Chapter VI

Informational Structures of Language in a Subfield of a Science

In the last chapter it was suggested how meaning, reconstructed as predication-created information, can serve as a criterion of the adequacy of grammars formulated in accordance with the explicit constraints of a general theory of language structure which consist of three fundamental relations: of a partially ordered word dependence requirement, of gross differences in likelihood (or 'expectability') [1] of occurrence of operator words upon different words of their argument class, and of reduction in phonemic shape consequent upon the entry of an operator word into a sentence which contains words (arguments) for which the entering operator has a high likelihood of co-occurrence. There we also attempted to substantiate a conceptual linkage between redundancy, i.e., restrictions upon combinations of elements (e.g., the operator/argument dependence requirement among words), and a notion of information. In the present chapter we survey the results of a detailed discourse analysis, framed within these constraints, of a corpus of texts in a subfield of a science.[2] The express purpose of this work was justificatory: to show that the methods employed were adequate to provide a canonical form for the information of a particular research problem within a science; that they reproduced, in an empirically controlled manner, the antecedently known results and developments within this scientific subfield.

[1] On some problems raised by the notion of 'likelihood', see fn 3 p. 305.

[2] The term 'subfield of a science' is used in accordance with the sense provided in Shapere (1976:281):"The real give-and-take of science, the real wrestling with concrete problems, takes place at levels of <u>subfields</u> of (scientific fields -- physics, astronomy, chemistry, biology, etc)."

Several times in previous chapters we have alluded to the fact that the analysis of sentences in connected discourse reveals many restrictions on word combinations whose domain extends beyond sentence boundaries, and which therefore do not 'surface' in a sentence grammar of the language as a whole. Due to the additional constraints of word recurrence and regularities of word combination in connected discourse, the elements comprising the sentences of discourse may be redefined. The sentence-forming elements of discourse can be identified as members of specific word classes (which may include word sequence as well as single word members); the word classes are purely positionally defined by the occurrence of the sentence-forming elements with respect to each other. Sentences may then be recognized as types comprised of strings of word classes (taken in an ordered 'normal form' sequence); these are called "elementary information units" or, since each word class may be represented by a designated symbol, as "formulas of information". Beyond this, sequences of sentence types may be established ("macrosentences") representing sentence types conjoined under particular operators or in terms of an ordered sequence of such operators. Again, the criterion of the identification of the larger and larger discourse and sublanguage (see below) elements remains the relative positioning of these elements with respect to one another.

Major portions of texts pertaining to this scientific subfield can be seen to be paraphrastically characterized and hence described by a 'grammar' of these subject=matter specific constraints on word combination. In turn, representing the texts in terms of these constraints serves to exhibit the objects, ~~and~~ relations and discussions with which this subfield of a science is concerned. This is to accord the scientific subfield an explicit, and, in many ways, quite precise informational representation.

As argued above in Chapter 5, an informational representation of this kind is not a rational reconstruction of the science in a language to which an antecedent interpretation is assigned. Starting only from an informal conceptual connection between redundancy and information, the methods of discourse analysis provide a means of eliminating from linguistic description variant forms which 'say the same', as empirically attested by informants who have a specialized knowledge of the subject matter and who are, in addition, native speakers of English. The resulting description ('grammar') is then a structure of ineliminable restrictions on (redundancies of) word combinations. This is the sense in which it has been said, several times above, that meaning (in particular, the recognition that two <u>apparently different forms</u> 'say the same') is required to determine repetition, or, alternately, that linguistic structure cannot be adequately specified in an external metalanguage (Chapter 5 §1).

An exhaustive presentation of this analysis will soon appear in print [1] and the treatment of many details and fine points must be

[1] Harris <u>et al</u> (in press); hereafter designated 'FIS'.

Are you asking informants for paraphrase as well as for repetition? Or are you just asking them to confirm your prediction that 2 things "say the same"?

Some will read this as a claim that informants can judge paraphrase relations. Perhaps an explicit recalling of the two versions of economics would help — but they were not repetitions, and their relation emerges from distributional analysis, not from informant judgements. That they "say the same": Just so, 2 apparently different forms turn out here to "say the same" under discourse analysis, and informants confirm that prediction. right?

deferred to this larger work. Our interest here is to provide an indication of how the output of the analysis of the text of research reports concerning a particular problem in the formative period of cellular immunology corroborates the general theory of language structure and grammar of English on which the analysis is based.[1]

After some initial orientation (§1) concerning the specific research problem treated in the corpus of reports analyzed in FIS and a clarification concerning the notion of a 'scientific sublanguage', we proceed in §2 to a summary presentation of the methods of analysis and the informational considerations guiding them. A partial survey of the results -- the output of the analysis -- is given in §3. And in §4, it is shown how the sublanguage formulas are adequate as an informational representation of this subfield of cellular immunology: as compact summaries of the articles, whose changes accord ~~changing in accordance~~ with known changes in methods and results in the subfield; and even as providing a basis from which to critique the course of research presented in the actual texts themselves. In doing so, the output of the analysis may be considered as taking a major step towards providing an instance of a 'grammar of a science', i.e., a structural characterization of the objects and relations of a science and an exact specification which situates a scientific domain in relation to neighboring domains.[2]

[1] Harris (1982) and Harris (forthcoming).

[2] By 'domain' of a science, we refer, loosely, to the set of things studied in that inquiry, a formulation given by Shapere (1976) who makes the congenial point that that the characterization of the objects of a domain strongly reflects the influence of (changing) knowledge-claims about them.

6.1 The Research Problem and Corpus of Research Reports. Following the first experimental demonstration of antibody formation in lymphatic tissue by McMaster and Hudack (1935),[1] a primary question to be answered in the field of research which became known as cellular immunology was: Which cell type among those generally found within the lymph node is involved in the actual synthesis of antibody. Beginning in the 1940's, a controversy ensued, largely waged between European and American scientists; the former generally holding that cells of the lymph system termed "plasma cells" (for their abundant content of cytoplasm) are responsible, the latter in the main favoring the view that other, generally smaller and presumably distinct, cells of the lymph system termed "lymphocytes" (Dougherty, Chase and White (1944))[1] are the cellular site of antibody formation. In the course of nearly thirty years of experiment, it eventually emerged (by around 1965) that both were antibody producers and that these names were being used for what were, in fact, different stages of the same cell line. It should be noted that beginning about 1961, the first indications appeared of a distinction within the cell population termed "lymphocyte", i.e., between thymus-dependent (T) lymphocytes and thymus-independent, bone marrow derived (B) lymphocytes. Together with macrophages, another kind of white blood cell, these different cell types work together to elaborate antibodies directed towards specific antigens in ways that are yet to be fully understood.[2] It is the study of this cooperative interaction

[1] See the list of research articles below.

[2] Sato and Gefter (1981), Chapter 1; on the distinction between T and B lymphocytes (the latter are the precursors of plasma cells) see Benacerraf and Unanue (1981). As has been subsequently recognized, there are several different subtypes of T lymphocytes.

between different cell types to which the disciplinary appellation "cellular immunology" is now applied. In addition, it may be mentioned that the latter stages of the "which cell?" controversy transpired against the backdrop of another, largely theoretical, controversy (termed "selectionist" versus "instructionist") regarding the actual mechanism of antibody-antigen specificity: how it is that a body can 'recognize' the enormous number of different antigen molecules by forming antibodies specific to them.[1] However, the resolution of this question (in the work of the Nobel prize winners Macfarlane Burnet, Medawar and Jerne, among others) had little or negligible impact upon the "lymphocyte"/"plasma cell" dispute which was resolved by electron microscopic studies of specially prepared single cells, in which the effects of a continuing synthesis of antibody could be observed.[2]

The main reason the "which cell" problem was chosen as a test case of the linguistic methods was that it had a clearly identifiable beginning and resolution, and because the two sides of the controversy could be clearly delimited in a reasonably sized selection of papers, which was done by workers who actually participated in the research.

The standard format for research reports in the bio-medical sciences divides the article under the following headings: "Introduction", "Methods and Materials", "Results", "Discussion" and "Conclusion". With the sole exception of the "Methods and Materials" sections (see further below), the entire text of ca. 20 articles were analyzed. Due to limitations of size, however, only portions of the analysis of 14 articles are published

[1] For a philosophically insightful discussion, see Edelman (1974); more generally, Jerne (1967).

