The Methodology off Sememic Analysis with Special Application to the English Preposition*

by James H. White, Mechanolinguistics Project, The University of California, Berkeley; Consultant, System Development Corporation, Santa Monica, California

This paper summarizes stratificational theory, and applies its linguistic methods in a sememic analysis of English prepositions. The phenomenon of interlocking diversification is shown to be quite generally present among the prepositions. Also, the analysis of prepositions is shown to entail a partial sememic analysis of other words; it therefore provides a starting point for the sememic analysis, on stratificational principles, of the rest of the language.

Introduction

In the past two decades, a number of linguistic theories have been developed whose primary purpose is to give a formalized method of handling linguistic data, i.e., natural language. One of the significant reasons for this is that in recent years problems of linguistic automation and machine translation have required that a great deal more structure be given to linguistic theory than had been previously. One of the more significant theories or models set up to handle ordinary linguistic data as well as the problems of these new fields is the stratificational theory of S. M. Lamb.[1,2] In this model, languages are viewed as complex systems whose structures are made up of a series of strata which are related by certain linguistic rules. These rules, called rules of realization, make it possible to deal systematically with the linguistic entities which exist on a certain stratum and the relation of those entities to neighboring strata. This paper will describe the stratificational theory with emphasis on sememic analysis and then will give in detail a sememic analysis of the major prepositions in the English language and some conclusions about the linguistic environment of the individual prepositions.

The Stratificational Theory

THE STRATA

The strata of written language have been given the names graphemic, morphemic, lexemic, and sememic—the graphemic being the lowest stratum and the sememic being the highest stratum. The graphemic stratum has letters or symbols and strings of letters of symbols. The morphemic stratum has segmented strings of letters which are minimal meaningful elements. The lexemic stratum combines the strings of meaningful elements into the proper meaningful expressions. Finally, the sememic stratum has the structural elements of meaning in a given concept.

A few examples will give a good indication of the differences between the strata. Consider the *-es* in the noun *taxes* and the *s* in the noun *books;* graphemically these are different, but morphemically they are the same entity which can be labeled ^{M}/s/, where the M indicates that the item between the / / is a morphemic entity. The reason for this is that the *-es* is completely predictable after such an expression as *tax,* or that no reason of meaning requires the *-es,* but only a reason of spelling.

Next, consider the two expressions *good* and *better;* morphemically these are entirely different, but lexemically they are partially the same. *Better* from the lexemic point of view consists of a *good* followed by the comparative suffix, the lexeme ^{L}/-er/; here the L indicates that the item between the / / is a lexemic entity.

Finally, *can* and *be able to* are lexemically different, but they are both the same sememe, ^{S}/can/; here the s indicates that the expression between the / / is a sememic entity.

Some expressions as they appear on each of the strata are:

waterfalls

1. Graphemically:
   ^{G}/w + a + t + e + r + f + a+ l + l + s/
2. Morphemically:
   ^{M}/water + fall + s/
3. Lexemically:
   ^{L}/waterfall + plural/
4. Sememically:
   ^{S}/WATERFALL + plural/

rams

1. Graphemically:
   ^{G}/r + a + m + s/
2. Morphemically:
   ^{M}/ram + s/

3. Lexemically:
   ^{L}/ram + plural/
4. Sememically:
   ^{S}/SHEEP + male + plural/

fillies

1. Graphemically:
   ^{G}/f + i + 1 + 1 + i + e + s/
2. Morphemically:
   ^{M}/filly + s/
3. Lexemically:
   ^{L}/filly + plural/
4. Sememically:
   ^{S}/HORSE + young + female + plural/

The units on these strata have the names: grapheme, morpheme, lexeme, and sememe, respectively. Thus, in the example of *waterfalls:* ^{G}/w/ is a grapheme; ^{M}/water/ is a morpheme; ^{L}/waterfall/ is a lexeme; and ^{S}/WATERFALL/ is a sememe. There is certainly more than one structural element of meaning contained in the concept *waterfall.* These components may be called semons. Figure 1 shows the strata and the entities.

Strata	Entities
sememic	sememe, semon
lexemic	lexeme
morphemic	morpheme
graphemic	grapheme

A CHART OF LINGUISTIC STRATA AND ENTITIES:
STAGE 1

FIGURE 1

On each stratum certain operations may be performed on these entities: combinations and classes may be formed. For example, on the graphemic stratum one may form vowel or consonant classes or perhaps classes of mathematical or biological symbols. On the lexemic stratum there are classes of nouns and verbs, prepositions and adjectives, etc. On the sememic stratum one may classify sememic entities which have certain basic semons in common such as the semon ^{S}/sense/ which can be found in the sememic entities ^{S}/THINK/, ^{S}/TELL/, ^{S}/KNOW/, ^{S}/SEE/, etc.

There are several types of combinations. On the lexemic stratum there are tactic rules which show how to combine the classes of nouns, verbs, adjectives, etc. to get clauses. On the sememic stratum, there are semotactic rules which explain how to put semons together to get discourse blocks. Farther down the scale there are graphotactic rules which tell how to put letters and symbols together to form syllables.

REALIZATION THEORY

With this brief discussion of the operations which may be performed *on* each stratum, we come to a relation which exists *between* neighboring strata. This relation is known as "realization." Realization is a relation of a higher stratum to a lower stratum. (Incidentally, this was the initial reasoning for the ordering of the strata in Figure 1.) Linguistically, an entity on a certain stratum has a realization (or realizations) on the next lower stratum. Thus, for example, a sememe would have realizations on the lexemic stratum, or the morpheme would have realizations on the graphemic stratum.

The realizations of the units of the higher strata have been given special names. Realizations of morphemes are called "morphs"; realizations of lexemes are called "lexes"; realizations of sememes are called "semes". In general, morphs are combinations of graphemes, and lexes are combinations of morphemes. However, semes are usually single lexemes. Figure 2 relates these entities to the strata.

Stratum	Basic tactic unit	Realization unit
sememic	sememe	
lexemic	lexeme	seme
morphemic	morpheme	lex
graphemic	grapheme	morph

A CHART OF LINGUISTIC STRATA AND ENTITIES:
STAGE 2

FIGURE 2

We complete the full picture of the stratificational theory with the introduction of the "realizates." The realizates are those elements on the higher stratum which are realized on the lower stratum by the realizations. Thus a morpheme is the realizate of the morph, or again, the lexeme is the realizate of the lex. The grapheme, morpheme and lexeme also have realizates. The realizate of the grapheme is the morphon. The morphons are the elements which make up the morpheme. Thus, for example, the morphons ^{M}/w/, ^{M}/a/, ^{M}/t/, ^{M}/e/, ^{M}/r/ make up the morpheme "/water/. One way to express it is to say that the morphon is a graphemic-sized element of the morphemic stratum.

