$1)))
               (reabsorb/intestine
                         (much/BA$1 *(secrete/liver/BA$1))))
            (for(and(deliver to *again
                                 /
                                 /CH$1
                                 /liver)
                    (deliver to *again
                                 /
                                 /BA$1
                                 /liver))
               (and(excrete/
                            /CH$1)
                   (excrete/
                            /BA$1))))
      (form/and$1
           /enterohepatic circulation))
```

U BKG 2.7.4: During each cycle a portion of the cholesterol and bile acid escapes reabsorption and is lost in the feces.

S BKG 2.7.4:1

```
(and(and(escape *during each cycle
                (a portion of/CH$1)
                (reabsorb/intestine
                          /CH$1))
        (be lost in/CH$1
                   /feces))
    (and(escape *during each cycle
                (a portion of/BA$1)
                (reabsorb/intestine
                          /BA$1))
        (be lost in/BA$1
                   /feces)))
```

[Note: escape(reabsorption) - NEG(reabsorption)
'intestine' is supplied from 2.7.3.]

U BKG 2.7.5: With the typical American diet, which is high in cholesterol, about 1100 mg of sterol is lost from the body each day.

S BKG 2.7.5:1

```
          (lose *(with the typical American diet$1 *(be high in /diet$1
                                                                 /CH))

                *each day
                /body
                (1100mg/sterol$1))

     [Note: each day - per day]

U BKG 2.7.6: In the steady state, about 850 mg of this sterol is derived from
             endogenously synthesized cholesterol and approximately 250 mg
             from previously absorbed dietary cholesterol.

S BKG 2.7.6:1
             (and *in the steady state
                  (be derived from (850 mg/CH$1BKG2.7.5)
                                   /CH$1BKG2.7.5 *(synthesized *endogenously
                                                   /CH$1BKG2.7.5))
                  (be derived from (250 mg/CH$2BKG2.7.5)
                                   /CH$2BKG2.7.5 *(absorbed *previously
                                                 /CH$2BKG2.7.5 *dietary)))

U BKG 2.8.1: The liver converts carbohydrates and fatty acids into triglycerides
             ,which it packages into lipoproteins for transport to adipose
             tissue.

S BKG 2.8.1:1
             (and(convert into/liver
                          /carbohydrates
                          /TG$1 *(for(package into/liver
                                                  /TG$1
                                                  /LP$1)
                                     (transport to/LP$1
                                                  /TG$1
                                                  /adipose tissues)))
                 (convert into/liver
                             /FA
                             /TG$1 *(for(package into/liver
                                                     /TG$1
                                                     /LP$1)
                                        (transport to/LP$1
                                                     /TG$1
                                                     /adipose tissues))))

U BKG 2.8.2: These lipoproteins also contain cholesterol, which will be
             delivered to extrahepatic cells.

S BKG 2.8.2:1
             (contain *also
                      /LP$1BKG2.8.1
                      /CH$1 *(will(deliver to/LP$1BKG2.8.1
                                             /CH$1
                                             /extrahepatic cells)))

U BKG 2.8.3: When dietary cholesterol is available, the liver uses that source
```

of sterol, derived from the receptor-mediated uptake of chylomicron remnants, for lipoprotein synthesis.

S BKG 2.8.3:1
```
        (for(when(use/liver
                      /CH$1 *(derived from/CH$1
                                          (uptake$1 *(mediate/RC
                                                              /uptake$1)
                                                     /liver
                                                     /REM)))
                 (be available to(QAMT/CH$1 *dietary)
                                     /liver))
             (synthesis/LP))
```

[Note: "that source of sterol" → CH *dietary
This is a referential substitution where
source = dietary
sterol = classifier of CH

CH be available - CH be available to the liver
(zeroing by argument requirement and proximity of "liver".
"Proximity" to be defined later.)]

U BKG 2.8.4: When dietary cholesterol is insufficient, the liver synthesizes its own cholesterol by increasing the activity of a rate-controlling enzyme, 3-hydroxy-3-methlglutarylcoenzyme A reductase (HMG CoA reductase).

S BKG 2.8.4:1
```
(when(by(synthesize/liver
                   /its own CH)
        (increase/liver
                 (activity/rate-control enzyme$1
                          *(IDENT/rate-control enzyme$1
                                 /HMG CoA reductase))))
     (be insufficient for(QAMT/CH *dietary)
                         /   ))
```

U BKG 2.9.1: For export of triglycerides and cholesterol, the liver incorporates the lipids into VLDL (300-800A).

S BKG 2.9.1:1
```
        (for(incorporate into/liver
                             /lipids
                             /VLDL *300-800A)
            (and(export/liver
                       /TG)
                (export/liver
                       /CH)))
```

U BKG 2.9.2: The VLDL particles interact with lipoprotein lipase in capillaries, releasing most of their triglycerides.

S BKG 2.9.2:1
```
        (ing(interact with *in capillaries
                           /VLDL particles$1BKG2.9.1
```