[2] See T.N. Harris and S. Harris, "The Cellular Source of Antibody: A Review", Chapter 8 of FIS.

in Appendix 1 of FIS. In addition to the 20 articles written in English, 3 articles in French were also analyzed, by a native speaker of French. The output of this analysis showed that the structures obtained for the English corpus were also sufficient for the French. Portions of this analysis are presented in Appendix 2 of FIS. For ease of reference below, we list here the 14 English articles:

(abbreviations: JEM = Journal of Experimental Medicine,
JI = Journal of Immunology)

1. McMaster, P. and Hudack, S. (1935): "The Formation of Agglutinins within Lymph Nodes," JEM, 61, 783-805.
2. Dougherty, T., Chase, J. and White, A. (1944): "The Demonstration of Antibodies in Lymphocytes," Proc. Soc. Exp. Bio. Med., 57, 295-298.
3. Björneboe, B., Gormsen, H. and Lundquist, F. (1947): "Further Experimental Studies on the Role of Plasma Cells as Antibody Producers," JI, 55, 121-129.
4. Fagraeus, A. (1948): "The Plasma Cellular Reaction and its Relation to the Formation of Antibodies in Vitro," JI, 58, 1-13.
5. Harris, S. and T.N. Harris (1949): "Influenzal Antibodies in Lymphocytes of Rabbits Following the Local Injection of Rabbits," JI, 61, 193-207.
6. Ehrich, W., Drabkin, D. and Forman, C. (1949): "Nucleic Acids and the Production of Antibody by Plasma Cells," JEM, 40, 157-167.
7. Keuining, F. and van der Slikke, L. (1950): "The Role of Immature Plasma Cells in the Formation of Antibodies, as Established in Tissue Culture Experiments," Journal of Laboratory and Clinical Medicine, 36, 167-182.
8. Coons, A., Leduc, E. and Connolly, J. (1955): "Studies on Antibody Production I. A Method for the Histochemical Demonstration of Specific Antibody and its Application to a Study of the Hyperimmune Rabbit," JEM, 102, 49-60.
9. Leduc, E., Coons, A. and Connolly, J. (1955): "Studies on Antibody Production II. The Primary and Secondary Responses in the Popliteal Lymph Node of the Rabbit," JEM, 102, 61-72.

10. McGregor, D. and Gowans, J. (1963): "The Antibody Response of Rats Depleted of Lymphocytes by Chronic Drainage from the Thoracic Duct," JEM, 117, 303-320.

11. Harris, T.N., Hummeler, K., and Harris, S. (1966): "Electron Microscopic Observations on Antibody-Producing Lymph Node Cells," JEM, 123, 161-172.

12. Leduc, E., Avrameas, S., and Bouteille, M. (1968): "Ultrastructural Localization of Antibody in Differentiating Plasma Cells," JEM, 127, 109-118.

13. Gudat, F., Harris, T.N., Harris, S., and Hummeler, K. (1970): "Studies on Antibody-Producing Cells, I. Ultrastructure of 19S and 7S Antibody-Producing Cells," JEM, 132, 448-474.

14. Yoffey, J. and Courtice, F. (1970): Lymphatics, Lymph and the Lymphomyeloid Complex. 3rd edition. NY, Academic Press, 573-88.

subhead: Sublanguage

A sublanguage is defined in Harris (1968) as a proper subset of the sentences of a language that may be closed under some or all of the operations defined in the language.[1] Hence we may think of a sublanguage as ~~comprised of~~ comprising all sentences which can be described by a 'grammar' of specific word classes and sentence types which may be constructed (positionally defined) by the application of methods of discourse and transformational analysis to a corpus.

It has been objected that the definition of a sublanguage given here is insufficient or incomplete, in that no role is assigned to belief in the characterization of a sublanguage. Of course, such an objection has no place for certain sublanguages, e.g., the sublanguage of a language that is its grammar. ~~But,~~ On this objection,

[1] Harris (1968:152):"Certain proper subsets of the sentences of a language may be closed under some or all of the operations defined in the language, and thus constitute a sublanguage of it."

there is no clear separation to be drawn between meaning and belief in the sense that an informant's discernable judgements about the meaning of sublanguage expressions are not to be gauged as independent of either his beliefs about specific subject matters or of what is often referred to as his 'background beliefs'. Further, the objection may run, since beliefs are likely to differ among different informants or users of the sublanguage, the way seems clear to encounter difficult problems of incommensurability and translation. Moreover, the sentences of a discourse reporting scientific research are typically asserted as true by the authors of the report, although, given conflicting opinions, interpretations of data, standards of evidence, and the like, a collection of such discourses within what may be considered to be the same sublanguage is liable to contain truth-functional inconsistencies. Even so, it seems not improbable that a rather sizeable intersection can be located in the aggregate of sentences of all such discourses, the truth or warrant for which all competent researchers would assent to. Does the definition of a sublanguage require reference to this communality of belief? There appears to be no compelling reason for such a requirement. A sublanguage of a science can initially be approached by choice of a particular research problem, perhaps as selected by a knowledgeable researcher or by closure under citation of texts. But this is only an interim demarcation. A sublanguage is specified by its grammar;

as will be suggested below, a corpus of texts may be found to ~~be comprised of~~ several distinct and partially intersecting sublanguages so characterized. There is no necessity that the word classes be closed, or the vocabulary frozen. Any sentence that can be 'housed' within the defined structures of a sublanguage grammar belongs to that sublanguage. A case could be made that the appeal to belief appears most unavoidable in the case of sentences which have not occurred in any particular sublanguage discourse but which, perhaps because of shared vocabulary, may be entertained to be sentences of the sublanguage. Thus it might be maintained that e.g., _the antibody developed inflamation_, _plasmacytogenesis occurs in liver tissue_, and _lymphocytes produce renal fat_ could only be ruled out of the sublanguage of cellular immunology on grounds that any researcher in the field would assert their falsity or even incoherence. But independently of the appeal to belief, it can be shown that no existing sentence type of the grammar of this sublanguage characterizes any of these sentences, each of which involves displacements of established patterns of word occurrence which are the basis of ~~existing~~ word classes and subclasses. In general, unless such violations of patterns of occurrence recur or can be systematically linked to structural patterns which do recur, no convincing reason ~~has been given~~ for including such sentences within the sublanguage. Under these conditions, even should such sentences occur, they could be viewed as anomalies, perhaps misprints, and a principled case made for their exclusion from the sublanguage, all

This is the step of ex post facto verification noted above, on p. 375 (per my suggested changes there).

without appeal to the beliefs of sublanguage speakers and users. This is not to say that sublanguage informants should not be consulted as to the legitimacy of this, purely grammatical, exclusion. The significant issue is really whether the notion of a sublanguage, so defined, is a workable one, whether any significant consequences follow from its adoption.[1] If this is so, then also no special problems of incommensurability or translation arise in considering a sublanguage under this definition. And, in any case, each user of a sublanguage has available the resources of the containing language (e.g., English) with which to ensure communicability and mutual understanding, even where beliefs and conclusions diverge.

As noted above in Chapter 3, to formulate sentence types is already to project from the described corpus to many other sentences not in the corpus which may never be observed. We may, for example, wish to say that certain extensions of a sublanguage grammar (allowing for the introduction of new word classes, and perhaps new sentence types, and for new members of established word classes and sentence types) also characterize sentences of the specified sublanguage though it is at best unclear, at present, how far we might push in this direction. How

[1] Accordingly, the adequacy of the proposed closure of a sublanguage under certain elements and operations definable in the language as a whole is above all a pragmatic matter. In particular, the methods of sublanguage analysis outlined here, based upon this definition, are not wedded to a notion of analyticity. It may be that other, more philosophical inquires into sublanguages or specialized uses of language will perhaps require reference to a theory of belief in order to explain certain aspects of behavior.

much extension such a grammar can be allowed and yet still be said to characterize the same sublanguage is a matter which requires further study (and to some extent, at least, further definition).

