THEORETICAL PROBLEMS IN
FULL-TEXT INFORMATION RETRIEVAL

Working Paper No. 365

David C. Blair
The University of Michigan

The design and implementation of large-scale, computerized, document retrieval systems has been an activity of increasing recent importance. The amount of information in both the public and private sectors has been growing at a dramatic rate. Resnikolf has shown that the volume of information for any country increases at the same rate as the increase in that country's Gross National Product. This growth of information has been facilitated by the widespread use of computerized word processing systems which have sped up the transition from drafts of documents to final copy. Word processing systems have also exacerbated the information retrieval problem in another way. A growing number of word-processed documents never become "hard-copy" but are stored in computer data bases and are circulated via electronic mail systems. This creates a problem because documents stored on-line in machine readable form frequently have fewer access points than documents stored in more conventional information retrieval systems. [Blair]

This need for sophisticated computerized document retrieval systems coupled with the increasing amount of machine-readable text being stored in data bases have made the use of full text retrieval systems, such as STAIRS, LEXIS or GESCAN, an appealing solution to the document retrieval problem.

The Appeal of Full Text Retrieval

The retrieval of texts or documents by subject content occupies a special place in the province of information retrieval for, unlike data retrieval the richness and flexibility of natural language have a significant influence on the conduct of an inquirer's search. The inquirer must describe his information need using subject descriptors actually assigned to documents on the data base he is searching, while the indexer must choose appropriate subject terms to describe the "information content" of the documents to be

included in that data base (see Figure 1). But there are no clear and pre-
cise rules which an indexer can follow to select the "appropriate subject
terms" describing a particular document. This means that even trained in-
dexers may be inconsistent in the selection of subject terms to describe
documents. Experimental studies on indexing have confirmed this by demon-
strating that different indexers will generally index the same document
differently. [Zunde and Dexter] (Even the same individual will not always
select the same index terms if asked at a later time to index a document
he has previously indexed.) Such problems with the manual assignment of
subject descriptors to documents make computerized, full-text document
retrieval appealing. By entering the entire, or the most significant part
of, the text of a document onto the data base one is freed, it is argued,
from the inherent evils of manually creating document records which reflect
the subject content of a particular document. The evils avoided include:
the construction of an indexing vocabulary, the training of indexers, the
excessive time needed to scan/read the documents and assign context and sub-
ject terms to documents. Such economies are appealing, but for full-text
retrieval to be worthwhile it must also provide satisfactory levels of
retrieval effectiveness.

Full-Text Retrieval Problems: Empirical Evidence

In a detailed study of the retrieval effectiveness a full-text retrieval
system [Blair and Maron] · found that the Recall levels of the system were
unacceptably low (less than 20% when 75% was considered the minimum required
level). While this result is quite dramatic, several explanations for such
poor results must be considered.

1. The inquirers confused Recall and Precision levels in their

```
┌─────────────────┐                           ┌─────────────────┐
│    INCOMING     │                           │    INQUIRER     │
│    DOCUMENTS    │                           │  IN SEARCH OF   │
│                 │                           │   INFORMATION   │
└─────────────────┘                           └─────────────────┘
         │                                             │
         ▼                                             ▼
┌─────────────────┐     ┌─────────────┐       ┌─────────────────┐
│ IDENTIFICATION  │ ──> │  VOCABULARY │ <──   │      QUERY      │
│       OF        │     │  THESAURUS  │       │   FORMULATION   │
│    CONTENT      │ <── │             │ ──>   │                 │
└─────────────────┘     └─────────────┘       └─────────────────┘
         │                                             │
         ▼                                             ▼
┌─────────────┐      ┌─────────────┐          ┌─────────────┐
│    INDEX    │      │   SEARCH    │          │   FORMAL    │
│   RECORD    │ ──>  │    AND      │ <──      │    QUERY    │
│             │      │  RETRIEVAL  │          │             │
└─────────────┘      │   SYSTEM    │          └─────────────┘
                     └─────────────┘
                           │
              OUTPUT  TO  INQUIRER
                "RELEVANT" ITEMS
```
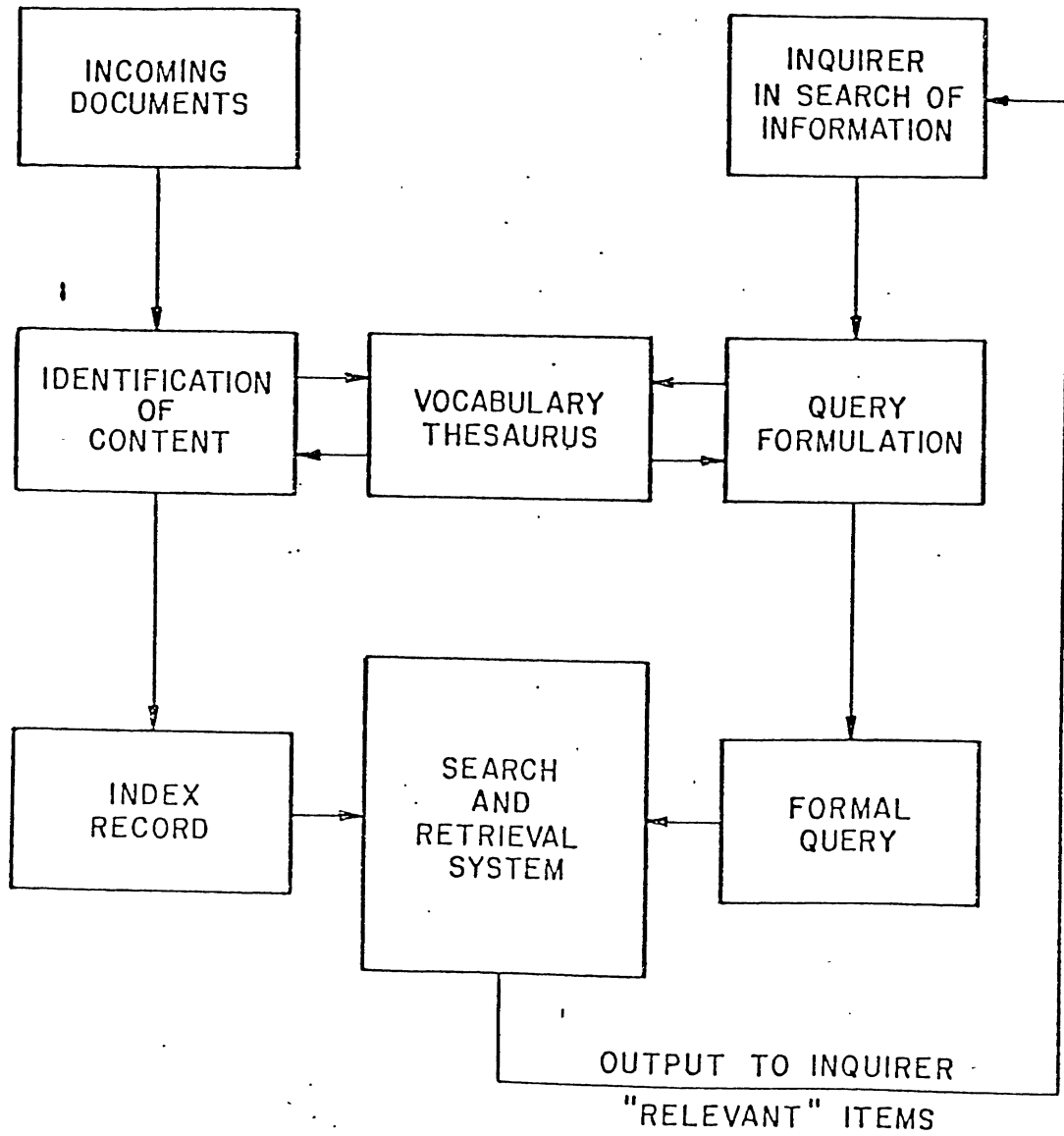
Figure 1

searches. Since the mean Precision level was over 80%, it is possible that the inquirers were confusing the percentage of relevant documents retrieved (Recall) with the percentage of retrieved documents which were relevant (Precision). This is a plausible explanation for the low Recall levels, but an examination of the experimental data reveals that the explanation must lie somewhere else. If the inquirers were confusing Precision with Recall, then they would not request further searching to be done if the initial retrieved set of documents had a Precision level higher than the desired Recall levels (.75). But several of the information requests for which further searching was requested had Precision levels of over .75 (up to 1.0). Other information requests began by retrieving sets of documents with high levels of Precision, but, as further searching was done, the levels of Precision dropped dramatically. Nevertheless, for each of these information requests, the inquirer eventually expressed satisfaction with the search (in spite of the degrading Precision levels). Finally, if we look at the mean Precision values for those information requests where no further searching was requested, we find a value of .84. Now, while this value is higher than the mean Precision level for all requests, the difference in these two means is not statistically significant at the .05 level. Thus, it seems that the reason for the inquirers' poor Recall estimation is more complex than just a confusion of Recall and Precision.