```
                        /LPL)
        (release/VLDL$1BKG2.9.1
                 (most/TG)))
```

U BKG 2.9.3: The interaction of VLDL with lipoprotein lipase is less efficient than the interaction of chylomicrons with this enzyme.

```
S BKG 2.9.3:1
          (less than(Q(be efficient(interact with/VLDL
                                                 /LPL)))
                (Q(be efficient(interact with/CY
                                                 /LPL))))
```

U BKG 2.9.4: Thus, the half-life of the VLDL particle in the circulation of humans is 1 to 3 hours, in contrast to the previously cited half-life of 4 to 5 minutes for chylomicrons.

```
S BKG 2.9.4:1
          (in contrast to(half-life of 1-3 hours(be in *humans
                                                  /VLDL
                                                  /circulation))
                    (half-life of 4-5 min *previously cited BKG 262
                                              (be in *humans
                                                  /CY
                                                  /circulation)))
```

U BKG 2.9.5: As the size of the VLDL particle diminishes owing to its interaction with lipoprotein lipase, its density increases, and the particles are converted to intermediate density lipoproteins (IDL).

```
S BKG 2.9.5:1
          (and(as(increase(density/VLDL))
               (owing to(diminish(size/VLDL))
                     (interact with/VLDL
                                   /LPL))
               )
             (are converted to/VLDL
                              /IDL))
```

U BKG 2.9.6: The excess surface materials, mostly phospholipids and cholesterol, are transferred to HDL.

```
S BKG 2.9.6:1
          (be transferred to/excess surface material$1
                           *(be *mostly
                               /excess surface material$1
                               (and/PH
                                   /CH))
                      /HDL)
```

U BKG 2.9.7: The HDL particles interact with the plasma enzyme, lecithin-cholesterol acyltransferase (LCAT), which esterifies the excess cholesterol with fatty acids derived from the 2-position of lecithin, the major phospholipid of plasma.

S BKG 2.9.7:1

```
(interact with/HDL
              /plasma enzyme$1
                 *(be/plasma enzyme$1
                     /LCAT)
                 *(esterify$1 with/plasma enzyme$1
                                 /excess CH
                                 /FA$1 *(be derived from
                                            /FA$1
                                            /2-position of
                                             lecithin$1
                                             *(be of *major
                                                 /lecithin$1
                                                 /PH
                                                 /plasma))))
```

U BKG 2.9.8: The newly synthesized cholesteryl ester is transferred back to the IDL particles from HDL, apparently through the action of a plasma cholesteryl ester exchange protein.

S BKG 2.9.8:1

```
(through *apparently
         (be transferred to from *back
                /CE$1 [result nom esterify with$1BKG2.9.7]
                    *(synthesized *newly
                                  /CE$1)
                /IDL
                /HDL)
         (action of(CE exchange protein *in PL)))
```

[Note: "esterify X" - X ester]

U BKG 2.9.9: The net result of the coupled lipolysis and exchange reactions is the replacement of most of the triglyceride core of VLDL with cholesteryl esters.

S BKG 2.9.9:1