The realizate of the morpheme is the lexon. The lexons are the entities which make up the lexemes. Continuing the example, we find that it is the lexons ^{L}/water/ and ^{L}/fall/ that go to make up the lexeme ^{L}/waterfall/. Once again in the terminology above a lexon is a morphemic-sized element of the lexemic stratum.

Lastly, the realizate of the lexeme is the semolexeme. Here the situation is different, because the semolexemes are often not elementary units, i.e., semons, but rather are composed of semons; for example, the semolexeme ^{S}/RAM/ is composed of the semons ^{S}/male/ plus the bundle of semons in the concept ^{S}/SHEEP/, or in other words, the semolexeme ^{S}/SHEEP/. The sememe is the unit of the sememic stratum which corresponds to a

referent. It often coincides with the semolexeme. Figure 3 summarizes the structural units discussed here.

Stratum	Elementary component	Basic tactic unit	Realization unit	Complex tactic unit
sememic	semon	sememe		discourse block
lexemic	lexon	lexeme	seme	clause
morphemic	morphon	morpheme	lex	word
graphemic	graphon	grapheme	morph	syllable

A CHART OF LINGUISTIC STRATA AND ENTITIES:
STAGE 3

FIGURE 3

PHENOMENA OF REALIZATION

The phenomena of realization fall into two categories for the most part: "vertical discrepancy" and "horizontal discrepancy." One type of vertical discrepancy is "diversification." Diversification occurs when an element of a higher stratum has more than one realization on the lower stratum. For example, the morpheme ^{M}/s/ is realized by the two morphs (called allomorphs when there are more than one) ^{G}/s/ and ^{G}/es/ as in the expressions *boxes* and *books;* again, the lexeme ^{L}/good/ has three allolexes: ^{M}/good/, ^{M}/bett/, and ^{M}/be/ as in the expressions *good, better* and *best.*

Another type of vertical discrepancy is "neutralization." Neutralization is the opposite of diversification; that is, two elements of the higher stratum are said to be neutralized when they are realized by the same element on the lower stratum. For example, the two lexemes ^{L}/plural/ and ^{L}/third-person-singular-present-tense/ both have the same realization on the morphemic stratum, namely ^{M}/s/. Also the sememes ^{S}/LARGE/ and ^{S}/IMPORTANT/ may both be realized by the lexeme ^{L}/big/. Consider the two expressions *the big rock* and *he is a big man around the town.* One *big* is in free variation with *large* and the other is in free variation with *important;* in other words, *large* can be substituted in the first expression without a change of meaning and *important* can be substituted in the second without a change of meaning. The significance of this type of decision will become very clear later on.

Horizontal discrepancy is divided into two main types: "composite realization" and "portmanteau realization." Composite realization is present when an element of a higher stratum is realized by a combination of elements on the next lower stratum. For example, the sememe ^{S}/WATERFALL/ is realized by the two lexons ^{L}/water/ and ^{L}/fall/; the lexon ^{L}/pin/ is realized by the three morphons ^{M}/p/, ^{M}/i/, ^{M}/n/.

Portmanteau realization is the opposite of composite realization; that is, a combination of elements on the higher stratum is a realizate of a unit on the next lower stratum. Striking examples of this type of discrepancy occur between the sememic and lexemic strata. For example, the combination or bundle of semons which form the semolexeme ^{S}/RAM/—among which are ^{S}/male/ and ^{S}/SHEEP/—are realized by the single lexeme ^{L}/ram/. Another example between two different strata is the realization of the two lexons ^{L}/bad/ and ^{L}/er/ by the single morpheme ^{M}/worse/.

Other types of discrepancy exist but are not essential for this paper and so will be omitted for the present. Table 1 shows the two kinds of discrepancy and two types of each which we have discussed, with examples.

Vertical discrepancy:

A) Diversification: ^{M}/s/ :R: ^{G}/s/ and ^{G}/es/
B) Neutralization: ^{L}/plural/ and ^{L}/third-person-singular-
present-tense/ :R: ^{M}/s/

Horizontal Discrepancy:

A) Composite realization: ^{L}/pin/ :R: ^{M}/p + i + n/
B) Portmanteau realization: ^{S}/male + SHEEP/ :R: ^{L}/ram/

In the table ":R:" is to be read: "is (are) realized by."

KINDS OF LINGUISTIC DISCREPANCY

TABLE 1

LINGUISTIC ANALYSIS

As a result of these various types of discrepancies, there are various types of analyses necessary for a complete stratification of the language. This section discusses some of these types of analyses. The analysis problem is this: what does the language and the utterances it produces look like on each of the strata? We have indicated what individual expressions look like on each stratum* but not what an expression such as *he hit the ball with a bat* would look like on each of the strata. This problem is solved partially by the methods of analysis described below.

The first method is that of "grouping" which is necessary because of diversification. If two or more units of a lower stratum realize the same unit of a higher stratum, they are grouped together. Thus in our example of the sememe ^{S}/can/, we must group together the lexemes ^{L}/can/ and ^{L}/be able to/. Or, again, in the case of the lexon ^{L}/plural/, we group together the morphemes that realize it: ^{M}/s/ as in the expression *books,* ^{M}/en/ as in *oxen,* ^{M}/ren/ as in *children,* ^{M}/Ø/ as in *deer.* Finally, if we take the morpheme ^{M}/s/, we group together the combinations of graphemes which realize it: ^{G}/s/ and ^{G}/es/.

The second method is that of "differentiation" which is necessary because of neutralization. We recall that neutralization is present when two or more entities on the higher stratum are realized by the same element on the lower stratum. There are three types of criteria for differentiation. First there is "interlocking diversi-

* See pages 15-16.

fication." Here the presence of neutralization is discovered because it is interlocked with a case of diversification. For example, the morpheme ^{M}/-ed/ is a neutralization of the lexemes ^{L}/past-tense/ and ^{L}/past-participle/. However, the lexeme ^{L}/past-participle/ also has the realization ^{M}/-en/, but the ^{L}/past-tense/ lexeme does not. Hence we are able to differentiate the two lexemes which are realizates of ^{M}/-ed/.

The second of the three criteria is that of finding "different portmanteau analyses." Take for example the lexeme ^{L}/soft/. This has two different realizates on the sememic stratum, and the way to differentiate them is by the presence of different portmanteau analyses as shown in the tables:

| loud | loudness | soft |
| hard | hardness | soft. |

Here ^{L}/soft/ realizes two semon bundles ^{S}/not + loud/ and ^{S}/not + hard/.

The third criterion, perhaps the most often used, is that of "distribution." The lexeme ^{L}/big/ has several different sememic realizates and the following is the way to distinguish two of them. ^{S}/big/, as in the expression *the big rock,* can occur in *the rock is big* and still retain the same meaning. However, ^{S}/big/ in the expression *the big fool* does not have the same distributional freedom, so we must set up two different sememic units to take care of the situation.