For example, FIS distinguishes a residue class M of metascientific material which operate grammatically on what are termed "science sentences" (those sentences which are instances of specific sentence types); semantically, the M segments state the scientist's views of action in respect to matters described in the science sentences: McMaster and Hudack demonstrated the production of antibody in lymph nodes has the M sentence segment McMaster and and Hudack demonstrated with the remainder a science sentence of AV_pT type. And M also includes certain sentence 'introducers' (however, etc) and connectives (...serves to indicate that..., etc.) and other material which may more properly be seen, not only as metascientific but ~~as well~~ also part of a sublanguage of laboratory procedures and techniques: In single cell droplet studies which involved a more sensitive antibody assay than that of Nossal, Attardi, Cohen, Horibata and Lennox found (11, 161.3.2)[1] or since the volumes on which calculations of volume were based were derived from a graph which agreed closely with one based on hematocrit-readings of packed cells (5,205.2.2). Since certain members of M are common to much scientific writing, not merely to the subfield of immunology dealt with in FIS, there may be reason to distinguish an immunology sublanguage of the so-called "science sentences" and another, metascientific sublanguage, of operators on these. Still another sublanguage of laboratory procedures and techniques may perhaps be defined by word classes and sentence types

[1] References of this form are to the listing of articles above, as follows: paper number (in this listing), page, paragraph number (counting from the initial text on a page) and sentence number within the paragraph. Thus this reference is to paper 11, page 161, paragraph 3, sentence number 2.

which may be set up for the text of the "Methods and Materials" sections, an analysis omitted from FIS since the regularities of word combination existing in the other sections of the articles could not, in the main, be shown to occur here.

Thus the complete text of each article is demonstrably not included within the immunology sublanguage closed under the objects and relations of a 'grammar' of only the science sentences of these articles and their undifferentiated M operators. On the other hand, given the coincidence of structures obtained for the "science sentences" in both English and French research articles, we cannot say that the sublanguage of these sentences belongs to one or the other language. As Harris has pointed out, this sublanguage is more appropriately looked upon as "an independent linguistic system sufficient for articles in a particular research area".[1] It may be fitting to speak here of a sublanguage of (a subfield of) a science, where the language of a science could be considered to be a collection or 'envelope'[2] of partially intersecting, partially independent sublanguages.[3]

[1] Harris (to appear): "(I)t was found that when French articles in in field were analyzed, the same word classes and sentence types appeared as in the English articles. The language of each set of articles can be considered a sublanguage of its particular natural language. But the language common to them all, consisting of the word subclass symbols (which suffice for a vocabulary) and their sentence types, is not a sublanguage of either English or French. Instead, it can be looked upon as an independent linguistic system sufficient for articles in a particular research area."

[2] Harris has used the term with this sense, see e.g., his (1981). — but in his lectures, he said the whole lg is not an envelope of its sublgs.

[3] On the interrelations between sublanguages, see Hiż (1975).

6.2 Procedures of Analysis: (a) Gross Grammatical Analysis: Identification of Word Dependence (Operator/Argument) Relations. The analysis of (most) written discourse is facilitated in that segmentation of the discourse into sentences is given. Given sentence boundaries, the initial step of the analysis is to identify the operator/argument relations of the words of the individual sentences. However, since many occurrences of 'the same' word' are often in what appear to be different grammatical environments, the classification of word dependence relations can most reasonably get underway by considering at first only those words which, in almost all occurrences, have the same grammatical position (e.g., the 'noun subject' of a sentence). Most notable among these are words with null argument requirement,[1] i.e., many nouns such as antigen, antibody, cell, tissue, as well as names for various cellular ultrastructure -- endoplasmic reticulum, nucleus, etc. -- and names for various tissues -- serum, renal fat, liver etc. In addition to these words of null requirement, there are also words with non-null requirement which appear to occur with nearly always the same relation of word dependence. Included here are operators on sentences (operating on the 'highest' (least upper bound in the directed semi-lattice in which the partial order of word dependences may be represented) operator of that sentence) which have

[1] I.e., words which require no other words for their 'entry' into a sentence and which, therefore, are not operators. It should be noted, however, that nearly all null requirement words can also be treated as predicates of a noun (null requirement word). For example. 'classifier' occurrences of tissue, and cell (The serum is a tissue, A macrophage is a cell) are rare, if not non-existent in research reports, but they may be common in elementary texts. Almost all proper and common nouns can in turn be considered as 'classifiers' (O_n operators) of the indefinites, someone, something, e.g., Something is a whale, etc.

the word dependence classification 0_o, such as continue or begin, as in Mature plasma cells began to appear in large numbers only on the 4th day...(from 6,164.3.2). And there is also a good deal of stability among certain binary operators, e.g., certain words of the classification 0_{nn} which operate on two null requirement words, as contain in The rare ergastoplasmic cisternae in this cell sometimes also contain antibody... (from 12,112.3.3) operating on (cisternae, antibody), produce in antibody is not produced by small lymphocytes...(from 10,317.1.1) operating on (lymphocytes, antibody), or synthesizes, stores in the cell synthesizes antibody at a higher rate but stores some of it... (from 11,167.1.4) each operating on the argument pair (cell, antibody).

Another example of 0_{nn} binary operators are operators of a certain metalinguistic cast, e.g., were called, were identified as, as in large reticulum cells were called transitional cells...(from 4,9.3.4). Taking these operators as binaries rather than ternaries, as might be the normal case in the language as a whole (e.g., We called large reticulum cells 'transitional cells', etc) illustrates how the argument requirement classification of 'the same word' can diverge between its sublanguage occurrences and its wider occurrences. Since the style of much scientific writing of experimental findings favors the passive voice over the active, the great bulk of occurrences of certain 'meta' words, which in the active voice require specification or mention of the authors or investigators (find, identify, call, detect, etc) of the paper, are in passive form (was found, were identified as, were called, were detected, etc.) which permits elimination of mention of the active subject (We, investigators, Mitchel and Gowans, etc.). Taking the sublanguage occurrences of these words as binary operators, then, is already a departure form the method of assignment of operator/argument relations to the words of a sentence by an axiomatic grammar of the language as a whole (see Chapter 5 §3) whose concern is to show, by providing explicit derivations, that each word of the language (counting homonyms as different words) has a single 'regularized' word dependence classification in all of its occurrences. Thus, from the point of view of a grammar of the language as a whole, a sentence like these cells were identified as lymphocytes would be derived from a source with a reconstruction of the zeroed subject, We/investigators identified these cells as lymphocytes, since identify may be taken as having a 'regularized' base classification, 0_{nnn} (identify operating on the ordered triplet(We, cells, lymphocytes)). The binary classification in the sublanguage is justified, however, since there is a standard transformation which takes a sentence of the active voice and transforms it into an informationally equivalent sentence of the passive (These cells were identified as lymphocytes by us) and where the passivized subject may be zeroed on grounds of contributing no additional information.[1] This elementary example shows how, using transformations available in the language as a whole, a canonical or 'regularized' classification of operator/argument relations over the domain of a sublanguage may be obtained which is 'different' from that desired for the language as a whole. In particular, there is no compelling reason to regularize sublanguage word occurrences by going 'all the way down' to their base classification given in the axiomatic grammar of the sentences of a language. While this may be done on demand in each instance (with the possible exception of technical vocabulary which cannot be said to occur in the language as a whole outside of a particular sublanguge), thereby showing that the sublanguage is indeed a <u>sublanguage of English</u> (in this case), to do so often works against the sublanguage-grammatical objective of exhibiting the maximum degree of regularity of word recurrences that the constraints of the methods and empirical control permit. And this favors constructing sentence types relating as many word classes as possible, and sentence sequences relating as many sentence types as possible. For the objective of this analysis is to eliminate from grammatical description as many as possible of the variant forms which can be determined to 'say the same', thereby ensuring that the remaining redundancies of word combination are actually information-creating.[2]

[1] As is certainly the case with a written text.

[2] To this theoretical objection to full decomposition must be added another, practical and aesthetic, one: since each reconstruction is noted and justified in FIS, the comprehensibility and readability of the final output would suffer from unnecessary details of reconstruction.