2. Variability in the searching abilities of the inquirers. It is possible that one of the inquirers with very poor searching skills brought the overall Recall average down, and that the more skillful inquirer had good results. As in the case with the first objection (supra.), an examination of the experimental data shows that this hypothesis cannot be confirmed.

	Recall	Precision
Inquirer 1	22.7%	76.0%
Inquirer 2	18.0%	81.4%

While there does seem to be some difference between the results for each inquirer, the variance is not statistically significant at the .05 level. Although this is a very limited test, we can conclude that at least for this experiment the results were independent of the particular inquirer involved.

As with other large computerized retrieval systems, the inquirers did not actually use the system but submitted their requests for documents to searchers who acted as intermediaries and did the actual searching for the inquirers. This procedure raised one other objection to this experiment. Consider the following argument: Because STAIRS is a high speed, on-line, interactive system, the searcher at the terminal can quickly and effectively evaluate the output of STAIRS during the query modification process. Therefore, the retrieval effectiveness can be significantly improved if the person who originated the information request was himself doing the searching at the terminal. This means that if an inquirer worked directly on the query formulation and query modification at the STAIRS terminal, rather than use an intermediary searcher, the retrieval effectiveness would be improved.

We tested this conjecture in order to see if, in fact, we could find a significant difference in values of Recall when comparing the retrieval effectiveness of the inquirer and the searcher on the same information request. We selected (at random) five information requests for which the searches had already been completed by the searcher, retrieved sets had been evaluated by the inquirer, and values of Recall had been computed. (Neither the inquirer who made the relevance judgments of retrieved sets nor the searcher

knew the Recall figures for these requests.) We invited the inquirer to use
STAIRS directly to access the data base, and we gave him copies of his
original information requests. He "translated" these information requests
into formal queries, evaluated the text displayed on the video screen,
modified the queries as he saw fit, and decided when to finally terminate
the search. We knew which documents he had previously judged relevant, and
we had previously estimated (for each of the five information requests) the
minimum number of relevant documents in the entire file. Therefore, we
were able to compute for the inquirer (as we had already done for his
searcher) the values of Recall. Thus, if it were true that STAIRS would
give better results when the inquirers themselves work at the terminal, then
the values of Recall should be significantly higher than the values of Re-
call when the searchers did the retrieval. The results were:

Request Number	Recall; Searcher	Recall; Inquirer
1	7.2%	6.6%
2	19.4%	10.3%
3	4.2%	26.4%
4	4.1%	7.4%
5	18.9%	25.3%
Mean	10.7% (s.d.=7.65)	15.2% (s.d.=9.83

Although there is a marked improvement in the inquirer's Recall for informa-
tion requests 3, 4 and 5, and in the average Recall for all 5 information
requests, the improvement is not statistically significant at the .05 level
($z=-0.81$). Hence, we cannot reject the hypothesis that both the lawyer and
the searcher get the same results for Recall.

3. The experiment itself was an artifact. It is possible that although
the experimental procedure and its conclusions are unassailable, the experi-
ment itself recorded an anomalous situation, and that given a different

group of inquirers, a different data base, or a different retrieval situation, the Recall levels might have been satisfactory. This is a much more difficult objection to answer than the first two. But it can be answered, not by looking at the experimental data, but by looking at the fundamental assumption on which full text retrieval is based. Full-text retrieval should work, it is argued, because it is a relatively simple matter for inquirers to predict the words and phrases used by relevant documents and not used by irrelevant documents. This would insure that mostly relevant documents would be retrieved. Such an assumption about full-text retrieval is, in reality, a statement about how natural language works. It states that the authors of relevant documents will write about the subject of interest to the inquirer in a uniform and predictable way using words and phrases which are predictably different from the words and phrases used to write about subjects not of interest to the inquirer. To determine whether or not this assumption holds we must first consider how language works—how authors use language to convey certain ideas or subjects to their readers.

Linguistics

The study of signs and sign-systems has become an increasingly vigorous academic industry, dramatically expanding the scope of research which had been limited previously to linguistics. Linguistics is now perceived as one part of a large, and rather loose, field of investigation commonly subsumed under the rubric of Semiotics or Semiology. A careful analysis of Semiotics reveals that it is not a rigorous or established discipline either in its description or methods of analysis. At best, Semiotics is a provocative, but perhaps pre-paradigm field which, as it now stands, is capable of

offering suggestions and adumbrations, but few hard "facts" or reliable methods. Nevertheless, the field of Semiotics contains some of the most important current work in understanding natural language.

Semiotics

Semiotics ("Semiology," "The Semiotic," "Theory or Signs," etc.) is a wide-ranging field concerned with the many aspects and uses of signs. It is an old field whose roots have been traced back as far as Heraclitus [Luigi] but whose formal beginnings find their origin in Locke and their early development in Peirce and Saussure. It is not my purpose to offer a discussion of the development and scope of Semiotics. This has been done well in several other sources [e.g., Sebeok, Guiraud, or Jakobson]. The following introduction to Semiotics is meant to give the reader a basic familiarity with its theory.

The Sign

Semiotics concerns anything that can be taken as a sign, and a sign is anything which can be taken as a substitute for something else. "This something else does not necessarily have to exist or to actually be somewhere at the moment in which a sign stands for it. Thus semiotics is in principle the discipline studying everything which can be used in order to lie," [Eco, p. 7]. A sign, therefore, is not a single entity, but at least two "things," or, more accurately, two "relata": "(signs) all necessarily refer to us to a relation between two relata." [Barthes, p. 35] These two relata are, after Saussure, often known as the "signifier" and the "signified."

Now while a distinction can be made between these two relata, we must be careful when talking about either by itself: " . . . language wields its signifiers and signifieds so that it is impossible to dissociate and

differentiate them . . ."; or, "The nature of the signifier suggests roughly

the same remarks at that of the signified: it is purely a relatum, whose

definition cannot be separated from that of the signified." [Op. Cit.,

pp. 44 and 47, resp.]

Signification

Properly speaking, then, we cannot talk of a signifier by itself, but

only of a relation, or, more precisely, a process in which a signifier and

a signified are correlated. This process is call signification. "The

signification can be conceived as a process; it is the act which binds the

signifier and the signified, an act whose product is the sign." [Op. Cit.,

p. 48] The signifier and the signified are what Eco, following Hjelmslev,

referred to as the "expressions" and "contents," respectively: " .... .

it appears more appropriate to use the word sign as the name for the unit

consisting of content-form and the expression-form and established by the

solidarity that we have called the sign-function [signification]."

[Hjelmslev, p. 58] Thus, the signification imposes a transitory relation

between two functives (expression and content). The signification is tran-

sitory because the expression plane and the content plane remain independent.