Another type of stratificational analysis is "segmentation," which is necessary because of horizontal discrepancy. The most obvious example of this is the segmenting of strings of graphemes into morphs—a situation arising because of composite realization. The string of graphemes ^{G}/w + a + t + e + r + f + a + 1 + 1 + s/ must be segmented into the morphs ^{G}/water + fall + s/. On a higher stratum, such as the lexemic, we must segment idiomatic phrases which represent a single sememe, such as the strings of lexons ^{L}/with regard to/ or ^{L}/call up/ as in *call up on the phone.*

Table 2 shows the different types of discrepancy and the methods of analysis that correspond.

Grouping ....necessary because of diversification
Differentiation ....necessary because of neutralization

    Criteria:

      1. Interlocking diversification
      2. Different portmanteau analyses
      3. Distribution

Segmentation . . necessary because of horizontal discrepancies

      1. Composite realization
      2. Portmanteau realization

TYPES OF LINGUISTIC DISCREPANCY AND
CORRESPONDING METHODS OF ANALYSIS

TABLE 2

Sememic Analysis of the English Preposition

INTRODUCTION

The rest of the paper will deal mainly with the upper two strata. However, it is precisely the kind of analyses used on the lower strata that one uses on the upper strata. In other words, sememes are not simply "picked out of the blue" but are rigorously demanded by the structure of a given language. This concept is vital to an understanding of the analysis that is to follow.

An important example of linguistic analysis between the sememic and the lexemic strata is an analysis of the major English prepositions. The prepositions link all the important words of the language and not only relate their meaning but often determine it. One of the ways this analysis differs from other studies is by showing the difference between sememic analysis and other types of language analysis.

PROCEDURE

A large body of text was examined to determine the various sememic realizates of the prepositions. A sample of the text, taken from the *Golden Book Encyclopedia[3],* is given in the appendix. To illustrate the procedure, let us consider a particular preposition which may be called PREP. The entire corpus of data, that is, linguistic data, was scanned by a computer which printed out every sentence containing the preposition PREP. Using the method of differentiation, along with the two criteria interlocking diversification and distribution, a comparison of all these sentences was made by the experimenter to determine the sememic realizates of the lexon PREP.

ADDITIONAL NOTES ON THE METHOD OF ANALYSIS

Although the methods of analysis are technically described by the criteria interlocking diversification and distribution, there follows a brief description of the process in everyday terminology. To do an exact linguistic analysis (in this case, sememic analysis) of the English prepositions, one should examine every sentence (or utterance) which contains a preposition or any phrase which functions as a preposition (that is, which is substitutable for a preposition). One then is able to determine the various phrases in which one or more prepositions can occur. However, this task is obviously unlimited and so a heuristic must be introduced to make the problem feasible. For my analysis this heuristic was to use myself as an informant, that is, to use my knowledge, or perhaps view, of the language to solve the problem. I used myself as an informant in this sense: no new sememes were set up for a given preposition unless the text—to my mind—required it. For example, in the expression *it moves about on the branch of a tree,* using myself as an informant, I am able to substitute *around* for *about* and still have the

same meaning retained for the utterance. Continuing down the body of text, I came to the expression *shouting about the things they had to sell.* If I try the substitution of *around* for *about* here, the meaning of the sentence is no longer the same. Therefore, I set up here a new sememe which is realized by *about.* This new sememe is more in the sense of *concerning* and not in the sense of *around.* As I proceeded further, I came to the expression *about 600 B.C.* If I substitute *around* here, the meaning of the sentence remains unchanged. Moreover, if I substitute *approximately,* the meaning also remains unchanged. However, I cannot say *it moves approximately on the branch of a tree* and still keep the same meaning, nor can I say *shouting approximately the things they had to sell* and retain the same meaning. This, therefore, indicates two things: first it indicates a new sememe is realized by *about,* and secondly a new sememe is realized by *around.* So we can draw a chart (below) of interlocking diversification, labeling the sememes, for the lack of better names, ABOUT 1, ABOUT 2, and ABOUT 3. (We say technically that *about* is in interlocking diversification with *concerning* and *around).*

From this small example, one can get an idea of the
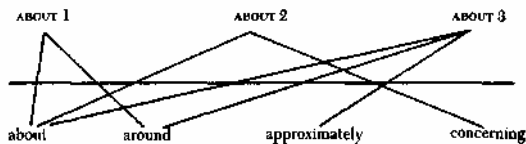
nature of such an analysis. Many times, of course, there isn't always a made-to-order word to substitute for the preposition (or for that matter a made-to-order phrase). To show the contrast, the substitution word must be a word which is equivalent in its function to a preposition and not simply any kind of long utterance. When this happens, the other criterion of distribution (perhaps better called transformation) must be used. For example, in the corpus that was analyzed there occurred the following expression: *fables* (were) *told by storytellers.* Here, it is difficult to find a substitution to distinguish this sememic realizate of *by* from the others. However, it happens in this case that one can make a transformation *storytellers told fables* and keep the same meaning. It turned out that this distribution criterion was enough to distinguish it from the other sememic realizates of *by.* For example, *the tree was planted by the house* does not transform to the *house planted the tree* and keep the same meaning. Hence there must be two different sememes here that are realized by *by.*

Finally, if an instance occurred where no substitution (or interlocking diversification) criterion or distribution criterion sufficed for a preposition, it was labeled as idiomatic usage, or as a new sememe because

it satisfied none of the criteria that the other examples satisfied. In idiomatic usages, the preposition is *part* of the realization of a sememe, instead of being a complete realization.

RESULTS

In addition to the sememic realizates of the prepositions the analysis also yielded classifications of semolexemes according to their constituent semons. This happened more than a few times when semolexemes were classified together because they occurred with a particular sememic realizate of a preposition. The analysis also sometimes yielded the criterion that determines which semolexemes can be associated with other semolexemes, or in layman's terms, which concepts can be associated with other concepts.

Since the main purpose of the procedure was to clarify the nature of sememic analysis and to show it to be a useful tool in problems of handling natural language, the analysis was not exhaustive. The partial analysis of a number of major English prepositions is presented here. With each sememic realizate associated with a preposition there will occur a certain class of semolexemes sharing one or more semons (unless the occurrence is an idiom). The semon may or may not be named, for as yet I have found no convenient system for naming each semon. The same principle applies to the sememic realizates of the prepositions.

EXPLANATION OF THE NOTATION FOR THE FOLLOWING ANALYSIS

The following is a brief explanation of the notation in the pages to follow. The sememes will be labeled with the name of the preposition in capital letters followed by 1, 2, 3, etc., to indicate a different sememe, for example, ABOUT 3 or BY 5. Sometimes, additional names or information will be given concerning the sememes in the space below the sememe name. For example, below ABOUT 1 we have written "this sememe is required by the semon shared by the following list."