Another, perhaps more important, instance of how transformations valid for the language as a whole are already employed in obtaining an initial segmentation of sublanguage sentences into operator/argument relations is the case of the word *inject*. In all but the "Methods and Materials" sections of the articles, the bulk of occurrences of *inject* are as nominalizations, e.g., *following the injection of inactivated influenzal virus into the foot-pad of rabbits...* (from 5,204.2.1) where it may be seen to have two arguments, a word of the G ('antigen') class -- here, *influenzal virus* -- and a word of the B ('body' or 'body part') class -- *the foot-pad of rabbits*. So characterized, *inject* is a word of the J word class (which includes -- in some of their occurrences -- other words such as *immunize* as in *rabbits were immunized strongly for several weeks with a formolized mixture of 8 pneumococcus types...*(from 3,121.1.2) and *sensitize* as in *the animals were sensitized by means of subcutaneous injections of horse serum* (4,1.3.1)); words of the J class are binary O_{nn} operators having a first argument from the G class and and second argument from the B class. On the other hand, in the "Methods and Materials" sections which describe the laboratory methods and experimental procedures and measurement techniques, there are perhaps sufficiently many occurrences of *inject* as O_{nnn} (of the form *we injected the rabbits with influenzal virus*) that, as regards these sections, the most appropriate characterization of these occurrences is as O_{nnn} with an explicit argument place for a class of words like *we*, *investigators*, *Mowat*, etc. Although, as noted above, the "Methods and Materials" sections were not analyzed in establishing the initial result in FIS on the grounds that many of the "Methods" sentences contained no or only one member of the otherwise established word classes, preliminary work indicates that the "Methods" section may comprise a distinct sublanguage or confluence of sublanguages (including, perhaps, a sublanguage of measurement) dealing with procedures, instruments, techniques of measurement and the like, whose applicability is wider than that of the part of cellular immunology reported in the remainder of the article. One, relatively minor, indication of how these distinct sublanguages may be related is therefore the transformation, similar to that noted above, which establishes a connection between 0_{nn} occurrences of inject and its occurrences as 0_{nnn}.

It should be noted that there are also occurrences of inject which are not members of J; these are found exclusively in this corpus, in paper 10 (which also contains J occurrences of inject) where an unusual experimental technique of transferring lymphocytes obtained in one animal into another, lymphocyte depleted animal, is referred to: (An attempt was therefore made to reverse the unresponsive state by injecting lymphocyte-depleted rats with thoracic duct cells from normal non-immunized rats...(10,310.1.2)). Here inject occurs as 0_{nnn}, having a first argument of the C ('cell') word class, and a second and third argument of the B class. These occurrences of inject are taken as members of a new class I, defined for the context C_BB. This is but one of many cases of 'sublanguage homonymity' which are, however, readily resolved by noting the word class operator/argument relations of each occurrence of a word.

(b) Sentence types. The initial segmentation of the sentences of the corpus into their sublanguage operator/argument relations yield a first approximation to the word classes of the sublanguage. Here are found the noun (null requirement) classes: G ('antigen'), A ('antibody'), C ('cell'), T ('tissue'), S ('untrastructure of the cell'), B ('body') and a few others. Having these, we may then construct operator word classes whose members require words of one or more of the noun classes; for example, a class V of words (including produce, contain, is found in, etc.) whose members require words of the A and C (or S of C) or T classes, a class W of operators whose members (enlarge, development, change, contain, etc.) require one or more words of the C, S or T classes, a class J (as seen above) whose members require words of the G and B classes, and so on. The next step is to represent the recurrence of these word classes with one another in a canonical form, as sentence types. Thus a sentence like plasma cells produce antibodies can be represented as a sequence of the chosen symbols for the word classes to which the different words, as determined by the initial segmentation, belong, e.g., mirroring the left to right order of their occurrence as CVA, or for the production of antibody by plasma cells as VAC, or the production by plasma cells of antibody, VCA, or antibody is produced by plasma cells, AVC, etc. Since all of these sentences or sentence fragments are transformationally related, one word class symbol sequence may be selected as a "normal linear form" to designate this type of sentence: here the choice is AVC. In the last pages of Chapter 5 §3 above, we briefly indicated how, with the use of a re-linearizing 'transformation', which can be viewed as particular to the sublanguage, and a leftward pointing reading instruction, which instructs that the reading of the sentence should begin with the rightmost segment and proceed leftward, word sequences whose word-class linear order varies may be represented as instances of a common sentence type.

Here we need only point out how transformations again are employed in the analysis in obtaining normal form representation of sentence types. However, in distinction to the previous use of transformations in obtaining a segmentation into operator/argument word classes, the re-linearizing transformation and the reading instruction are not really operations which have validity for the language as a whole.[1] Yet this is not a burdensome cost to the analysis since each of these supplementary operations could be eliminated, the same result being attainable by transformations available for the language as a whole, e.g., by nominalization or denominalization. So a sentence like plasma cells produce antibody which can be segmented, in AVC normal form, using a left-pointing arrow (¦ antibody ¦ produce ¦ plasma cells ← ¦), may also be nominalized into AVC normal form (antibody's ¦ production by ¦ plasma cells ¦). However, since nominalization or other transformations involve change of phonemic shape of the sentence transformed,[2] these special sublanguage devices can be considered to be 'shortcuts' which, by

[1] For example, the re-linearization which produces of antibodies the production by plasma cells, is only of dubious standing in the language as a whole.

[2] And all such alterations of phonemic shape must be justified.

The arrow suffices for only one of the several orderings (the VCA order). What of the CVA, VAC, etc orderings? What indicates where to start & what comes 2nd?

easing the burden of reconstruction, aid the alignment of transformationally relatable sentences into a canonical structural form. As such they are distinct from the special sublanguage reconstructions of zeroed material (considered below) which have no standing in the grammar of the language as a whole.

(c) WH-. A further use of transformational methods in obtaining repeating sentence types (sublanguage formulas), which are available for the sentences of the language as a whole, is in the decomposition of many text sentences into a 'primary' sentence S_1 and one or more 'secondary' sentences S_2, S_3,... S_n, appended (in order) to the primary sentence and to the primary as modified by previously appended secondaries. In Harris (1982) secondaries are connected to a primary by semicolon, where (e.g.,) a word in S_2 may be reduced to *wh-* pronoun on the condition that it is 'the same' as a word in S_1, as asserted in a metalinguistic 'sameness' statement (Chapter 5 §3): *Max bought a notebook; said notebook was green* → *Max bought a notebook which was green* → *Max bought a green notebook*. In English, all modifiers may be transformationally derived from this process of appended secondary sentence, statement of sameness, reduction to *wh-* pronoun, and in the case of most adjectives and some adverbs transposition of the modifier to the immediate left of its "host" (predicand) word in S_1.[1] Another example shows how this mechanism enables text sentences to be decomposed into instances of the recurring sentence types. A sentence like *the cells*

[1] Harris (1982:130 ff).

were found in antibody-producing lymph nodes (from 11,161.2.2) may be transformationally decomposed into a primary and an appended secondary: the cells were found in lymph nodes; said lymph nodes are antibody-producing whereupon the reduction to wh- pronoun, the cells were found in lymph nodes which are antibody-producing. Upon decompounding antibody-producing to produce antibody (or are producing antibody), factorization of the relative pronoun which into a wh- conjunctional element (a morphophonemic variant of semicolon) and an -ich proform, and substitution of the S_1 word (lymph node) for its proform in S_2, we have, in the normal form segmentation

¦ these cells ¦ were found in ¦ lymph nodes ¦

WH- ¦¦¦ antibody ¦ produce ¦ (lymph nodes) ← ¦

where the WH- element, as a conjunction, is partitioned off by three vertical bars.[1] The normal form segmentation is formulaically represented as

$$C\ W_i\ T_n^w$$

$$⦙\ A\ V_p\ T_n$$

the w superscript indicating the host word in S_1 to which the appended S_2 modifier is attached (i.e., operates upon).[2]

The transformational treatment of modifiers enables a single sentence to be structurally represented as composed of several already existing sentence types, rather than as a possibly ~~sui generis~~ unique type of sentence; a reduction in the complexity of grammatical description and an elimination therefore of redundancy of description. This is notably the case where

[1] The capitalization of wh- emphasizes its conjunctional status; parentheses enclose reconstructed material not appearing in the text.