```
(The net result of coupled lipolysis and exchange reactions is
  (replace with(most/TG$1 core *(be of/TG$1 core/VLDL))
              /CE))
```

U BKG 2.10.1: After lipolysis, the IDL particles are released from the capillary wall into the circulation.

U BKG 2.10.2: They then undergo a further conversion in which most of the remaining triglycerides are removed and all of the apoproteins except apoB are lost.

U BKG 2.12.3: When the plasma level of LDL rises, these scavenger cells

degrade increasing amounts of LDL.
[Note that 'when' is not in Polish notation]

S BKG 2.12.3:1
```
      (when(rise(level *plasma
                        /LDL))
            (degrade/scavenger cells
                       (amount$1 *(increase/amount$1)
                                  /LDL)))
```

U BKG 2.14.2: At least six genetically determined disorders disrupt the endogenous or exogenous transport system in ways that produce an increase in the plasma level of one or more lipoproteins.

S BKG 2.14.2:1
```
  (in ways that produce(disrupt(six/disorders$1 *(determine/genetics
                                                          /disorders$1))
                                  (or/endogenous transport system
                                    /exogenous transport system))
                         (increase(level *in PL
                                          /LPs)))
```

U BKG 3.1.7: The number of receptors in fibroblasts can be increased by certain hormones that stimulate cell growth, including insulin, thyroxine, and platelet-derived growth factor.

S BKG 3.1.7:1
```
  (by(can(increase(number *in fibroblasts
                          /receptors)))
     /hormones$1 *(stimulate/hormones$1 *(ing(and(include/hormone$1/INSLN)
                                                 (include/hormone$1/thyroxine)
                                                 (include/hormone$1/platelet-
                                                    derived growth factor)))
                               (grow/cells)))
```

U BKG 3.3.5: The mutant FH cells grow in tissue culture because they adapt to the unavailability of LDL-cholesterol by increasing HMG CoA reductase activity and cholesterol synthesis.

S BKG 3.3.5:1
```
  (because(grow *in tissue culture
                /mutant FH cells)
        (by(adapt to/mutant FH cells
                   (unavailability of /CH
                                      /( )))
           (and(increase/mutant FH cells
                         (Q(activity/HMG CoA)))
               (increase/mutant FH cells
                         (Q(synthesis/CH))))))
```

U BKG 4.3.3: The power of correlations has been increased by treatment of the animals with drugs that alter the number of receptors.

S BKG 4.3.3:1
```
  (by(increase(power/correlations))
```

```
    (treatment of with/animals
                      /drugs$1 *(alter/drug$1
                                    (number/RC))))
```

U BKG 4.4.3: The membrane binding assay showed that the livers of estradiol-treated rats have a tenfold increase in the number of lipoprotein receptors.

```
S BKG 4.4.3:1
  (show that/the membrane binding assay
           (increase[tenfold]/liver$1 *(be of/liver$1/rats$1
                                         *(be treated with /rats$1
                                                      /estradiol))
                      (number/lipoprotein receptors)))
```

[Note: X has Q increase = there exists Q increase in X]

U BKG 4.6.1: In young beagle dogs an increase in hepatic LDL receptors, as measured in vitro, has been shown to be correlated with an increase in removal of 125-I-labeled LDL from the circulation as measured in vivo.

```
S BKG 4.6.1:1
 (It has been shown that(correlate with *young beagle dogs
                                    (increase *in vitro
                                            (QNUM/RC-LDL))
                                    (increase *in vivo
                                            (QAMT(removal of from
                                                    /125-I-labeled LDL
                                                    /circulation)))))
```

U BKG 4.6.3: By trapping bile acids in the intestine, colestipol causes the liver to convert more cholesterol to bile acids and hence increases the heptic demand for cholesterol.

```
S BKG 4.6.3:1
 (and hence(by(cause/colestipol
                    (convert to/liver
                              (more/CH)
                              /bile acids))
               (trap in/colestipol
                      /bile acids
                      /intestine))
            (increase(QAMT(demand of for/liver
                                        /CH))))
```

U BKG 4.6.4: Mevinolin blocks the compensatory increase in cholesterol synthesis that would ordinarily occur under these conditions and thus forces the liver to rely on lipoprotein cholesterol for the synthesis of bile acids.