The notation *about * around* for ABOUT 1 indicates that *around* was the substitution criterion used to determine the sememe ABOUT 1. Sometimes there will be no substitution criterion, as we have said earlier, and this fact will often be indicated by the name of the preposition in small letters followed by "*——." In these cases, the distributional criterion by which the sememe was determined will sometimes be given as in the case of IN 15. In some of the cases, criteria have not been given due to the lack of linguistic data. (There were about 3000 sentences to examine, all of medium length, i.e., about 15 words. Therefore, there was not enough data to substantiate completely some of the sememic categories set up, and these might be considered conjectures unless it is obvious that there is a new sememe. Moreover, some of the sememes which

are realized by a particular preposition will not be discovered by the analysis of such a limited amount of data).

Most of the examples where realization of a particular sememe occurs are taken from the text which was analyzed. Occasionally, I have made use of other examples, and these will be found below the dotted lines, as in BY 8.

The notation "Environment" will be found where word classes were listed rather than the full examples from the text. The notation /LIST——/ or /——LIST/ indicates that the list of words goes before or after the preposition being considered, respectively.

The interlocking diversification charts further explain the analysis. Above the line is the sememic stratum and the sememic realizates of the particular preposition under consideration; below the line is the lexemic stratum and the lexemic realizations of the sememes above the line. Of course, there may be more than are pictured but the essentials listed will help to clarify the analysis for the reader.

At the end of each list of sememes, the idiomatic usages are listed under "Idiomatic Usage" where, as mentioned above, the preposition is a part of a realization of a sememe.

ANALYSIS OF THE ENGLISH PREPOSITION

ABOUT 1: about * around

This sememe occurs with the semon shared by the following list:

move	
travel	
go	Environment
roam	
journey	
romp	/ LIST *about* /
strew	
batter	
scatter	

ABOUT 2: about * concerning

This sememe occurs with the semon shared by the following list:

shout	
hear	
myth	
sing	
talk	Environment
boast	
legend	
know	/ LIST *about* /
worry	
happy	
story	

learn
warn
study
curious
find out
agree
wrong
superstition
tell
read
quarrel
write
puzzle
fables
discoveries

ABOUT 3: about * approximately

This sememe occurs with the sememe shared by the following list:

QUANTIFIERS such as NUMBERS or NUMBERS of something

ready	
set	Environment
time	*/about* LIST/
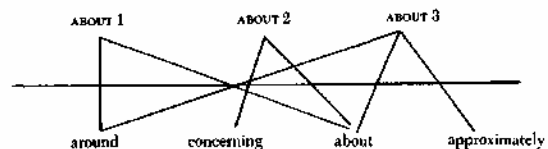
INTERLOCKING DIVERSIFICATION CHART FOR ABOUT

Idiomatic usage

1. to bring about
2. am about to

AT 1: at * ——

This is the most general sememic representate of *at* and can be distinguished by its contrast with the other representates.

at night
at day
at one meal
at the time
at the end
at the beginning
at midnight
at a place
at home
at the University
at the airport
at the control tower
at the mouth of the river
at the hospital
at the farmhouse

at the point
at the surface of the earth
at the top of a mountain
at the level of the sea
at the bottom of a pond

AT 2: at * on

work at it
- - - - - -
struggle at it
labor at it

AT 3: This sememe is determined by a distributional or transformation criterion.
increasing at a rapid rate
- - - - - - - - - - - - - - - -
running at a fast pace

Note: we can make a transformation:

rapid rate of increase
fast pace of running

This is sufficient to distinguish AT 3 from the other realizates of AT. For example, if we had an expression such as *stopping at the best hotel,* we cannot make a transformation to *best hotel of stop.*

AT 4: at * for

at the cost of
at the price of
at the rate of
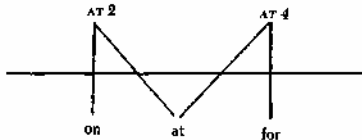
INTERLOCKING DIVERSIFICATION CHART FOR AT

Idiomatic usage

1. not at all
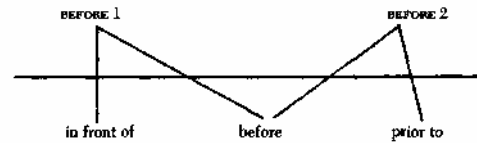2. at last
3. at least

BEFORE 1: before * in front of

stand before him
lay before him
fall before him

BEFORE 2: before * prior to

before Alfred's time
before 1800

just before the performance
before the middle of the nineteenth

INTERLOCKING DIVERSIFICATION CHART FOR BEFORE

BY 1: by * near
- - - - - - - - - - - - -
the house by the sea
the tree by the lake
the wastebasket by the desk

BY 2: by * during (the)

by night
by day

BY 3: by * ——

by the next New Year's day
by the middle of January
by (SOME EXACT YEAR)

BY 4: This sememe is also called AGENT; it expresses a relationship between ACTOR and ACTION.

(   ) told by storytellers
(   ) written by an author
fruits carried by ponies and camels
Africa explored by the people
Albania ruled by Turkey
apparatus made by alchemists
library destroyed by conquerors
Alhambra was built by the Moors
alphabets invented by the Semites
ambrosia brought by pigeons
animals cared for by their parents
base camps set up by the American explorer
diseases caused by germs
story told by his wife
Syracuse captured by the Romans
region was bought by the United States
bulletproof vests are worn by soldiers
guns pulled by horses

Here we have a distributional or transformational criterion: if we have A (Passive Verb) by B, we can transform to B (Active Verb) A. This is the active-passive transformation.

Thus we have the examples:

ponies carried fruit
people explored Africa
conquerors destroyed library
etc.

BY 5: by * with

This sememe is also called MEANS or INSTRUMENT
-- - --------------
  fastened together by pins
  tied together by rope

This sememe differs from BY 4 in the following sense. We can say *John tied the packages together by rope,* and so the agent is *John* and the MEANS is the rope. If we change this to the passive form, the language requires *with; the packages were tied together by John with a rope.*

BY 6: by * via

  travel by land
  by boat
  by train
  by plane
  by bus
  by air
  by (MEANS OF TRANSPORTATION)

BY 7: by *——

This sememe is also called DISTRIBUTIONAL MEASURE

  by the hundreds
  by the thousands
  little by little
  -- - -------
  step by step
  bit by bit

BY 8: by * (multiplied by)
-- - ---------
  two by four
  five by five
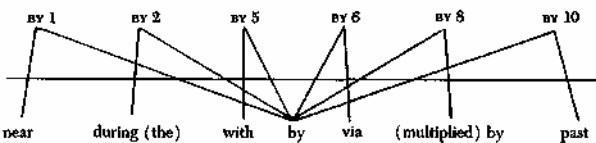 NUMBER by NUMBER

BY 9: by *——
-- - ----------
  differs by quite a lot
  better by far

BY 10: by * past
  roar by
  go by
  -- - -------
  sail by
  flow by

INTERLOCKING DIVERSIFICATION CHART FOR BY

FOR 1: for *——

This sememe is also called GOAL. It occurs with the sememe shared by the following list:

hunt	Environment
search	
fish	/ LIST *for* /
-- - ----	
look	

FOR 2: for * as

  wearing a saucepan for a hat
  using a cave for a shelter
  for example; for instance