[2] The subscripts designate a particular subclass of the word class; see below.

the modifier (operator) in the secondary sentence is a binary or ternary operator (i.e., has two or three argument places), as produce (binary) in the above example. Where S_2 introduces a unary (single place) operator upon a word in S_1, the objective of compressing as much predicational information as possible (modulo recurrence of existing sentence types) into a single sublanguage formula or row of conjoined formulas can be accomodated by indicating the secondary predication as a superscript to the word class symbol representing its host word. For example, among the modifiers of certain nouns (members of S, C or T classes) are included words of the W class which, in other of their occurrences, are the main (highest or latest entering [1]) operator upon their noun. Occuring as main operator, they are, together with their noun argument, an elementary sentence (an instance of an existing sentence type). Thus large occurs as main operator on the cells in the cells were large...(from 4,1.3.5), which is a CW_g instance of the CW sentence type. But upon entry of a further operator into the sentence upon their noun, they are modifiers (or adjuncts) of their noun, which becomes their host. For instance in the presence of large cells in these cultures...(from 7,12.2.1) large occurs as a local modifier of its host cells which is under the higher operator presence in. In the normal form segmentation, large as a left modifier is included within the segment of its host; by relinearization this is:

¦ of large cells ¦ the presence in ¦ these cultures ¦

[1] I.e. l.u.b. in the semi-lattice of the partial ordering.

In the formula representing this segmentation, the *g* subclass subscript of the W predicate *large* is written as a right superscript to the word class symbol of its noun host: $C^{g}\ W_{i}\ T^{u}$ (the *u* superscript similarly indicating ~~the~~ that *culture* is a modifier of a non-reconstructed *tissue*). Similarly, the W predicate *mature* occurs as the main operator in *the cells were fully mature*...(from 6,164.3.4), indexed $C\ W_{m}^{+}$, whereas *mature* occurs as local modifier of *plasma cells* under the higher operator *found to be present* in *mature plasma cells were found to be present in large numbers* (from 6,164.5.3), which is $C_{z}^{m}\ W_{i}^{+}$.

The inclusion of local modifiers (and sometimes other, e.g., M (metascience), material) within the segments of their host nouns obviously means that words belonging to different word classes are united within a single segment in the normal form projection. In principle there are transformational methods of constructing a segmentation in which only material belonging to a single word class appears, but once again the objective of eliminating redundancy from description and thus exhibiting, to the extent possible by the methods, the regularities of occurrence within the sublanguage overrules a more exact decomposition ~~if~~ when doing so allows representation of text sentences by existing sentence types. The amount of transformational decomposition of a sentence (reconstruction of less reduced stages in the derivation of that sentence) required to 'regularize' the characterization of the sentence cannot be stipulated *a priori*, and is, in the final analysis, a matter of overall best fit. However, it may be emphasized, this does not mean that the transformations employed are *ad hoc*. With noted exceptions, the transformations have applicability to the sentences of the entire language. Moreover, the actual significance of the use of transformational analysis in the sublanguage analysis can ultimately be attested by determining with what adequacy and facility they enable a representation of information in the sublanguage.

(d) Subclasses. The word [1] classes of the immunology sublanguage are set up on the basis of ~~determining~~ the operator/argument relations obtaining among the various words of the text sentences; as seen above, this is done initially by noting that some words have null requirement, that others are operators upon one or more null requirement words, that still others are operators upon these operators, and so on. Thus an operator word class V can be defined by determining which operator words require an argument word from the class A and possibly also one from any of the classes C, S or T; that is, V consists of the class of operator words occurring in environments receiving the 'regularized' grammatical designation as A _, A _ C, A _ S, A _ T, (or A _ SC, A _ CT, since we have fragments or sentences of the form S *of* C, C *in* T designated S W C and C W T, respectively). Having done ~~so~~ this, material like *the antibody detected in and around small lymphcytes* (from 9,68.3.5) and *antibody is not produced by small lymphocytes* (from 10,317.1.1) and *the antibody was formed by plasma cells* (from 6,164.5.5) may be segmented, respectively, as

| *the antibody* | *detected in and around* | *small lymphocytes* |

[1] Understanding 'word' here in the sense of operator grammar; i.e., as distinguished by predicational relations. In particular, as noted above, a sublanguage subclass may contain phrases as well as words, as these terms are normally understood.

antibody *is not produced by* *small lymphocytes* and

the antibody *was formed by* *plasma cells*.

Within the broad confines of the gross classification, each of these segmented sentences are formulaically representable as sentences of the AVC type. But a useful result obviously requires further differentiation within the gross word classes, i.e. setting up word subclasses. This is possible generally in two ways: determining whether (a) a word in a given class occurs only with a restricted set of words in its argument (or operator) class, or (b) words within a given class may be distinguished by the occurrence of different "local modifiers" operating upon them. It is this latter condition which enables *detected in and around* to be set apart from *is not produced by*, *was formed by* which in the above examples are all occurrences of words of the gross word class V. For we can find sentences or parts of sentences such as *the participation of lymphocytes in the formation of antibody* (from 4,12.4.1), *the lymphocyte was instrumental in the formation of antibody* (from 7,2.2.7), *lymphocytes act as producers of antibody* (from 3,128.3.1), *a minor contribution of lymphocytes to antibody synthesis* (from 8,58.3.3), and *the role of immature plasma cells in antibody formation* (from 7,14.2.1), whereas there are no sentences of the form *lymphocytes are instrumental in detecting antibody in and around them* or *plasma cells play a role in the presence of antibody*. The "local modifiers" *act as*, *have a role in*, etc. serve to distinguish *produce*, *form*, *synthesis*, as a subclass of words (designated V_p , with 'p' mnemonic for 'produce') within the gross class V. In turn, since we also have sentences like the antibody was formed by plasma cells (above), these local modifiers -- whose first argument is a word of the C word class and whose second argument a member of the subclass V_p -- themselves form a word class, designated by superscript 'r' (mnemonic for 'have a role in') which is attached to the symbol of their second argument in the formulaic representation. Accordingly, the above sentences and fragments containing members of V_p are represented as instances of the formula A V_p^r C.

Similarly, detected in, found in, contained members of V may be distinguished by the occurrence of particular modifiers upon them, e.g., titers of, concentration of as in lymphoid cells from minced lymph nodes also contained high concentrations of antibody (6,157.1.4) and the demonstration of antibodies in higher titer in the local lymphatic system than in the serum...(from 5,205.1.3); these form a subclass V_i (with 'i' mnemonic for 'in').[1] All other subclasses of V may be distinguished in this manner, e.g., a subclass V_s ('secrete') in a cell could be producing and secreting enough antibody to produce a rosette or plaque without containing, at a given time enough completed antibody to be detectable...(from 13,471.1.2), a subclass V_t ('store') which selects a unique member of S endoplasmic reticulum as in in some of the large lymphocytes ...the endoplasmic reticulum...channels were slightly and variably distended, and appeared to have deposits of protein-like material (from 13,456.2.1) and so on. Difference of

[1] On inclusion of 'metascience' terms like detect and demonstrate in word classes of the immunology sublanguage, see fn 1 p. 403 at the beginning of §3 below.

grammatical environment also serves to differentiate subclasses within the various argument word classes. For example, lymphocyte (C_y) may be distinguished from plasma cell (C_z) on the basis of sentences such as so far we can only state that the possibility that plasma cell's produce antibody is just as good as the possibility that lymphocytes do (3,128.5.1) and one of the major points at issue, then as now, was whether the plasma cell was derived from lymphocytes, or not (14,584.2.2).

(e) Sublanguage Transformations. With these comments in mind, we proceed to briefly consider how certain transformations, special to the sublanguage and hardly available in the grammar of the language as a whole, permit a further regularization of the analysis.

Some words (and word sequences) can be said to occur in text sentences in zero phonemic form in as much as the informational contribution they make to their sentence is negligent or highly redundant (expectable) in a given occurrence. Since all reductions and zeroings leave a trace in the environing words, namely, a recognizable departure from established word dependence requirements,[1] the reconstruction of these words in many cases enables a reduction in the apparatus of grammatical description, e.g., in showing that an apparent modifier of one word (which is not in its argument class, and so constitutes an apparent exception to be listed) is, in fact, a modifier of another word which occurs only in zero form. A striking instance of this may be seen in the homologous-antibody titers of extracts of a given lymph node (from 5,205.1.1) where homologous is an apparent modifier of

[1] Harris (1982:19):"Each reduction leaves a trace: the trace of a zeroing is a recognizable emptiness in matching the argument requirement of operators to their required arguments."

antibody. However, on the basis of sentences such as the difference in titers to the homologous and heterologous virus are clearly marked (from 5,204.5.2) and experiments in which opposite legs of each rabbit received injections of different serological types of influenzal virus (from 5,204.5.1) as well as the demonstration of antibodies in higher titer in the local lymphatic system than in the serum (from 5,205.1.3) we can reconstruct homologous-antibody titers of extracts of a given node as titers of antibody in extracts of a lymph node which is homologous (local) to the leg injected with a strain of influenzal virus where homologous asserts that the location of the lymph node from which the extracts were made is on the same side of the body (here, on the same leg) as is injected with a strain of virus (hence also the occurrence, above, of homologous and heterologous virus).