(" . . . a sign is not a fixed semiotic entity but, rather, the meeting

ground for independent elements." [Eco, p. 49]) An "expression" is corre-

lated with a "content" by an actual usage, and remains correlated only inso-

far as individuals use that expression to signify this particular content.

A particular content, C_1, may be correlated with ("represented by") a

particular expresssion, E_1, but the natural evolution and change of language

(the sign system of most importance here) permit the free movement of C_1 and

allow it to be correlated with a new expression, say E_n, or E_1 to be

correlated with a new content, say C_n. The new correlation signification

may coexist with, overshadow, or even supersede the original correlation

between E_1 and C_1.

> We shall therefore say in general terms that in the language the
> link between signifier and signified is contractual in its prin-
> ciple, but that this contract is collective, inscribed in a long
> temporality (Saussure says that 'a language is always a legacy'),
> and that consequently it is, as it were, naturalized; in the same
> way, Levi-Strauss specified that the linguistic sign is arbitrary
> a priori but non-arbitrary a posteriori. This discussion leads us
> to keep two different terms, which will be useful during the semio-
> logical extension. We shall say that a system is arbitrary when
> its signs are founded not by convention, but by unilateral decision:
> the sign is not arbitrary in the language but it is in fashion; and
> we shall say that a sign is motivated when the relation between its
> signified and its signifier is analogical. (Buyssens has put for-
> ward, as suitable terms, intrinsic semes for motivated signs and
> extrinsic semes for unmotivated ones.) It will therefore be poss-
> ible to have systems which are arbitrary and motivated, and others
> which are non-arbitrary and unmotivated. [Barthes, p. 51]

This brief description of signification assumes a great deal on the

reader's part, and I will try to make it clearer by explicating the con-

stituents which are correlated by the process of signification (viz., ex-

pressions/signifiers and contents/signifieds). It is important to remember,

though, that while semiotics distinguishes between contents and expressions,

these entities cannot be analyzed by themselves; they find their full defi-

nition in the act of signification—in the act of representing something

that is absent. As Hawkes [p. 130] puts it:

> . . . any semiotic analysis must postulate a relationship between
> the two terms signifier and signified . . . What we grasp in the
> relationships is not the sequential ordering whereby one term leads
> to the other, but the correlation which unites them.

Hjelmslev even warns about the dangers of considering the sign apart

from the act of signification:

> If sign is used as the name for the expression alone or for a part
> of it, the terminology, even if protected by formal definitions,

will run the widespread risk of consciously or unconsciously giving rise to or favoring the widespread misconception according to which a language is simply a nomenclature or stock of labels intended to be fastened on pre-existent things. [p. 58]

Speaking more figuratively, Hjelmslev emphasizes the unity of the sign:

> . . . we see no justification for calling the sign a sign merely for content-substance, or . . . merely for the expression substance. The sign is a two-sided entity, which a Janus-like perspective in two directions, and with effect in two respects: 'outwards' toward the expression-substance and 'inwards' toward the content-substance. [ibid.]

Expressions

Expressions are often called "sign vehicles," and in this sense they "carry" the signification. They are the words or symbols which signify some content. But there is something disconcerting about this definition. That is, if an expression must signify (loosely, "stand for") some content (signified), then how can we speak of an expression by itself (i.e., an expression that does not signify some content)? In truth, we cannot. To speak of an expression "by itself" is merely to adopt an analytical convenience rather than state a linguistic fact. When Semiotics speaks of an "expression" it means to speak of a word or symbol without regard for the content it signifies or could potentially signify. When Semiotics refers to a word or symbol in relation to its content or meaning, then it is dealing with the word/symbol as a sign, not expression. As Barthes [1972] states:

> . . . take a black pebble; I can make it signify in several ways, it is a mere signifier [expression]; but if (weight it with a definite signified [content] (a death sentence, for instance, in an anonymous vote), it will become a sign. [p. 113]

For the present, the expressions with which I will concern myself will be the words or phrases of natural language.

Contents

Now the expression, according to Semiotics, cannot stand by itself.
It must be (correctly or incorrectly) correlated with a content or meaning,
otherwise it ceases to be an expression and remains a cipher. The marks
XZ#&% are correlated with no content (loosely, "meaning") that I am aware
of (though they may be in the future). Therefore, it is not, in the strict-
est sense, an expression. But notice a curious phenomenon here: once I
have written the marks down and explained how they are correlated with no
commonly understood content (meaning), they cease to be meaningless and
become an expression for a "sign without content/meaning." In this sense,
XZ#&% can now be writen "XZ#&%." While this sleight-of-hand is unnecessary
to present discussion, it is important for the reader to understand how
elusive the study of signification may be.

The subtle interplay of signifier and signified can be seen more
clearly in the following statement by Hawkes:

> . . . a bunch of roses . . . can be used to signify passion. When
> it does so, the bunch of roses is the signifier, the passion the
> signified. The relation between the two (the 'associative total')
> produces the third term, the bunch of roses as a sign. And, as a
> sign it is important to understand that the bunch of roses is quite
> a different thing from the bunch of roses as a signifier: that is,
> as a horticultural entity. As a signifier, the bunch of roses is
> empty, as a sign it is full. What has filled it (with significa-
> tion) is a combination of my intent and the nature of society's con-
> ventional modes and channels which offer me a range of vehicles for
> the purpose. The range is extensive, but conventionalized and so
> finite, and it offers a complex system of ways of signifying . . .
> [p. 131]

Now, from a purely formal point of view we can speak of Semiotics as
dealing with three terms: the signifier (expression), the signified (con-
tent), and the sign (a signifier and signified united by the process of
signification).

The above quotation brings up a crucial issue in Semiotics—What is the nature of the signified (content) of a sign? This is not an easy question to answer. In fact, it involves issues of lasting difficulty such as "reference," and "meaning". But theorists in Semiotics have offered some suggestions. "Saussure himself has marked the mental nature of the signified by calling it a concept: the signified of the word signifier ox is not the animal ox, but its mental image . . ." [Barthes, 1968, p. 43] This is clearly not an unimpeachable definition of "signified," but it does make an important distinction, viz., that the signified, while it may be used in the process of mentioning an object, it is not the same thing as the object. To equate the content of the expression "ox" with the animal ox is to, in Wittgenstein's words, "confound the meaning of a name with the bearer of the name." [para. 40] This leaves us with the definition of the signified as a "mental image"—by itself an almost vacuous explication. Barthes goes on, "the signified is . . . neither an act of consciousness nor a real thing, it can be defined only within the signifying process, in a quasi-tautological way: it is this 'something' which is meant by the person who uses the sign. In this way we are back again to a purely functional definition . . ." [loc. cit.] We have made some, albeit modest, progress here, if only by explicating the signified in terms of what it is not. The signified is not a "real thing," even though the signifier to which it is correlated may be used to mention a real thing; and the signified is not an act of consciousness, it is somehow already "given." In other words, while we can see the "mental nature" of the signified, we can also see that it is not constructed in the mind of a person who "understands" it. That is, hearing the expression "ox" we do not construct, infer, or deduce its "mental image"