FOR 3: for * for the purpose of

This sememe is also called PURPOSE.

  system for bringing water
  lamps for killing germs
  wheel for steering
  points nose of the airplane into the wind for
    landing
  mills for making cotton
  factories for making (   )
  roads for cars
  houses for planes
  land for cotton
  home for water plants
  grave for a dead body
  books for girls
  large areas for ranches
  apples for eating
  apples for cooking

FOR 4: for * assigned to

  name
  sign
  word
  letter
  -- - ----
  nickname
  symbol
  password

As in the following examples:

  name for an animal
  signs for various sounds
  Mont Blanc is the French name for White Moun-
    tain
  nickname for Arizona
  word for amber

FOR 5: for * for the extension of

  thousands of years
  generations to come

the rest of the journey   Environment
more than a summer
a hundred years   */for* LIST/
many centuries
many minutes
twelve seconds
several hundred miles
long stretches
- - - - - - - - - -
several yards
many feet

FOR 6: for * in exchange for

     United States bought Alaska for $7,200,000
     sells for a high price
     paid money for it
     trade seeds for food
     offer a million dollars for the backbone of a
        mosquito
     for a dollar or so, one can go hunting

FOR 7: for * because of

     famous for its alligators
     whales are killed for their oil
     famous for its vineyards
     for this reason

FOR 8: for *——

This sememe occurs with the following relational-type expressions:

     hard for him
     easy for him
     - - - - - - - - - -
     difficult for him
     troublesome for him

FOR 9: for * used for

This is a relationship sememe which occurs with the sememe shared by the following MATERIAL—PRODUCT pairs:

     cotton—clothes
     tobacco—cigarettes
     sisal—rope
     palm—soap
     pulp wood—lumber
     iron—bridges
     silver—knives
     wood—arrows
     stone—building
     chain mail—armor
     silver—fillings

Notice here we also have a transformational criterion; where we have A for B, we may say B made of A. This is completely sufficient to distinguish FOR 9 from the rest of the realizates.

FOR 10: for * ——
     for the first time
     for the second time
     - - - - - - - - - - - -
   - - - - - - - - - - - -
     for the last time

FOR 11: for * corresponding to

     an airline may have more than fifty men on the
        ground for every plane it flies
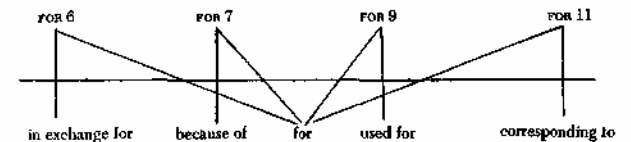   different ages for different kinds of plants

FOR 12: for * despite

     for all its great size, the Amazon is a lazy, sluggish river

FOR 13: for * for the benefit of

This is also called the BENEFACTIVE sememe.

     singing for the other workers
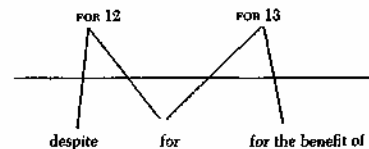   carved whistles for the people

FOR 14: for * ——

     they are old even for mountains
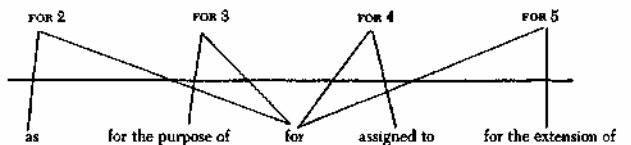
FIRST INTERLOCKING DIVERSIFICATION CHART FOR FOR

SECOND INTERLOCKING DIVERSIFICATION CHART FOR FOR

THIRD INTERLOCKING DIVERSIFICATION CHART FOR FOR

Idiomatic Usage
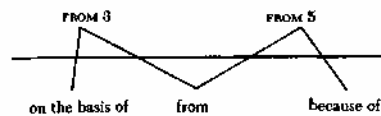
     1. cost too little for me to bother

FROM 1: from * (out of)

This sememe is called SOURCE.

     this name is made from two words
     it gets water from its food

get ivory from the coast
people from other lands
come from
 the forests
cotton from Africa
men from the University
gases from burning fuel
oxygen from the air
wind comes from different directions
power from dams
products manufactured from steel
steel from Birmingham
"Albino" comes from a Latin word
color comes from its blood
grain alcohol made from corn
they hatch from eggs
from there on
alphabet came from the Greeks
aluminum comes from mineral bauxite
begged for water from another ship
name comes from his
crossed over from Europe
sticking out from its forehead
get milk from the mother

FROM 2: from *——

SEPARATION is another name for this sememe

55 miles from the Soviet Union
stretched westward from Alaska
straight south from North America
keeps water away from bear's skin
separates from them
hide from enemies

FROM 3: from * on the basis of

as anyone can guess from their sharp teeth
from its fur one can tell that it is a mammal

Here, of course we have a list of words not too unlike some of the other SENSE categories; however, it is more restricted:

guess
tell
surmise
see
figure

One might call this the INFERENCE sememe.

FROM 4: from *——
one place may differ from another
they are a different kind from (   )
- - - - - - - - - - - - - - - - - -
animals in Africa are usually different from those in America

FROM 5: from * because of * (of)

die from pneumonia
die from a fatal wound

INTERLOCKING DIVERSIFICATION CHART FOR FROM

IN 1: in * ——

in some ways
in like manner

That is, words like *way* and *manner* fall into this special category.

IN 2: in * inside of

This sememe is called LOCATION.

curled up in a hole
in the ground
in Japan
lemons have acid in them
in a forest
pigment in skin
picture in the book

IN 3: This sememe is called STATE

rose in bloom
Alfred was in hiding
twisting in pain
- - - - - - - - - - - - -
the body was in state
in suspense
in trouble

IN 4: in * during

in the War
in one battle
in the days of the cavemen
in the early days
in the middle ages
in a single day
in the last part of the war
in the long history of (   )
in summer

IN 5: in * for

used a bead in counting
are important in playing
used in manufacturing
help in traveling
planes use runways in landing
used in building it
used it in hunting

WHITE

IN 6: in * at

    in the end
    in the beginning

IN 7: in * into

    dipped in the river
    eat holes in the cloth
    - - - - - - - - - - - -
    jump in bed
    fall in the pit
    throw in the arena

IN 8: in * ——

    interested in

IN 9: in * after a time of

    in about ten weeks
    in a few weeks
    in three or four months

IN 10: in * with respect to

    in size, Alabama ranks twenty-ninth
    in population, Alabama ranks seventeenth
    light in weight
    match it in size
    change in looks

IN 11: A transformation determines this sememe, for where we have A in B we may also say B of A.

    ships in the fleet—fleet of ships
    -
    people in the group—group of people
    cattle in the herd—herd of cattle
    fish in the school—school of fish