```
S BKG 4.6.4:1
 (and thus(blocks/mevinolin
                 (increase$1 *(compensate for/increase$1
                                           /          )
```

```
                    (synthesis/CH)))
        (force/mevinolin
            (for(rely on/liver
                    /CH)
                (synthesis/BA))))
```

U BKG 4.6.6: This change was associated with a twofold increase in the fractional clearance rate of intravenously administered 125-I-labeled LDL from the plasma.

```
S BKG 4.6.6:1
 (associated with
  /this change
  (increase[twofold](rate(FC clearance of from/125-I-labeled LDL$1
                                          *(administer *intravenously
                                                                  /
                                                  /125-I-LDL$1)
                                          /plasma))))
```

U BKG 4.6.7: This increased efficiency of LDL clearance, plus a 50 percent reduction in the synthetic rate of LDL produced by mevinolin, contributed to a remarkable 75 percent drop in the plasma level of LDL-cholesterol in the treated dogs.

```
S BKG 4.6.7:1
 (contribute to(plus(increase(efficiency(clearance of from/LDL
                                                          /   )))
                (reduction$1[50%] *(produce/mevinolin
                                          /reduction$1)
                                  (rate(synthesis/LDL))))
            (drop[remarkable 75%](level *LDL *LOC:plasma *POP:dogs
                                    /CH)))
```

U BKG 4.7.1: Of all animal species studied, only the rabbit has the propensity to develop a massive increase in plasma cholesterol within days after being placed on a high cholesterol diet.

```
S BKG 4.7.1:1
 (propensity(after(develop(increase(QLEVEL *in rabbit *in plasma
                                    *within days
                                    /CH)))

                (placed on/rabbit
                        /high CH diet)))
```

[Note: X has the propensity to = there exists a propensity to...in X]

U BKG 4.7.3: These particles accumulate in plasma because their removal rate fails to increase in proportion to the increased production of chylomicrons and VLDL that occurs during cholesterol feeding.

```
S BKG 4.7.3:1
 (because(accumulate in/these particles
                        /PL)
        (fail to(in proportion to(increase(rate(removal of from/
```

```
                                   /   )))
                     (increase(QAMT(production(and/CY
                                                /VLDL)))))))
```

U BKG 7.2.3: This increase in LDL receptor activity explains the long-known ability of these two agents to lower plasma cholesterol levels.

```
S BKG 7.2.3:1
  (explains(this increase(activitity/RC LDL))
           (ability$1 to *(be long known/ability$1)
                         /two agents
                         (lower(level *plasma
                                     /CH))))
```

U BKG 7.2.4: These results in humans have been supported by animal and cell culture experiments showing that bile acid-binding resins increase hepatic LDL receptors in rabbits and dogs and that L-thyroxine increases LDL receptors in fibroblasts and probably in the liver.

```
S BKG 7.2.4:1
  (These results in humans have been supported by animal and cell culture
   experiments showing that
       (and(increase/BA-binding resins
                    (QNUM *rabbits *dogs *liver
                          /RC-LDL))
           (increase/L-thyroxine
                    (QNUM *fibroblasts *liver probably
                          /RC-LDL))))
```

U BKG 8.4.3: Epidemiologic evidence suggests that excessive dietary intake of fat, cholesterol, and calories is responsible for the increased blood cholesterol.

```
S BKG 8.4.3:1
  (suggests that/epidemiologic evidence
    (be responsible for (QAMT$1  *(exceed/QAMT$1
                                         /QAMT$2 *normal)
                           (intake of from(and/fat
                                                /CH
                                                /calories)
                                          /diet))
                        (increase(QLEVEL *blood
                                         /CH))))
```

U BKG FIG.16.2: This pharmacologic manipulation takes advantage of the knowledge that the receptors are normally under feedback regulation and that the number of receptors in the liver can be increased when the liver's demand for cholesterol is increased.

```
S BKG FIG.16.2:1
 (takes advantage of the knowledge that/this pharmacologic manipulation
    (and(be under feedback regulation *normally
```

```
                                        /RC)
  (when(can(increase(number *liver
                                   /RC)))
       (increase(QAMT(demand of for/liver
                                            /CH))))))
```