(signifier); it is, somehow already there. This is not to say that we cannot modify this mental image or learn a new one. We can, of course. What is important to the Semiotician is that we already have the mental image correlated with the expression "ox". This is the sense in which Saussure speaks of language as a legacy. In other words, for a sign to signify something (to correlate an expression and a content), the mental "something" (content) that is signified must already be "given" to the person who understands the signification. That is, for a person to understand that a rose may signify passion, the individual must not only know what a rose is, he must <u>already</u> <u>have</u> a conception ("idea") of passion. The primary new information which is provided by the correlation of a rose and the idea of passion is not the content but the <u>correlation</u> of content and expression. When the reader of James Joyce's works learns that Leopold Bloom signifies Ulysses, he must already have some idea of what both Bloom and Ulysses are. It is their correlation which gives him new information. Thus Semiotics finds it primary activity as the restructuring or explication of already existent systems of signification (or, as Saussure might say, " . . . restructuring the legacy of language.").
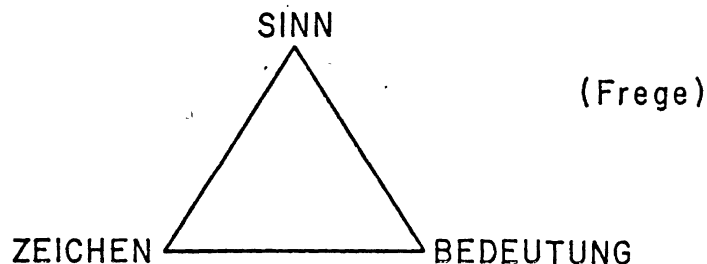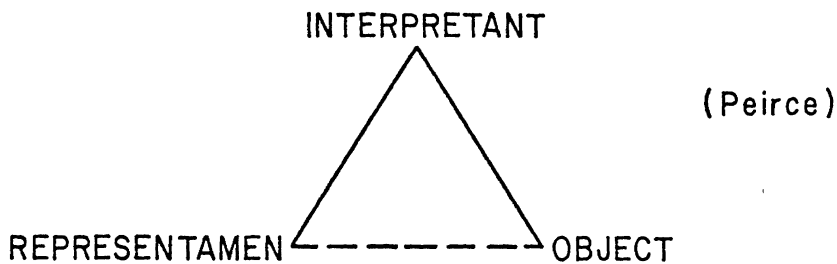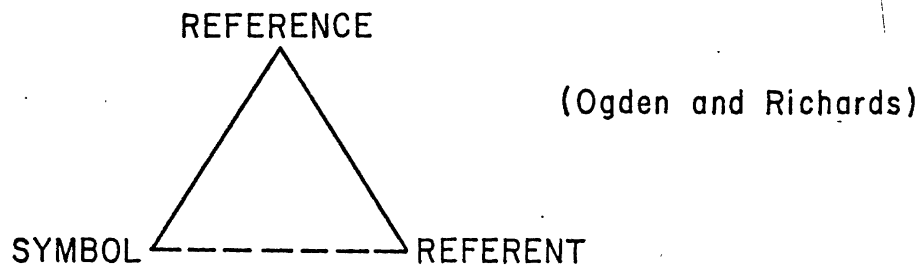
<u>Signification and Communication</u>

For Eco, the proprium of Semiotics is the possibility of lying. Semiotics itself has two main concerns: signification and communication. A Semiotics of signification is what Eco refers to as a theory of codes. A Semiotics of communication is a theory of sign-production. Here we are chiefly concerned with the theory of codes and, hence, signification. As I have stated before, signification is only possible when it is possible to lie. But note an important distinction here. It is not necessary for

Semiotics (in terms of a theory of codes) to know whether an expression is being used to lie, it is only necessary for lying to be possible. Thus we have drawn a distinction between conditions of signification and conditions of truth, or between intensional semantics and extensional semantics. The theory of codes is concerned primarily with intensional semantics and not with truth conditions. Truth conditions, and, hence, a theory of mentions, fall under the rubric of Semiotics of communication, not a theory of codes. The well-known triangles of Odgen and Richards, Peirce, and Frege, respectively, (below) are somewhat misleading when used to discuss signification.

> ". . . . such triangles can, indeed, be useful in discussing a theory of sign-production and particularly a theory of 'mentioning', but they become something of an embarrassment when studying the problem of codes signification." [Eco, p. 60]

REFERENCE

(Ogden and Richards)

SYMBOL ← — — — — —→ REFERENT

INTERPRETANT

(Peirce)

REPRESENTAMEN ← — — — — —→ OBJECT

SINN

(Frege)

ZEICHEN ← — — — — —→ BEDEUTUNG

For Eco, by including an extensional aspect, they confuse signification with communication.

If, as Eco claims, the process of signification does not have an extensional facet (i.e., if the signified/content does not refer to an object), then what do signified/contents signify? For Eco the signified is a <u>cultural content</u> or unit. The signification, in producing a sign, unites an expression (signifier) with a cultural unit (signifier/content). This sign, once it is established, may then be used to mention an object, but this mentioning, though it is important in communication, is not a part of the process of signification. Signification is prior to communication. "A signification system is an autonomous Semiotic construct that has an abstract mode of existence independent of any possible communicative act it makes possible. On the contrary . . . <u>every act of communication to or between human beings</u>-- or any other intelligent biological or mechanical apparatus--<u>presupposes a signification system as its necessary condition</u>." [Eco, p. 9] Let us look more closely at the signified as cultural unit.

<u>Signified as Cultural Unit</u>

By positing a signifying relationship between expressions/signifiers and contents/signifieds, Semiotics forces itself into the unenviable position of explicating the content of a signifying relation. Surprisingly, though, Semioticians (and Linguists) have been satisfied with an impressionistic rendering of this issue--a market contrast with their more precise discussions of coding, sub-coding, componential analysis, and similar issues. Semioticians have been content with this inconsistency, explaining signifieds as the "meaning" of a sign, or, even more discomforting, as the " . . . 'something' which is meant by the person who uses the sign." [Barthes, <u>loc.</u>

cit.] These adumbrations are tolerated only within the somewhat naive con-

text of Semiotics' concern about the problem of meaning or signification.

In other words, no Semiotician questions whether or not signs mean/signify

anything. If the signs did not mean anything, they would not be signs and

would not fall within the purview of Semiotics. Thus the fact that signs

must be meaningful in order to be studied by Semiotics appears to relieve

Semiotics of any concern for the exact nature of this meaning. Semiotics,

then, prefers to busy itself with the componential analysis of signs rather

than call into question whether or not this method of analysis has anything

significant to say about the meaning of these signs. As John Kenneth

Galbraith said on another issue, but no less applicable here, "To many it

will always seem better to have measurable progress towards the wrong goals

than unmeasurable, and hence uncertain progress toward the right ones."

Componential analysis survives, not because it solves any problems of mean-

ing/signification, but because it forms the basis for a vigorous (and

"measurable") cottage industry of linguistic analysis.

Eco has tried to improve this condition by equating contents/signifieds

with what he calls "cultural units."

> What, then is the meaning of a term? From a semiotic point of
> view it can only be a cultural unit. In every culture "a unit
> . . . is simply anything that is culturally defined and distin-
> guished as an entity. It may be a person, place, thing, feeling,
> state of affairs, sense of foreboding, fantasy, hallucination,
> hope, or idea. In American culture such units as uncle, town,
> blue (depressed), a mess, a hunch, the idea of progress, hope,
> and art are cultural units." [p. 67]

Eco goes on to explain that " . . . a cultural unit can be defined

semiotically as a semantic unit inserted into a [signification] system,"

[ibid.] and then makes the final bridge to a pragmatic theory of

signification with the statement (truism?) that "Recognition of the presence of these cultural units (which are therefore the meaning to which the code makes the system of sign-vehicles correspond) involves understanding language as a social phenomenon." [ibid.] Eco's purpose is clear in his attempt to incorporate pragmatic elements into a traditional form of linguistic analysis. This represents the next logical step in the development of contemporary linguistic theory which has its antecedents in such works as Chomsky's Syntactic Structures, Lakoff's Irregularity in Syntax, and Katz and Fodor's "The Structure of a Semantic Theory." Theoretical linguistics has long aspired to the formal elegance of mathematics and logic, and the work of Chomsky, et al., has been characterized by its efforts to develop models of language which are as formal as possible (Chomsky's model of a purely syntactic deep structure of language is the extreme example of this formalization). The discovery of many anomalies (notably by Lakoff) contributed to the breakdown of the purely syntactic paradigm of language. Chomsky responded by including semantic and phonological elements in his language model which would act as "interpretants" for deep syntactic structures as they were transformed into surface structures [Chomsky, 1965] (natural language). But even this did not represent a sufficient model of language. Katz and Fodor argued for a more fully developed semantic model of language, feeling that linguistic competence could never be sufficiently described by a largely syntactic model. But although Katz and Fodor insisted on semantics as a sine qua non of a satisfactory theory of language, they resisted any attempt to expand their model to include such pragmatic elements as the context or circumstances in which words may be articulated. Eco includes in his book a criticism of the Katz-Fodor language model,

arguing that it is insufficient because, among other things, it does not
include necessary pragmatic elements. [pp. 96-105] This is the reason
behind Eco's interpretation of contents/signifieds as cultural units and
his emphasis on language as a social phenomenon. Eco has revised the
semantic method of componential analysis, as developed by Katz and Fodor,
to include pragmatic markers (or elements) such as circumstances and con-
texts.