This classifies (or quantifies) groups of objects by the terminology used for a group of them.
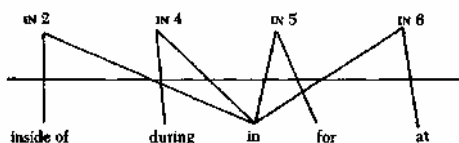
FIRST INTERLOCKING DIVERSIFICATION CHART FOR IN
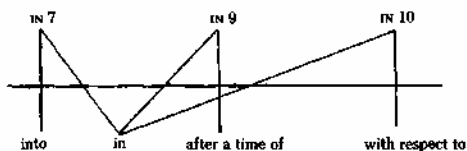
SECOND INTERLOCKING DIVERSIFICATION CHART FOR IN

Idiomatic Usage

1. in fact

OF l: of *——
This sememe is also called the PARTITIVE sememe.

    ends of an accordion
    part of an accordion
    keyboard of an accordion
    back of our heels
    branch of a tree
    much of it
    most of it
    parts of Africa
    edge of the plateau
    half of the continent
    the rest of the world
    all of Africa
    rest of the journey
    part of the atmosphere
    every square inch of this page
    one fifth of the air
    door of the building
    wings of a giant bird
    top of the wing
    cockpit of an airplane
    walls of Troy
    countries of Asia
    plains of Pakistan
    surface of the earth

OF 2: of * made of

    strips of metal
    coat of fur
    land of stone
    rivers of ice called glaciers

This sememe contrasts strongly with OF 1, for we cannot say *ends made of an accordion* and retain the same meaning.

OF 3: of * from or (taken from)

    one of the heroes
    few of them
    three of the strong acids
    one of our biggest businesses
    largest of the seven continents
    most of them
    several of Alaska's towns

There are some interesting features which distinguish this sememic realizate from the others. First of all, in OF 1 where we have an expression like *A of B,* B can either be a singular or plural noun. However, in OF 3 in an expression *A of B,* B is always plural.

Secondly, in most of the cases of OF 3, the *of* can be removed or neutralized and we can have such expressions as *one hero, three acids, one business, several towns,* etc. However, in OF 1 and OF 2 we cannot perform this

kind of transformation, else we get such expressions as *part accordion* or *coat fur.* Interestingly enough, though, we do have for OF 2, transformations to such expressions as *fur coat* or *metal strips,* i.e., A of B is transformed to B A.

OF 4: of * in

> heroes of Homer's Iliad
> merchants of Ancient Egypt
> people of Afghanistan
> natives of Africa
> mountains of Switzerland
> trees of California

OF 5: of * about

> the story of the Trojan War
> think of it
> teacher of natural history
> history of airplanes

This sememe is, of course, the same as ABOUT 2.

OF 6: This sememe is called POSSESSION. We have a well-known transformation to determine this, namely that of *A of B* to *B's A.*

> - - - - - - - - - - - - - - - -
> the hat of the King of England
> the jewels of the Duke of Rochester

This contrasts with:

OF 7: This sememe is called RELATIONSHIP.

> gods of the sea
> the king of Troy
> the governor of Louisiana
> the son of Thetis
> the father of Achilles

In the first example—that is, OF 6—the *of* is in the sense of *belonging to* or *in possession of.* Thus we can say:

> the hat belongs to or is the King of England's
> the jewels belong to or are the Duke of Rochester's

but we cannot say so readily:

> the gods belong to or are the sea's
> the king belongs to or is Troy's
> the father belongs to or is Achilles's

and still retain the same meaning.

OF 8: of * given to

> name of (anything)

OF 9: of *——

This sememe is determined by a transformation criterion:

X of (VERB LIST) + ing transforms to
X to (VERB LIST) Y

> ways of finding out—ways to find out
> ways of doing the job—ways to do the job
> idea of copying birds—idea to copy birds

OF 10: of *——

This sememe is a sort of MEASURE sememe. That is, in the expression *A of B, A* is the quantizer and *B* the thing quantized

> flock of sheep
> group of children
> quantities of cotton
> box of breakfast cereal
> cartons of milk
> - - - - - - - - - - - - - - - -
> set of silverware

It is worth noting, then, that this sememe determines a special set of expressions—group expressions, such as flock, herd, school, group, etc.

OF 11: of *——

> different kind of chemical
> every kind of place
> - - - - - - - - - - - - - - - -
> type of ( )

OF 12: of * ——

This sememe is called PROPERTY

> pressure of the water
>
> - - - - - - - - - - - - - - - -
> color of the chair
> warmth of the body

Where we have *A of B,* we may say that A is a property of B.

OF 13: of * ——

> full of advertising
> it is full of air

OF 14: of * ——

This sememe can be determined transformationally:

> *A of B* goes over to *B is an A*
> strait of Gibraltar
> city of New York
> isthmus of Suez
> state of Texas

OF 15: of * containing

This sememe is admittedly similar to OF 10, but is clearly not the same

> city of 1000 people
> town of 3500 people

We are not using the city to measure the number of people, which would be the case if this were an example of OF 10, but rather we are talking about a *city* which contains 1000 people.

OF 16: of * from (a different *from* from OF 3)

In fact, we can pretty much say that the *from* which realizes OF 3 is that which realizes FROM 1, whereas this *from* is that which realizes FROM 2.

> Africa is south of Europe
> south of the Sahara
> north of Mobile
> - - - - - - - - - - - - - - -
> west of Alaska
> east of the Mississippi

OF 17: of * because of * from

> die of old age
> die of pneumonia

OF 18: of * by

This sememe is also called AGENT.
> - - - - - - - - - - - - - - -
> growling of the lion
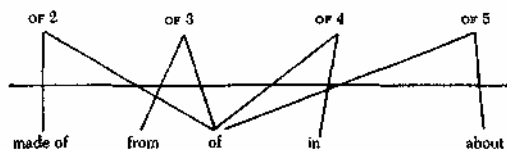> shooting of the hunter
> thinking of the student

OF 19: of *——

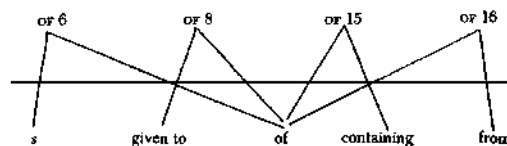A transformation determines this sememe:

> (VERB LIST) + ing of the Y transforms to (VERB LIST) the Y

> making of the stew—make the stew
> shooting of the hunters—shoot the hunters

This sememe is also called GOAL.

FIRST INTERLOCKING DIVERSIFICATION CHART FOR OF

SECOND INTERLOCKING DIVERSIFICATION CHART FOR OF

Idiomatic Usage

1. of course

ON 1: on * ——

This sememe is also called DEPENDENCY. It is a relational sememe and thereby indicates sememic properties of the expressions it relates

> aardvarks live on termites
> pilots depend on the instruments of their planes
> most counted on magic
> live on food
> - - - - - - - - - - - - - - -
> rely on ( )

ON 2: on * ——

This sememe occurs with a special class of verbs.