Another, more widely employed, reconstruction of 'sublanguage appropriate' low-information words enables the construction of many "macrosentences" of the GJB:AV(C or T) and GJB:C(or T)W type, i.e., an antigen injection sentence (GJB) conjoined (by a class of words represented by ':') to a 'response' sentence. The trace on which this reconstruction is based may be only the observable occurrence of a "local modifier" of the zeroed colon conjunction (e.g., after, following, etc.) whose reconstruction in turn serves as trace of a zeroed GJB first argument sentence. For example, in mature plasma *cells began to appear in large numbers only on the 4th day* (from 6,164.3.2) we can reconstruct *mature plasma cells began to appear in large numbers only on the 4th day after the injection of antigen*, a particular instance ($GJB :_t C_z^m W_i^{b+}$) of a GJB: response macrosentence (here the subscript *t* to the colon indicates the modifier *on the 4th day*, C_z^m represents *mature plasma cell*, and W_i^{b+} the W_i subclass member *appears*, with its aspectual modifier *began* and a quantity modifier *in large numbers*).

6.3 *Results of the Analysis*. The major word classes and some of the subclasses are listed below; a complete listing is given in Chapter 2 of FIS. We reiterate that the metascience (M) segments, grammatically distinguished as operating on sentences of the immunology sublanguage or segments of these, are a polyglot class including sentence introducers, assertion words, various kinds of connectives, and so forth which require further analysis and differentiation. At present, it may be said only that in most cases, M material can be separated on grammatical grounds from the immunology sentences proper.[1]

M verbs - *demonstrate*, *observe*, *detect*, *described* etc.

N' subjects of M verbs - *we*, *investigators*, *Sundberg*, etc.

M' procedural terms - *prepare*, *extracted*, *centrifuged*, etc.

M'' assertion (O_o) operators - *is probable*, *is insignificant*, etc.

M conjunctions - *indicates*, *has obvious bearing on*, *therefore*, *however*, etc

[1] This is not always readily achievable however, e.g., in *only insignificant amounts of antibody were detected in the follicle culture fluids* (from 4, 12.4.3) the 'meta' terms *insignificant* and *detected* are included within the segmentation of the immunology sentence.

Representative members of the major word classes and some of the subclasses of the immunology sublanguage are listed below. The status of members of word classes is established in that free substitution of one member for another in the grammatically characterized word-class environment yields a resultant sentence which may be said to belong to the immunology sublanguage, though it may not occur in the analyzed corpus or an expanded corpus, and may in fact be contradictory or unacceptable to users of the language ('sublanguage informants'). Word class members are, according to the term revived by Geach, substitutable salva congruitate: replacement yields a syntactically (i.e., grammatically) coherent string of words. In general, word class members are not freely replacable in a stronger sense, salva veritate or salva significatione. For example, in a cell could be producing and secreting enough antibody to produce a rosette or plaque without containing, at a given time, enough completed antibody to be detectable (from 13,471.1.2) produce (in either of its occurrences) is not substituable for secrete or contain though all are here members of the gross word class V. Nor is antigen, a 'classifier noun', substituable, in either of the stronger senses, for its 'classificands' paratyphoid bacterin or diptheria toxin in inflammation was induced in the cervical nodes of both sides and in both ears by the injection of paratyphoid bacterin on one side and diptheria toxin on the other (from 1,791.3.1) though these are here members of the gross word class G ('antigen'). On the other hand, members of word subclasses can, in general, be said to be "locally synonymous", i.e., replacement of one member by another in a sentence preserves, ex hypothesi, information. Thus, produce, synthesize, members of the subclass V_p, may be substituted (as was produced by, was synthesized by) for form (another member of V_p) in e.g., the antibody was formed by plasma cells (from 6,164.5.5). Again, in there were regularly present, in sections of the homolateral popliteal lymph node stained for antibody, large cells (from 9,64.5.1) stained for, occurring here as a member of V_i may be replaced by other members of V_i occurring in the environment regularized as A__T which, as stained for, select a member of T as first argument: containing, positive for, etc. Other members of V_i occurring in this environment select A as first argument: present in, found in, visible in (or visibly present in), demonstrable in (demonstrably present in), detectable in (detectably present in). Replacement in these 'inverse' cases requires transformation; for example, in the above case, the A V T segment may be transformed under replacement to sections of the homolateral popliteal lymph node in which antibody was visibly present. As the above examples show, "local synonyms" of the immunology sublanguage in many cases can hardly be considered to be synonymous in the language as a whole, i.e., English.

G antigen, killed cholera spirilla, S. typhi, B. prodigiosus

J injected, administered, immunized

B mice, rabbits, hind foot of the rabbit

U travels, is distributed to

A antibody, agglutinin, rosette (A_r)

V is present in (V_i), is produced by (V_p), secretes (V_s)

T tissue. liver (T_v), lymph node (T_n), serum (T_b)

C cell, plasma cell (C_z), lymphocyte (C_y), blast (C_b)

W enlarged (W_g), enflamed (W_f), mature (W_m) eccentric (W_e)

S cellular ultrastructure, nucleus (S_n), mitochondria (S_m)

Y were identified as, are, according to our terminology, develop into (Y_c^t)

I were injected into...from (I^{tf})

: after, following

These combine into the following sentence types (formulas of information):[1]

A V T the tissues demonstrably contained antibody

A V C antibody is synthesized by plasma cells

A V S C antibody is distributed in the perinuclear space of the cell

G J B antigen was administered intravenously in rabbits

G U T antigen is absorbed by the serum

G U C the uptake of antigen by the cell

S C W the ergastoplasmic cisternae of the cells became distended

C W the cells multiply

C W T these cells are scattered in the medullary cortex

C Y C these cells eventuate in mature plasma cells

S Y S these structures include Golgi bodies

C Y C C large lymphocytes are regarded as transitional between small lymphocytes and plasmablasts

[1] This list is not exhaustive; subclass designations omitted.

The sentence types, in turn, combine into sentence sequence types ("macrosentences"):

1) G J B : response sentences (i.e., sentences whose main operator is a member of V, W, or Y_c), e.g.,

G J B : C W T - After intravenous injections of horse serum into rabbits, cellular changes occurred in the spleen

G J B : A V C - On the first few days after the last of a series of injections, antibody is present in groups of plasma cells

G J B : C Y_c C - Members of this cell family arise from some undifferentiated precursor as the direct result of antigenic stimulus.

2) G J B : G U T - After injection antigen is carried by the lymph

3) G J B_1: C $I^{ft}B_1B_2$ - Cells taken from sensitized rats are injected into other rats

4) G U T(or C) : response sentences - The synthesis of antibody is not the usual result of the uptake of antigen by the cell

5) C $I^{ft}B_1B_2$: response sentences - Antibody response can be restored by injecting small lymphocytes from other rats of the same highly inbred strain.

In these examples, subclass and local modifier designations have largely been omitted; see further in § 4 below.

6.4 <u>Adequacy as a Representation of Information.</u> A completely adequate informational representation of these articles in cellular immunology would involve extending the analysis to the M portions of the text (exhibiting thereby something of the hierarchical connections between sentences -- argument structure) as well as to the detailed descriptions of procedures, techniques and measurement methods used in the experiment ("Methods and Materials" sections). Such an analysis would appropriately be comparative, seeking to establish which, if any, portions of the extended analysis were special to the subfield of science treated here, and which were of more general applicability. This is obviously a major undertaking for the future. However, the adequacy of the formulas for what has been termed above the "immunology sublanguage" can be ascertained in several ways.