While the trend of linguistic development from Chomsky through Eco
follows a logical progression, and it seems clear that pragmatic elements
must be considered in any reasonable language model, it is not clear that
Eco's "meaning-as-cultural unit" makes linguistic/semiotic analysis any more
tenable than it once was (this may be why Katz and Fodor resisted moves in
this direction). Eco's cultural units do not solve the problems of
linguistic/semiotic analysis because it is unclear what these units are.
Yet because the cornerstone of Semiotics is, by definition, the relation
between expressions/signifiers and contents/signifieds, the nature of the
latter demands a clearer explication. Eco's (and Schneider's) claim that
a cultural unit " . . . is simply anything that is culturally defined and
distinguished as an entity . . ." won't resolve this ambiguity. Eco does
show that he rejects a referential theory of meaning/signification, so the
"culturally defined entities" are not tangible things. But Eco does not
show how his cultural units differ from an ideational theory of meaning.
His (Schneider's) explication of cultural units as including " . . . feel-
ing, state of affairs, sense of foreboding, fantasy, hallucination, hope,
or idea . . ." can only be seen (without some clear caveat) as a somewhat
more complicated version of Locke's " . . . words in their primary or

immediate signification stand for nothing but the ideas in the mind of him
that uses them . . ." [op. cit. 3.2.2] In other words, the content/signi-
fied is an idea (feeling, sense, fantasy, etc.). In this sense, Eco's
pragmatic theory of Semiotics falls prey to the same difficulties as idea-
tional theories of meaning.

Problems with Ideational Theories of Meaning

The difficulties with ideational theories of meaning are well known in
the field of Philosophy of Language. [Black] Briefly stated, these problems
are of two major types:

1. The problem of verifiability. If there does exist, as Semioticians
claim, a "content-plane" (or set of contents/signified) to which expressions
are somehow related in the action of signification, and these contents can
only be explicated in some kind of mentalistic way (as ideas, feelings, con-
cepts, etc.), then how does the speaker of a language determine whether he
has the "correct" (or, culturally acceptable) idea (content) which an ex-
pression may signify? As Black puts it,

> It is part of the "mentalistic" conception to assume that the
> "idea" is something "private" to its possessor, something of which
> only he can be directly aware. But if so, how am I to convey my
> "idea" to you—or to be sure that the idea you have corresponds
> sufficiently closely to my own? [p. 194]

Verification is important here because language is learned, and
speakers of a language are continally acquiring new words and phrases. An
individual must have a way of determining whether he correctly understands
words (expressions/signifiers) in the language. Otherwise, there would be
no way that an individual could acquire a language or expand and refine
his command of it.

2. The problem of the nature of the "idea" as content/signified. Even

if we could assume that we have a clear mental image/idea/concept whenever

we hear or read a word (expression)—and this, itself, is a major assump-

tion—we cannot easily ascribe the meaning/significance of the word to this

mental image. The reason for this is that the mental image (or "picture")

will usually require some interpretation itself before it can be linked to

an expression. As Black states,

> When I hear the word "three" I may "see" three white spots ar-
> ranged as a triangle on a green background (perhaps through
> association with green blackboards seen in school), but I know
> that the arrangement of the spots and their colour "does not
> count," is irrelevant. But this means that the image itself,
> even if it does occur, must be interpreted . . . (The sound
> "three" might conceivably evoke the image of a pentagon, and the
> name of a colour might evoke the image of a complementary
> colour.) Thus recourse to images, even when they do occur, ac-
> complishes nothing of importance in the search for an entity to
> serve as the meaning of a given word. The image itself stands
> for something, has meaning, and we have merely pushed the search
> for meaning one stage further. [p. 195]

Or, as Wittgenstein [1953] remarks in a similar vein,

> I see a picture; it represents an old man walking up a steep
> path leaning on a stick—How? Might it not have looked just
> the same if he had been sliding downhill in that position?
> [para. 139]

or,

> Imagine a picture representing a boxer in a particular stance.
> Now this picture can be used to tell someone how he should
> stand, should hold himself; or how he should not hold himself;
> or how a particular man did stand in such-and-such a place; and
> so on. One might (using the language of chemistry) call this
> picture a proposition radical. [para. 23]

At best, Eco's "cultural units" hint at how the contents/signifieds

(which are assumed to exist) are formed and alter accepted linguistic theory

to include pragmatic elements. But in spite of these changes, Eco (and mutatis

mutandis, other Semioticians) has not been able to change the substantially

ideational character of Semiotics.

Consequences of an Ideational Semiotics

The problems with an ideational theory of signification mean that there can be no clear definition of the contents/signifieds. Yet the objective of Semiotics remains the description of the signifiying relation between expressions/signifiers and contents/signifieds, and, frequently, the representation of this relation by means of some sort of Structural Semantics like Componential Analysis (the representation of the signifying relation between expressions and contents). Without a clear definition of contents/ signifieds—or at least a definition that does not have the ideational problems mentioned above—it is difficult to speak convincingly of (or represent) a signifying relation between expression and content. Furthermore, these ambiguities make it unclear exactly what Structural Semantics does in its representation of meaning/signification.

The Componential Analysis (probably the most popular form of Structural Semantics) of the signifying relation between expressions and contents qua "cultural units" usually consists of relating expressions ot other expressions; e.g.,

"father" = X parent of Y + Male X + (Animate Y + Adult X + Animate Y)

But clearly the expressions on the right side of the equation are something more than just expressions, otherwise the equation would reduce to a not very informative form of "de-coding". The right side of the equation, Eco claims, represents the "contents-as-cultural units" which "father" signifies. But here we find ourselves heading back into the cul-de-sac of ideational theories of meaning/signification (i.e., expression as content as cultural units as . . .?). Why, a reasonable and prudent individual might ask, are these ambiguous representations of meaning/signification tolerated? The

answer that comes immediately to mind is that they only represent what we

(collectively) already know. In this sense, the Componential Analysis is

no more revealing than a definition in the O. E. D. or Encyclopedia

Brittanica (Eco even draws an analogy between sememes (contents represented

in Componential Analysis) and an encyclopedia, [pp. 112-114] and the format

of representation adds a semblance of rigor to an otherwise ordinary pro-

cess of definition.

But to return to the central difficulty: without a clear definition

of contents/signifieds most Semiotics reduces to a topography of imaginary

landscapes. Furthermore, this ambiguity calls into question the very

existence of contents/signifieds themselves. In other words, Semiotics'

inability to give a satisfactory definition of contents/signifieds reduces

the status of the content plane to an implicit assumption. By insisting

that there are contents/signifieds which play an important part in the pro-

cess of signification, Semiotics has fallen prey to what Wittgenstein [1958]

called a "disease of thinking":

> There is a kind of general disease of thinking which always
> looks for (and finds) what would be called a mental state from
> which all our acts spring as from a reservoir. [p. 143]

The reason why Semiotics unquestioningly maintains the existence of the

content plane is that the fundamental question which orders Semiotic inquiry

("what does a given expression mean/signify?") leads ineluctably to some

"other" which is somehow "tied" to the expression in question. This "other"

has been variously described by Semiotics in referential, ideational or

behavioral ways, but all these theories, though they may differ as to the

nature of what expressions mean/signify, maintain an unavoidable dichotomy

between expressions and contents. In short, by asking the question in the

form "what does a given expression mean/signify?" Semiotics has limited

severely the ways it can be answered (or what would count as an answer).