> an aardvark may stand up on its hind legs
> gorillas walk on all fours

ON 3: on *——

This sememe is also called LOCATION and is used to answer the question "Where?"

> move up and down on wires
> keys on the keyboard
> on the branch of a tree
> ports on the coast
> on land
> - - - - - - - - - - - - - - -
> on the street
> on the table

ON 4: on *——

> play on (a musical instrument)

ON 5: on *——

> hear on the radio
> see on television
> - - - - - - - - - - - - - - -
> talk on the phone

ON 6: on * for

> money spent on advertising
> time wasted on advertising

ON 7: on * against

> pushing on the right pedal
> marched on Rome
> force demands on the English people

ON 8: on *——

> she went on an errand
> - - - - - - - - - - - - - - -
> traveled on a mission

ON 9: on * at

   work on it

ON 10: Idiomatic usage with the following expressions:
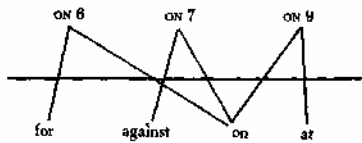
   the heat was hard on his eskimo dogs

   - - - - - - - - - - - - - - - -

   rough on him

ON 11: on *———

   play jokes on other people
   play tricks on other people

INTERLOCKING DIVERSIFICATION CHART FOR ON

Idiomatic Usage

   1. on the other hand
   2. set it on fire

TO 1: to *———

   farthest to the right
   path to one side
   lands to the east
   far to the north

TO 2: to * (motion) towards

   come to the edge
   went to them
   bring water to dry fields
   come up to the falls by boat
   turn to the right or left
   swings to the right
   forces the tail to the left
   plane banks to the right
   sent Hans to school
   moved on to giant planes
   bring to the region
   made his way to the big marshes
   journey to the Atlantic Ocean

   Note: TO 1 and TO 2 differ in the same respect as in
   and *into.*

TO 3: to *———

   to everyone's surprise

   - - - - - - - - - - - - - - -

   to my amazement
   to my wonderment

TO 4: to * ———

   This sememe occurs with the sememe shared by the
   following list which might be called JOIN.

      fasten
      connect
      tie
      anchor
      graft
      - - - - - - - - - - - - - - -
      confine
      belong

   Examples:

      fasten to the keys
      connected to the car's front wheels
      closely tied to the Soviet Union
      algae have holdfasts that anchor them to rocks
      graft them to the roots
      tying a purse to a string

TO 5: to *———

   This sememe occurs with those verbs which allow
   INDIRECT OBJECT; hence we get a transformation:

   (VERB LIST) X to Y transforms to
   (VERB LIST) YX

      sell it to more people
      give jobs to many
      did damage to it
      world owes a great deal to the people
      gives his name to half the world
      pay attention to him

TO 6: to *———

   This sememe is characterized by a special group of
   adjectives:

      attractive to Europeans
      agriculture important to Alabama
      known to everyone
      well known to the settlers

TO 7: to * with

   talk to the pilots

TO 8: to * in honor of

   a monument to freedom

   - - - - - - - - - - - - - - -

   a toast to the married couple

TO 9: to * with (in a different sense from TO 7)

   compared to the gold

   - - - - - - - - - - - - - - -

   in contrast to the fine job done by the painter

TO 10: to * into

>change metal to gold
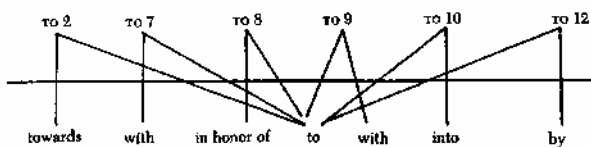>it turned to amber

TO 11: to *——

>sticky to the touch
>
>-- - - - - - - - - - - - - - -
>silent to the ear

TO 12: to * (by)

>close to the earth
>near to the equator

TO 13: to *——

>This is a sememe which occurs with the sememe shared by the following list:
>
>right
>claim
>
>Example:
>
>a right to the throne of England

INTERLOCKING DIVERSIFICATION CHART FOR TO

WITH 1: with * by means of

>This sememe is also called MEANS.
>
>an aardvark can rip open a termite nest with its strong claws
>an aardvark can lick termites with its tongue
>the boy is counting with an abacus
>it is weakened with water
>paper colored with dye
>connected to the door with hinges
>breathes with gills
>wings are fastened on with wax
>pushes the water with a paddle
>runways marked with lights
>plowing done with oxen and plows
>written with 26 letters
>written with signs
>built a house with lumber
>ditch walled with stone

WITH 2: with * together (along) with

>This sememe is also called ACCOMPANIMENT.
>
>a solo instrument not played with other instruments
>to work with him

joined with the southern states
joined with other materials *(note that this is ambiguous)*
set out for the pole with four men
could not take the tree home with him
the Bible he always carried with him
he camped with the Indians
study with Plato

WITH 3: with * (among)

>popular with sailors
>
>-- - - - - - - - - - - - - - -
>a favorite with teenagers

WITH 4: with * to

>joined with other materials *(note that this is ambiguous)*
>link rich mines with the port
>link Alaska with the other states
>-- - - - - - - - - - - - - - - - - - -
>join Philadelphia with New York

WITH 5: with * having

>animals with backbones
>animals with eardrums
>an aquarium with salt-water animals
>-- - - - - - - - - - - - - - -
>a man with courage

WITH 6: with *——

>This sememe is also called MANNER and is a part of the answer to the question "How?"
>
>push with force
>rush out with a terrible force
>burns with a hot blue flame

WITH 7: with * on

>experiment with gliders
>he experimented with mirrors *(note the ambiguity with WITH 1)*

WITH 8: with * to

>talk with pilots
>friendly with Latin American countries

WITH 9: with * ------

>filled with milk
>filled with bad odors

Compare this with OF 13.

WITH 10: with * against

>fought battle with the Danes *(note the ambiguity with WITH 2)*
>fought with the South

WITH 11: with * for * among

    one baby at a time is the rule with elephants

WITH 12: with * ——

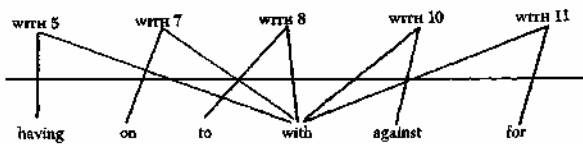   they furnish us with drugs
    they supply us with drugs

WITH 13: with * ——

    the use of antibiotics began with the operation in
      Boston
    the use of antibiotics ended with the discovery of
      .....