First, it can be seen that over the 30 year course of research, the formulas changed in ways that accord with the known development within the subfield. Paper 1 experimentally demonstrates the fact of antibody formation in lymph nodes ($A\ V_{p}\ T_{n}$) by a series of experiments carefully designed to eliminate the possibility that the antibody found there had been produced elsewhere ($A^{w}\ V_{i}\ T_{n} \vdots\vdots\vdots\ A\ V^{\sim}_{p}\ T_{n}$). No speculations are made in this early work concerning types of lymph node cell which may be responsible for antibody formation ($C^{w}\ W_{i}\ T_{n} \vdots\vdots\vdots\ A\ V^{r}_{p}\ C$). Paper 2 argues for lymphocytic involvement ($A\ V^{r}_{p}\ C_{y}$) on the basis of results showing that antibody is contained within lymphocytes ($A\ V_{i}\ C_{y}$); however, as the authors are careful to note, this finding does not establish that antibody is actually produced by lymphocytes

($A\ V_p\ C_y$). The authors of paper 3 employed hyperimmunized animals, i.e., animals repeatedly injected with antigen ($G\ J^3\ B$), a technique which (it was later recognized) results in massive plasma cell proliferation ($C_z\ W_p^+\ T$) and a relative diminishment of the lymphocyte population ($C_y\ W_i^-\ T$). As plasma cells were therefore the predominate cell type in tissues (fat of the renal sinus) found to contain high concentrations of antibody ($C_z\ W_i^{++}\ T_k^w \vdots\vdots\vdots\ A\ V_i^+\ T_k$), the authors concluded that plasma cells are responsible for this high concentration of antibody ($A^w\ V_p^r\ C_z \vdots\vdots\vdots\ A\ V_i^+\ T_k$) and lymphocytes only improbably involved ($A\ V_p^{r\sim} C_y$). Paper 4, again with hyperimmunized animals ($G\ J^3\ B$), reports the results of <u>in vitro</u> experiments which sought to determine the antibody-forming capacity of two different tissues, red splenic pulp and follicular tissue ($A\ V_p^k\ T_d$, $A\ V_p^k\ T_f$) which are primarily plasmacytic ($C_z\ W_i^+\ T_d$) and lymphocytic ($C_y\ W_i^+\ T_f$), respectively, in composition. The author finds the red splenic tissue to have a far greater capacity to form antibody <u>in vitro</u> ($A\ V_p^{k>+} T_d^u \vdots\vdots\vdots\ A\ V_p^k\ T_f^u$) and observes that plasma cells appear to develop from reticulum cells ($C_r\ Y_c^f\ C_z$). She concluded that large reticulum cells (C_r^g), which she terms "transitional cells" ($C_r^g\ Y\ C^c$), developing from reticulum cells, produce antibodies ($A\ V_p\ C^{cw} \vdots\vdots\vdots\ C_r\ Y_c^f\ C^c$), thereupon developing into a type of cell with the morphological characteristics of plasma cells ($C^c\ Y_c^t\ C_z$), and that her results establish no evidence of the participation of lymphocytes in the production of antibody ($A\ V_p^{r\sim} C_y$). Paper 5 examines the antibody content of tissues of the popliteal lymph node of rabbits, which is the sole node draining the site of antigen injection in the footpad (G J B : A V_i T_n^B). The authors conclusively demonstrate antibody specific to the antigen injected in the respective footpad is formed in the homologous "local" lymphatic system (G J B : A^G V_p T_n^B) and that the activity of this lymph node (T_n W_c) is characterized by a marked enlargement (T_n W_g) which is due to lymphocytic hyperplasia (C_y W_p^+ T_n). This, they conclude, is evidence for the lymphocyte as a primary source of antibody (A $V_p^{r+}C_y$). Paper 6 compares the formation of antibody protein within lymph nodes (A V_p T_n) with changes in the nucleic acid content of lymph node tissue (D V_i^{Δ} T_n). The highest concentration of RNA, known to be associated with protein synthesis, were found (D_r $V_i^{++}T_n$) at a time when mature plasma cells were present in highest number (C_z W_i^{++} T_n), whereas, at this time, lymphocytes were only in early stages of proliferation (C_y W_p^- T_n). This leads the authors to surmise that the antibody found in lymph nodes was formed by plasma cells (A^w V_i T_n $\vdots$ A V_p C_z). Paper 7 repeats the <u>in vitro</u> experiments of Fagraeus (paper 4) finding, contrary to her results, that antibody is found in follicular tissue which is largely lymphocytic (A V_i T_f^w $\vdots$ C_y W_i^+ T_f) as well as in the plasmacytic red splenic pulp (A V_i T_d^w $\vdots$ C_z W_i^+ T_d). However, the red splenic pulp was determined to contain approximately twice the antibody content of the follicular tissue (A $V_i^>$ T_d $\vdots$ A V_i T_f). Moreover, the difference in antibody content between these tissues was found to be strikingly proportional to the presence, in cultures of these tissues, of large cells (A V_i $T_{d,f}^u$ $\vdots$ C^g W_i $T_{d,f}^u$), which were identified as immature plasma cells (C^g Y $C_z^{m\sim}$). Smaller cells, lymphocytes, were found to be not capable of producing or multiplying antibody under the conditions of the experiment ($A\ V_p^{r\sim} C_y^w$ ⁞⁞ $C_y\ W_i\ T^u$). Papers 8 and 9 use a specific fluorescence-staining technique developed by one of the authors (Coons) to detect antibody within individual cells ($A\ V_i\ C^1$) for the first time. They report that the major site of antibody formation is a family of cells which first appears as a response to antigenic stimulation ($A\ V_p^r\ C^{\ell w}$ ⁞⁞ $G\ J : C^\ell\ W_i^b$). This response consists of cell multiplication ($G\ J : C^\ell\ W_p$) and cell differentiation ($C^\ell\ W_c$), of which the end stage is the plasma cell ($C^\ell\ Y_c^t\ C_z$) and the concurrent synthesis of antibody ($A\ V_p\ C^\ell$). However, small amounts of antibody were visible in individual lymphocytes in the lymphoid follicles of the spleen and lymph nodes ($A\ V_i^-\ C_y^{lw}$ ⁞⁞ $C_y^1\ W_i\ T_{f,n}$). Hence the authors caution, a minor contribution of the lymphocyte to antibody synthesis cannot be excluded ($A\ V_p^{r-}\ C_y$). Paper 10 employs a novel technique developed by one of the authors (Gowans) in which animals from which small lymphocytes has been removed by chronic drainage from a fistula inserted into the thoracic duct ($C_y^{g\sim}\ W^f\ T_h$), show a normal secondary response, i.e, response to secondary injection ($C_y^{g\sim} W^f\ B^w$ ⁞⁞ $G\ J^2\ B : A\ V^+\ B$), but only an impoverished primary response ($C_y^{g\sim} W^f\ B^w$ ⁞⁞ $G\ J^1\ B : A\ V^-\ B$). The primary response of lymphocyte-depleted animals can, however, be restored by injecting small lymphocytes obtained from normal rats ($C_y^{g\sim} W^f\ B_1$) of the same highly inbred strain ($C_y^{g\sim} I^{ft} B_1\ B_2$ ⁞⁞ $G\ J^1\ B_2 : A\ V^+\ B_2$). These results unequivocally showed that the depletion of small lymphocytes resulted in a severe immunological deficiency ($C_y^{g\sim} W^f\ B^w$ ⁞⁞ $A\ V^-\ B$), and hence established the fundamental importance of the small lymphocyte in the primary immune response ($A\ V_p^{r+}\ C_y^{g\sim}$ ⁞⁞ $G\ J^1 : A\ V$). The authors also pointed to the fact that the small lymphocyte can develop rapidly into a "large pyroninophilic cell" which divides ($C_y^{g\sim} Y_c^{it}$ C_h^w ⁝ C_h W_p) and morphologically resembles a plasma cell precursor (C_h Y C^w ⁝ C Y_c^t C_z). They thus hypothesize that the small lymphocyte may be the ultimate precursor of antibody-forming plasma cells ($C_y^{g\sim}$ Y_c^t C_z^w ⁝ A V_p C_z). This hypothesis is confirmed by the electron microscopic observations reported in papers 11 and 13. However, in paper 12, electron microscopic observations trace the ultrastructural distribution of antibody (A V_i S C) in a series of cells, beginning with a cell termed "hemacytoblast" and eventuating in mature plasma cells some three to five days after a booster injection of antigen (G J^2 $:_t$ C_b Y_c^t C_z^m). These results are taken to have demonstrated the fact of antibody production by cells of the plasmacytic series (A V_p C_z^ℓ) as posited in papers 8 and 9 on the basis of immunofluorescence techniques. Since the electron microscopic observations did not show the presence of antibody in small lymphocytes (A V_i $C_y^{g\sim}$), the implication is that the plasmacytic and lymphocytic families are distinct (C_y^ℓ $Y_c^\sim$ C_z^ℓ). Papers 11 and 13 present the results of electron microscopic observations of single cells which could be determined to be antibody-producing by an novel technique of mounting suspensions of individual cells on an antigenic preparation and observing whether a plaque or rosette was formed around them (A_g V_p C^1, A_r V_p C^1). The production of a plaque or rosette indicates that the cell is both producing and synthesizing antibody (A $V_{p,s}$ C^1). Paper 11 observes two classes of lymph node cells with distinct morphological features which were found to produce plaques after a single injection of antigen (G J^1 : A_g V_p $C^{\ell 2}$, A_g V_p $C^{\ell 1}$). These cells were of the category lymphocyte or plasma cell ($C^{\ell 1}$ Y C_y, $C^{\ell 2}$ Y C_z).