In effect, there is an implicit assumption on the question, namely, that

there does exist a "something else" that an expression means/signifies.

The very existence of this "something else" is not questioned in Semiotics,

only its nature is investigated (i.e., whether it is an object, concept, or

cultural unit). As a result, the existence of contents/signifieds (the

"something else") and the dichotomy between them and expressions form an

"inherited background" for Semiotics and make it difficult to question the

existence of contents/signifieds without radically changing the entire

structure and method of the field. This leaves the researcher with only

two courses to follow: 1. Argue convincingly for the existence of contents/

signifieds; or 2. Base his inquiry on a question that does not have this im-

plicit assumption.

Without a satisfactory and realistic definition of contents/signifieds,

it would be virtually impossible to argue convincingly for their existence.

Consequently, further inquiry into the nature of signs and signification is

contingent upon the formulation of a new fundamental question--one which does

not have implicit assumptions and is not beset by the same ideational or

referential problems, yet would be useful in the investigation of signs and

signification. Instead of asking "what does an expression mean/signify" we

shall now ask "how is an expression used?"

The Alternative to Ideationally Based Semiotics

By re-orienting the inquiry away from a search for "meaning" and towards

an examination of how expressions are used, we have avoided the problems of

ideational Semiotics, while preserving productive avenues of investigation.

In order for us to understand an expression as a meaningful sign we are no longer compelled to search for a content that somehow infuses it with significance. Instead, our investigation is now primarily concerned with how expressions are used. Not only does this perspective avoid ideational problems, it also solves the problem of verification (q.v.). In other words, a question about the meaning or signification of an expression now becomes a question about how an expression is used, and disputes about the "meaning" of an expression are resolved by looking at the patterns of usage for the expression. In effect, we have adopted a kind of Copernican Reversal in the investigation of expressions and signification. Recent linguistic theory, notably Chomsky's, takes syntax as the fundamental ordering principle of language, and semantics or pragmatics as being interpretive of syntactic deep structures. The orientation of this investigation is that pragmatics is the fundamental ordering principle of meaning/signification and that questions of syntax and semantics (if they are addressed at all) are derivate from pragmatics. This is, of course, not a new observation, as Yehoshua Bar-Hillel [1972] points out:

> Linguistics was considered by Peirce and his followers to be the theory of a certain specific subclass of signs, the linguistic signs, or symbols, and the tripartite division into pragmatics, semantics, and syntax, was carried over to it . . . According to this scheme, then, semantics and syntax are arrived at from pragmatics by successive abstractions."

Useful Aspects of Semiotic Theory. That Can be Accommodated in the Pragmatic Interpretation

In spite of the critical approach which we have taken towards Semiotics, there remain useful aspects of it which will be helpful in our discussion of the pragmatic aspects of inquiry.

1. The notion of "unlimited Semiosis." C. S. Peirce was the first

individual to do substantive work in the analysis of signification, and, although much of his work has the same inherent ideational problems which we have discussed before, some of his observations have relevance to our own discussion. Most important of these observations is his awareness of the unlimited nature of the meaning of signification:

> The meaning of a representation can be nothing but a representation. In fact it is nothing but the representation itself conceived as stripped or irrelevant clothing. But this clothing never can be completely stripped off; it is only changed for something more diaphanous. For there is an infinite regression here. Finally, the interpretant is nothing but another representation to which the torch of truth is handed along; and as representation, it has its interpretant again. Lo, another infinite series. [Peirce, 1931, section S.492]

In short, Peirce is pointing out that there can never be a necessary and sufficient explanation or description of the meaning of a sign/expression. In terms of signification, this means that there can never be a complete description of the kinds of allowable uses that can be made of a given expression. But this is not a despairing observation; in fact, it puts our analysis into a more thoughtful context. Instead of concerning ourselves with definitive uses of expressions, we can recognize this endless regression of meaning/signification and concentrate on elucidating conventional uses of expressions, realizing that new and creative uses of these expressions are inevitable. What is important, then, is not just the uses of an expression, but the conventional uses of that expression in relation to some situation or task at hand.

Disregard for "unlimited semiosis" and the accompanying belief in the necessary and sufficient definability of linguistic expressions have proved fatal for many projects in formal linguistics (notably in fully automatic machine translation and artificial intelligence efforts in natural language

processing). The logician Yehoshua Bar-Hillel [1964] recognized this combinatorial problem in the 1950's, drawing an a propos analogy between the representation of meaning of an linguistic expression and the endless (and undecidable) ways that the same number could be represented (e.g.,

$$3 = 1 + 2 = 4 - 1 = +\sqrt{9} = 1158 - 1155 = \ldots).$$

2. Denotative and connotative aspects of significance. The concepts of denotation and connotation have always held positions of importance in most theories of linguistics/semiotics, although there has been a great deal of controversey over what they mean. In some theories they distinguish between "sense" and "reference", in others they mark the division between "intensional" and "extensional" semantics, while in some theories they distinguish between "emotive" and "referential" meanings of expressions. This exegesis of denotation and connotation could be extended indefinitely. But what is of primary importance here is not the precise definition of these two "types of meaning." What is important is the recognition that (speaking loosely) there are "primary" and "secondary" levels of significance for expressions. In terms of this discussion, we may now identify primary and secondary uses of expressions.

How are Words Used: The Tool Analogy

The reasonableness of our concern for the use of expressions rather than their meaning/signification can be better seen if we consider words as tools suited to perform certain tasks at hand. A given tool has certain kinds of jobs for which it can be used, and its usefulness in these tasks depends on the design of the tool and the skill of the user. Like expressions in language, there are "primary" and "secondary" uses of tools, as well as an unpredictable assortment of "creative" uses. For example, the

primary use of a screwdriver would be to turn screws of a certain size and type, and an accepted secondary use of the screwdriver would be as a lever to pry open a can of paint. A creative use of the screwdriver might be as a wedge under a door to hold it open. These kinds of uses of tools are similar to denotative, connotative, and creative uses of words (in the sense that we explained these distinctions in the last section). Thus, the uses of a given tool depend not just on the tool's inherent physical qualities, but also on the skill and background of the person who uses the tool and the task at hand to be accomplished. Like tools, words can be used effectively, misused, used carelessly, and even abused. Also, the analogy of words-as-tools put the concern for the meaning/signification of expressions in a better light. To ask a carpenter what one of his tools "means" or "signifies" would be a very strange question, indeed. More appropriately, we would ask him how, or for what, a particular tool is used, and in what kinds of situations these uses would be appropriate. We can also see that in spite of the fact that a tool may have a wide variety of possible primary and secondary uses, there are certain conventional ways that a tool is used. These conventional uses for a tool provide a kind of core of primary usage patterns surrounded by an unbounded penumbra of secondary and creative uses of the tool. These secondary and creative uses of secondardy and creative uses of the tool. These secondary and creative uses of a tool may at any time (though this is not totally arbitrary) become primary, conventional uses of that tool. In a carpenter's shop, the use of a screwdriver to open a can of paint would definitely be a secondary usage. But in a painter's shop, this might be the primary use of a screwdriver. These same patterns of usage obtain for linguistic expression (and for signs in general), also.