WITH 14: with *-------

  This sememe is also called EXCHANGE

    replace gold with silver

    - - - - - - - - - - - - - - - -

    change places with him
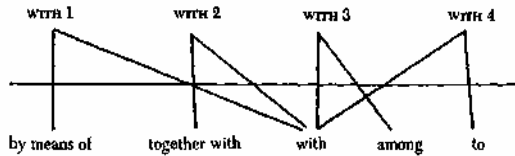
FIRST INTERLOCKING DIVERSIFICATION
CHART FOR WITH

SECOND INTERLOCKING DIVERSIFICATION
CHART FOR WITH

Idiomatic Usage

    1. she shakes hands with (him)
    2. in love with her

DISCUSSION OF RESULTS

From such a detailed sememic analysis of the English preposition, one can get a fairly good grasp of what language stratification is—at least as far as the lexemes and the sememes are concerned. However, we state again that the rules we used to obtain the sememes are exactly analogous to those which were used to determine the "emes" of the lower strata. In other words the analysis is not *ad hoc* but is in keeping with the rest of the theory of stratificational linguistics.

The best means of identifying the presence of an interlocking diversification is by noting the presence of ambiguity. We group together some of the more important ones, or rather more obvious ones. We will state the ambiguous phrase and then the two sememes (there may indeed be more than two for some ambiguities) which are realized by the preposition in the ambiguous phrase.

I feel about the room: ABOUT 1 and ABOUT 2
Come by night: BY 2 and BY 3
Go by a bus: BY 6 and BY 10
Killed by a poison arrow: BY 4 and BY 5
One can tell from the chair that the room is beautiful:
   FROM 1 and FROM 3
In a pickle: IN 2 and IN 3
Throw the ball in the arena: IN 2 and IN 7
Shooting of the hunter: OF 18 and OF 19
The thinking of the student: OF 18 and OF 5
Fleas live on dogs: ON 1 and ON 3
The child is playing on the piano: ON 3 and ON 4
I hear the clock on the radio: ON 3 and ON 5
The money was spent on the beach: ON 3 and ON 6
Pushing on the floor: ON 3 and ON 7
To work on the table: ON 3 and ON 9
The path to one side of the house: TO 1 and TO 2
I traded a big house for the company: FOR 6 and FOR 13
I see the man with a telescope: WITH 1 and WITH 5
To fight with the Danes: WITH 2 and WITH 10
Joined with other materials: WITH 2 and WITH 4
He experimented with mirrors: WITH 1 and WITH 7
He experimented with his friends: WITH 2 and WITH 7
A dog with a bone: WITH 2 and WITH 5

Another interesting phenomenon is the amount of equality among the sememic realizates of the various prepositions. Since I was not able to do an analysis of all the prepositions and preposition-like phrases, the results are not as complete as they could be. However, I list here some of the equalities:

AT 2 = ON 9	OF 4 = IN 2
BY 5 = WITH 1	OF 16 = FROM 2
FROM 5 = OF 17	ON 6 = FOR 3
AT 1 = IN 6	TO 12 = BY 1
IN 5 = FOR 3	TO 7 = WITH 8
OF 5 = ABOUT 2	WITH 4 = TO 4
OF 3 = FROM 1	OF 18 = BY 4

The "=" sign, of course, means that the sememes are the same. However, the actual realization of the sememe is often conditioned by the environment. For example, in the expression *I am worried about it,* one cannot substitute *of* for *about;* thus even though OF 5 and ABOUT 2 are the same, one cannot always interchange *of* and *about* in instances of their realizations. The environment is said to condition which realization one uses. This is the same principle as in the case between the morphemic and graphemic strata where one has the morpheme [M]/s/ realized by the graphemes [G]/s/ and [G]/es/. Which realization one uses is completely determined by the environment, so that after

[G]/x/, as in /tax/, one uses [G]/es/ whereas after [G]/k/ as in /book/, one uses [M]/s/.

The lists that accompanied some of the sememic realizates of particular prepositions, "share a common semon," as mentioned earlier. This means that each word in the list shares an element common to all the words in the list which partially determines its meaning; if we view the meaning of each of the words in the list as having a structure, then the element or semon which they share is one of the basic elements of the structure of all the words. For example, in ABOUT 2 analysis, we find a list of words: shout, hear, myth, talk, sing, etc. All of these words occur with *about* in a special sense. What do these words have in common? For one thing, all of them have something to do with the head senses, either directly, as in hear, see, etc., or indirectly, as in myth, legend, story. Hence, we can make a hypothesis that the particular semon here has something to do with "head sense", and that it is this semon which requires and is required by the special sememic realizate ABOUT 2.

Consider another example: in TO 4, we have a list of words such as: fasten, connect, tie, anchor, graft, tie, belong, etc. We find that what these words have in common is something like "joining", and so a basic element in the structure of these words is that of joining, and it is this semon that requires and is required by the sememic realizate TO 4.

There are other types of lists which might be called relational lists. Instead of lists of single words, these are lists of pairs which have a special relationship to one another as expressed by the particular sememic realizate of the preposition that relates them. This then gives some additional evidence for the structure of meaning in words.

An excellent example of this is found in FOR 9 where there are such pairs as cotton—clothes, tobacco—cigarettes, sisal—rope, etc. Here the *for* expresses the fact that something is a material of something else which, in general, is the product. So it is this material—product relationship (or perhaps "semon") which requires and is required by the sememic realizate: FOR 9.

Another example occurs in OF 1. Here there are such pairs as: keyboard—accordion, branch—tree, door—building, and the relationship then is that of part—

whole. And it is this "semon" that requires and is required by OF 1.

A third example is found in IN 11 and OF 10, such pairs as ships—fleet, people—group, cattle—herd, fish—school, etc. The obvious relationship between these pairs determines and is completely determined by the particular sememic realizate.

The basic result of this is, in fact, a partial sememic analysis of words of the language other than prepositions. In particular we get some interesting results. For example, we get criteria which determine when some concepts may be associated with other concepts. We also get lists of verbs and nouns which are determined by their particular function, such as those associated with BY 4. Most important, though, we get words classified by what structural elements of meaning they have in common. This particular result is excellent for providing portmanteau analyses of particular semolexemes, such as was done earlier for [S]/RAM/ which divided into /male/ + the bundle of sememes associated with the semolexeme [S]/SHEEP/. Consider for example, the portmanteau table:

(people) person	group
fish	school
(cattle) cow	herd
sheep	flock

Having set up such a table, one may do one of two things: either he can extend the table, using himself as an informant, by adding such things as

| wolf | pack |
| buffalo | herd |

or he can extend the table the other way setting up new categories such as the following:

(people) person	man	woman	group
fish	fish	fish	school
(cattle) cow	bull	cow	herd
sheep	ram	ewe	flock

and so forth.

In general, then, the sememic analysis of the prepositions turns out to be an excellent jumping off point for a sememic analysis of the rest of the language.

References

1. Lamb, S. M. *Outline of Stratificational Grammar.* Berkeley: University of California, 1962.
2. Lamb, S. M. The Sememic Approach to Structural Semantics. In *Transcultural Studies in Cognition,* Roy d'Andrade and A. Kimball Romney, eds. In press.
3. *Golden Book Encyclopedia,* Vol. 1. New York: Golden Press, Fourth Printing, 1960.