This finding is despite the prevalent belief that the synthesis of antibody could not be found in the lymphocyte, with its paucity of cytoplasmic differentiation (C_y S_c W_c^-). However, under the electron microscope, lymphocytes were observed to contain many of the ultra-structural components associated with protein systhesis (S^w W_i^+ C_y ||| A_p V_p^r S). These structures include Golgi bodies, nucleoli, and short channels of endoplasmic reticulum (S Y S_g, S_u, S_r^{h-}). In view of the relatively undeveloped endoplasmic reticulum of these lymphoctes (S_r^{h-} W_i C_y), the production of antibody by these cells may represent an instance in mammalian cells of a secreted protein synthesized by ribosomes not bordering on an organized endoplasmic reticulum (A_p^w V_p S_b^w C ||| A_p V_s ||| S_b $W_i^\sim$ S_r^h C). Paper 13 extends these electron microscopic observations to rosette-producing cells. Here the classification of cells into the lymphocytic or plasmacytic series (C Y C_y^i, C_z^i) was determined by the state of the endoplasmic reticulum of the cell (C S_r W_c^o). In some large lymphocytes organization of, and storage of protein within, the endoplasmic reticulum could be observed (C_y^g S_r^h W_c, A_p V_t S_r^h C_y^g). These large lymphocytes which still had a predominately ribosomal cytoplasm (S_b W_i^+ S_c C_y^g), were regarded as transition between small lymphocytes and early plasma-blast cells (C_y^g Y_c^{ft} $C_y^{g\sim}$ C_z^b). However, since small lymphocytes were also found to produce rosettes (A_r V_p $C_y^{g\sim}$), it is concluded that antibody production may start even before the development of the large lymphocyte or blast forms (A V_p^b ||| C_y^g Y_c^{ft} $C_y^{g\sim}$$C_z^b$). The rosette and plaque methods make possible the detection of individual antibody-producing cells which continuously secrete antibody and therefore do not contain enough antibody to be detected even under electron microsopic observation (A V_p C^{1w} ¦¦¦ A V_p C^1 <u>and</u> A V_s C^1 <u>and</u> A V_i C^1). Finally, paper 14 surveys the entire investigation.

In addition to providing a conveniently concise and informationally precise summary of the results of these investigations, the output of the analysis in some cases can serve as a basis from which to actually critique the presentation of these results. For example, as the research problem which is the subject of these articles was resolved by the finding that a single cell line (but with lymphocytic and plasmacytic stages) produced antibody, the individuation and naming of the various cell stages which could be distinguished was of primary importance. Given the outcome, however,the names of the different stages, rather than designating truly distinct entities, are less misleadingly construed as abbreviating certain characterizations of the morphologically different stages. With the advent of new techniques (e.g., plaque and rosette formation), individual cells could be identified as antibody-producing, even if (and somewhat paradoxically in the light of a long-standing assumption that an antibody-producing cell must demonstrably contain antibody) they could not be shown to contain antibody. Under the electron microscope, which greatly heightened the threshold of observation, new details were revealed of the cellular ultrastructure of antibody-producing and antibody-secreting cells. It thus becomes possible to see that what is termed the transition of one cell into another consists in specifiable changes along a number of distinct ultrastructural parameters. And, the grammatical characterization of this transition ($C_1\ Y_c^t\ C_2$ or $C_2\ Y_c^{ft}\ C_1\ C_3$) is, in effect, a **form** of short hand or 'classifier' of the various kinds and degrees of ultrastructural change (C S W). We have here the basis for reducing the Y_c operator to some conjunction of C_1 S W sentences, with C_2 as an appended meta-scientific designation for some suitably recognizable or agreed-upon stage characterized by a given aggregate of conjoined C S W sentences.[1]

For instance, as summarized above, the three electron microscopy papers (11-13) individuate cell types primarily by a characterization of the state or degree of organization of a cytoplasmic component termed "ergastoplasm" or "endoplasmic reticulum". In paper 11, the endoplasmic reticulum of the small lymphocyte is described as rough ($C_y^{g\sim}\ S_r\ W_r$) while that of the medium-sized lymphocyte (C_y^{g-}) is similarly characterized in paper 13. Whereas paper 13 reports small lymphocytes only with endoplasmic reticulum which is not widened ($C_y^g\ S_r\ W_{w\sim}$), paper 11 indicates some small lymphocytes having small, but widened amounts of endoplasmic reticulum ($S_r^w\ W_i^-\ C_y^{g\sim}\ S_r\ W_w$). Channels in the endoplasmic reticulum are reported in paper 11 in the large lymphocyte ($S_r^h\ W_i\ C_y^g$); these are not widened ($S_r^h\ W_{w\sim}$) and not parallel or lamellar ($S_r^h\ W_{y\sim}$). Organization of the endoplasmic

[1] This suggestion is due to Harris; Chapter 3 of FIS.

reticulum into cisternae-like structures is reported for the cell termed "hemacytoblast" ($S_r^s\ W_i\ C_b$) in paper 12. Paper 13 reports the transition of the large lymphcyte to the plasmablast ($C_y^g\ Y_c^t\ C_z^b$) is marked by various and small degrees of widening in the channels of endoplasmic reticulum ($S_r^h\ W_w^{o,-}$) together with a more nearly parallel-orientation of the channels ($S_r^h\ W_y$), whereas paper 12 shows that the transition of the cell, denominated "hemacytoblast", to plasmablast is marked primarily by the gradual development of the ergastoplasm or ribosome-associated endoplasmic reticulum ($S_r\ W_c^o$ or $S_r^w\ W_c^o \vdots\vdots\vdots\ S_b\ W_i\ S_r$). Thus a commonly recognized cell-stage (the plasmablast) is stated to arise from differently named cell stages (in two different papers), via a process of change which is similarly characterized. This permits a question to be raised concerning the relation between the cell stage named "large lymphocyte" (in paper 13) and that termed "plasmablast" (in paper 12). Through a close comparison of the sublanguage formulas corresponding to the ultrastructural descriptions given these differently named cell stages, the possibility is suggested that they are more appropriately noted as one cell stage, a clarification which, though perhaps merely "a matter of semantics",[1] (it can be conjectured) would have had considerable importance at the time.

contains?

[1] Paper 14 cites the following comment (584.2.6): "Exactly how much rough-surfaced endoplasmic reticulum a lymphocyte should display ...to be considered a full-fledged plasma cell, becomes a matter of semantics."

makes or contains the comment, or cites it as made elsewhere?

The increasingly fine level of ultrastructural description may also be seen as giving a more precise content to certain of the W operators on C, such as W_c (change, develop, differentiation) W_a (reaction, active), and even W_m (mature) and its negative $W_{m\sim}$ (immature). Here it is interesting to observe that these highly phenomenological or imprecise terms which are 'born' at the level of light microscopy to describe or otherwise characterize states or processes not further articulable at that level of observation, can be viewed as placeholders or incipient classifiers of the more refined (yet still phenomenological, e.g., rough, parallel) description of cellular ultrastructure subsequently made available with electron microscopy. It seems quite reasonable to suppose that replacements of this kind, i.e., of one form of discourse by another which is grammatically more articulated (as is indicated by the appearance of new subclasses, or by the restrictions of words of a particular word class to certain words of another class), are one way of gauging when, and at what points, a science or subfield of a science had changed, or is changing, and may even indicate where further change is likely to occur.[1]

[1] For further discussion, see Ryckman and Gottfried (in press).