This words-as-tools analogy also puts the problem of ambiguity in language in a less frustrating context. There is, no doubt, a great deal of ambiguity in language, but its extent is frequently overrated. Chomsky's famous phrase "flying planes can be dangerous" is a paradigm example of this kind of ambiguity. [1965, p. 21] But Chomsky is engaged in a linguistic shell-game here, and has created an ambiguity that does not exist if one looks at how the phrase "flying planes can be dangerous" is used. That is, the phrase "in vacuo", with no consideration for the situation in which it occurs, is, in fact, an ambiguous phrase. But the ambiguity arises because the application of the phrase is not given and thus remains unclear. On the other hand, if we meet a friend climbing out of the cockpit of a small plane, his face ashen, and his hands trembling, saying, "Flying planes can be dangerous," there is no ambiguity in the sense that Chomsky indicates. As Wittgenstein [1953] remarked,

> One cannot guess how a word functions. One has to <u>look at</u> its use and learn from that . . . But the difficulty is to remove the prejudice which stands in the way of doing this. It is not a <u>stupid</u> prejudice. [para. 340]

In the tool analogy, Chomsky's example would be like holding up a screwdriver and saying it is ambiguous. But if we ask how the screwdriver is used, as we asked how Chomsky's phrase might be used, the ambiguity is resolved, or at least lessened.

This argument is not meant to imply that there is no ambiguity in language; there is, of course. The point here is that much ambiguity can be lessened by an appeal of how expressions are used.

<u>George Kingsley Zipf</u>

At this point, our words-as-tools analogy is little more than an

interesting parallel which we have used to smooth the transition from idea-

tional semiotics to a pragmatic interpretation of language. For a more

developed analysis of this parallel it is helpful to look at the work of

George Kingsley Zipf.

Zipf offers much empirical evidence for the performance of words in a

tool-like manner, and although this evidence is beyond the scope of this

study, the implications which he draws are important. In the first place,

language use can be explained in terms of "tools-and-jobs." In other

words, tools do not exist by themselves; they exist in relation to certain

kinds of jobs for which they can be used.

> The problem of tools-and-jobs is the same as the problem of
> means and ends, or of instruments (or agents) and objectives.
> We shall adopt the homelier term, tools-and-jobs, to emphasize
> the commonplace nature of the problem under discussion [p. 10]

But this tools-and-jobs nature of language is fundamentally reciprocal.

It means that tools require certain kinds of jobs in order to be used, and

jobs require certain kinds of tools in order for the job to be accomplished.

This reciprocity has the function of fitting available tools to required

jobs, and, at the same time, altering the jobs to be performed to fit the

functions of available tools.

> . . . there are two aspects of the economy of the tools-and-
> jobs in question. In the first place, there is the economy of
> tools. In the second place, there is the economy of jobs . . .
> To clarify the significance of these two economies, let us il-
> lustrate them briefly in terms of carpentry tools and carpentry
> jobs . . . We all know from experience that when a person has
> a carpentry job to be performed, he directly or indirectly seeks
> a set of carpentry tools to perform the job. And, in general,
> we may say that jobs seek tools . . . But what is often over-
> looked is the equally obvious fact that when a person owns a set
> of carpentry tools, then, roughly speaking, he directly or in-
> directly seeks a carpentry job for his tools to perform. Thus,
> we may say that tools seek jobs. [p. 8]

Zipf calls this the "reciprocal economy of matching tools to jobs and jobs to tools." In linguistic behavior, this reciprocal economy manifests itself in the following way: The words and phrases in language are, of course, the tools, and (speaking loosely) the conveying or description of information, the asking of questions, or the discussion of subjects, comprise (inter alia) the jobs for which these words/tools must be used. For example, if I want to ask someone what time it is, I have certain words and phrases that are more-or-less suitable for this task (consider how the use of conventional English as opposed to the use of slang can be used to accomplish this job differently). We are so fluent in our native language that this reciprocal economy of tools and jobs is not readily apparent (it is too obvious). But we can see this reciprocity more clearly if we imagine ourselves in Germany with only a modest ability to speak German. Here, the jobs is the same--to find out what time it is. But suppose we don't know the standard phrase, "Wieviel uhr ist es?" Here, we must fit our limited tools (German words and phrases) to the task by using whatever German words we know that could possibly be used to find out what time it is. We may be able to get by with the contrived phrase, "Was ist die uhr?" (accompanied, perhaps, by appropriate gestures). But if we look closely at this situation, we can see another change taking place. Because our command of German is minimal, we would not comprehend a detailed or precise explanation of the time such as, "Es ist zehn nach halb acht." We can only comprehend such phrases as "Sieben uhr," or, "Ungefähr sieben uhr." Thus, the limited tools available to use have changed our task at hand from finding out exactly what time it is, to finding out roughly what time it is.

The reciprocal economy of tools and jobs is readily apparent in the

activity of inquiry. Consider a search through an indexed collection of
documents. The job that the inquirer has is to find several documents (in
the collection) that will be useful for a particular informational task.
The tools which he has at his command are whatever textual descriptions he
knows (or can discover) that are used to index documents in the collection.
Now, the nature of the informational task (the job) will determine what
textual descriptions (tools) will be used to search the collection. But
at the same time the number of useful textual descriptions that the in-
quirer knows (or has available) will determine how easily or well he per-
forms the task of finding useful documents. In other words, if the in-
quirer knows only a few of the textual descriptions in the indexing vocab-
ulary that are appropriate for his search, he may (consciously or uncon-
sciously) alter his objective from finding "exactly the documents he wants"
to finding "anything useful". Thus, the task for which the inquirer wants
information determines what textual descriptions he would like to use, and
the descriptions he actually uses in his search will determine the exact
definition of the final goal (in terms of this particular collection).

Ludwig Wittgenstein

So far we have demonstrated how the meaning/signification of expres-
sions can be explained usefully by looking at how these expressions are used
in language, and we have developed this shift by means of an analogy between
expressions and tools. Further, we have expanded this instrumentation anal-
ogy in terms of Zipf's notion of the reciprocal economy of tools and jobs.
But this concern with the use of expressions, while appealing, is somewhat
vague. To understand more fully what we mean by explaining the meaning/
signification of expressions by their employment, we must now turn to the

development of this theme in Wittgenstein's late philosophy.

For Wittgenstein [1953], as for Zipf, words could be seen more clearly as tools:

> Think of the tools in a tool-box: There are a hammer, pliers, a saw, a screwdriver, a rule, a glue-pot, glue, nails, and screws.--The functions of words are as diverse as the functions of these objects. [para. 11]

The parallel between words and tools is important in order to focus our attention on the functions of words rather than their meanings. As Pitkin points out:

> . . . understanding a language is not a matter of grasping some inner essence of meaning, but, rather, of knowing how to do certain things. "To understand a language means to master a technique." [p. 36]

The tool analogy also makes us aware of a common pitfall in linguistic analysis: the faulty inference from the uniform appearance of the same word in different contexts to the assumption that some uniform, essential meaning accompanies this word in its different contexts.

> Of course, what confuses us in the uniform appearance of words when we hear them spoken or meet them in script and print. For their application is not presented to us so clearly. . . . It is like looking into the cabin of a locomotive. We see handles all looking more or less alike. (Naturally, since they are all supposed to be handled.) But one is the handle of a crank which can be moved continuously (it regulates the opening of a valve); another is the handle of a switch, which has only two effective positions, it is either off or on; a third, the handle of a brake-lever, the harder one pulls on it, the harder it brakes; a fourth, the handle of a pump: it has an effect only so long as it is moved to and fro. [Wittgenstein, 1953, para. 11-12]

A word (like a tool), regardless of how it is used, always looks the same. It is this similarity in appearance that makes us think that there is some essence or meaning that accompanies the word at all times. This is the mistake that Semiotics makes when it ties a word/expression to some

"content." Here Wittgenstein admonishes us not to assume that the uniform-
ity of a word's appearance in different contexts indicates a uniformity of
meaning. Instead, he suggests that we look and see whether the words are
used in the same way each time. In this manner we must "Let the use teach
us the meaning." [1953, p. 212]

Semiotics begins from the perspective that certain words/expressions
exist and that they need explanation. Wittgenstein begins from a more prag-
matic perspective: " . . . we don't start from certain words, but from cer-
tain occasions or activities." [1972, p. 3] Thus for Wittgenstein, "An
expression has meaning only in the stream of life" [Malcolm, p. 93].

But if we don't begin with the word/expression and its "explanation",
how do we learn the use of new words (or new uses of familiar words). We
learn new words by means of what Wittgenstein calls "language games." These
language games provide a kind of framework in which the examples of how a
word is used will "make sense." In other words, the role of a word (or its
conventional usage) can only be understood in the context of some language
game which already is known. The individual who does not know the relevant
language games will not learn much from examples of how a word is used. As
Zabeeh points out:

> Linking the concept of language game to "kinds of use" of expres-
> sion is the stepping stone for stating that "speaking of language
> is part of an activity, or a form of life." [p. 341]

These language games are not hidden, arcane processes in language.
They are right before us at all times. There are "countless kinds" of them,
and new ones are continually coming into existence, while "others become
obsolete and get forgotten." Wittgenstein [1953] gives some examples:

Giving orders, and obeying them—
Describing the appearance of an object, or giving its measurements—
Constructing an object from a description (a drawing)—
Reporting an event—
Speculating about an event—
Forming and testing a hypothesis—
Presenting the results of an experiment in tables and diagrams—
Making up a story; and reading it—
Play-acting—
Singing catches—
Guessing riddles—
Making a joke; telling it—
Solving a problem in practical arithmetic—
Translating from one language into another—
Asking, thanking, cursing, greeting, praying.   [para. 23]

Wittgenstein resists the temptation to rigorously define language

games. In fact, though these language games are alike in many ways, no com-

mon, defining thread runs through all of them. They resemble each other in

the way that family members resemble each other. It is sufficient for

Wittgenstein to list enough examples of what he means by language games so

that "we get the idea" and know how to go on and enumerate our own examples.

Wittgenstein [1970] does give us a very general description (but not a defin-

ition) of what he means by language game: "We call something a language game

if it plays a particular role in our human life." [p. 177] Of course this

description assumes we already have some idea what language games are. Its

purpose is to focus our attention on one aspect of language games:

> The idea that a language game is something that "plays a parti-
> cular role in our human life" (though vaguely) is important.
> Since even at this early stage it connects language games with
> specifiable activities and in an oblique way shows that a mere
> use of words (or even use of a grammatically well-formed expres-
> sion in the absence of certain actions, such as informing or
> warning or referring) is not to be considered as playing a
> language game. [Zabeeh, p. 331]

The "mere use of words" is not enough to teach us anything about them.

The words must be used in a language game that the hearer always knows:

"The ostensive definition," says Wittgenstein, "explains the use—
the meaning—of the word only when the overall role of the word in
the language is clear. Thus, if I know that someone means to ex-
plain in colour-word to me, the ostensive definition 'That is called
"sepia"' will help me to understand the word." Only if I know what
a colour is am I fully ready for the meaning of "sepia." Here
again, knowing what a colour is means being able to do something,
knowing how colour terms are used. [Pitkin, p. 43]

Now, if we must understand the relevant language games before we can

understand the use of words which have a role in these games, the next

logical question to ask is how we learn these language games.

Wittgenstein answers that the child learns to master language
games not by explanation, but by training. "How do I explain
the meaning of 'regular,' 'uniform,' 'same,' to anyone?—I
shall explain these words to someone who, say, only speaks French
by means of the corresponding French words. But if a person has
not yet got the concepts, I shall teach him to use the words by
means of examples and by practice." Training differs from ex-
planation in at least these two ways: it is relatively nonverbal,
relying on gestures, facial expressions, and the like; and it aims
primarily at producing certain actions from the learner, quite
apart from what goes on in his head. [Pitkin, pp. 43-44]

Wittgenstein was attacking the traditional view that people learned

language by means of explanations or definitions alone. He challenged this

in two ways:

First, he seeks to show . . . that the grasping of definitions or
essences or universals cannot explain what needs to be explained.
And, second, he tries to show that even the mastery of definitions,
principles, generalities, depends ultimately on our natural human
capacities and inclinations, which do not themselves have any fur-
ther explanation . . . The kind of training that is necessary to
the acquisition of a natural language, Wittgenstein says, requires
"inducing the child to go on" in the same way, in new and different
cases. This is different from training for repetition, which "is
not meant to apply to anything but the examples given"; this teach-
ing "point beyond" the examples given. [Pitkin, p. 45]

What permits the teaching to "point beyond" the examples given is the

learner's familiarity with the language games involved. But it is important

to remember that the teaching does not "point beyond" to some kind of

essence of underlying meaning in language; it "points beyond" to an ability

to _use_ the expressions appropriately in the learner's everyday discourse.

These examples are not indirect means of imparting some further meaning to

. the student;

> . . . to put it baldly, there is no further knowledge that the
> teacher has at which his examples only hint. The examples con-
> stitute his knowledge, too. When I teach someone a new concept
> (as distinct from a new name to fit into a system of concepts,
> a language game he has mastered already) by example and
> practice, "I do not communicate less to him than I know my-
> self." Of course the teacher knows the formula, the rule, the
> definition; but that can be _explained_ to the pupil who has the
> necessary concepts, has mastered the relevant language games.
> For such a pupil, it does not need to be _hinted at._ The place
> where explanation fails and training is called for is where the
> pupil lacks the knowedge of how to _use_ the word. And that kind
> of knowledge is completely contained in the examples; about how
> to use the words, the teacher himself knows only from the
> examples he has mastered. The knowledge of language games is
> a "knowing how" rather than a "knowing that." [Pitkin, pp. 4-48]

All of this is not to say that explanations are not possible or are

irrelevant. The point to understand is that the language game of "giving

explanations" must be understood before an explanation itself can be

understood. As Paul Ziff commented in a similar vein, "I throw a cat a

piece of meat. It does not see where the meat fell. I point to the meat;

the cat smells my finger." The cat doesn't understand the "game of point-

ing. [p. 97-93]

These language games themselves fall within a broader context. "The

introduction of _actions_ into the fabric of language links the idea of

'language game' with the idea of 'form of life . . ." [Zabeeh, p. 333] For

Wittgenstein, the forms of life are the everyday human activities which make

up our lives in a social sense. Speaking loosely, the forms of life provide

a kind of context in which language, in general, and language games, in

particular, make sense. ["To imagine a language means to imagine a form of life." [1953, para. 19]) Language is one of the forms of life, one of these activities.

Here, too, we can see the reciprocal economy of tools and jobs. The language games are tools by which an individual engages in the task of learning a language. The mastery of these tools determines how well (or how easily or in what style) the language will be learned, and the required proficiency level (or style) of language ability determines how extensively the language games must be used in learning. Likewise, the learned language becomes a tool that enables the individual to participate in certain everyday activities, while the degree and kind of participation that is required in these activities determines how and in what way the language should be learned. In this way we can see how inextricably tied together and mutually influencing are language and human activities/forms of life.

> That notion [forms of life] is never explicitly defined, and we should not try to force more precision from it than its rich suggestiveness will bear. But its general significance is clear enough: human life as we live and observe it is not just a random, continuous flow, but displays recurrent patterns, regularities, characteristic ways of doing and being, of feeling and acting, of speaking and interacting.

> Because they are patterns, regularies, configurations, Wittgenstein calls them forms; and because they are patterns in the fabric of human existence and activity on each, he calls them forms of life. The idea is clearly related to the idea of a language game, and more generally to Wittgenstein's action-oriented view of language. "The speaking of language," he says, "is part of an activity, or form of life." How we talk is just part of it, is imbedded in, what we do. "Commanding, questioning, recounting, chatting, are as much a part of our natural history as walking, eating, drinking, playing." We all know our shared forms of life, these basic, general human ways of being and doing, though they have never been taught to us and we could not begin to be able to put into words what we know about them. Wittgenstein says that they are part of our "natural history," regularities "which no one has doubted, but which have escaped remark only because they are always before our

eyes." . . . The notion of forms of life should help us to under-
stand the sense in which language may be said to be conventional.
[Pitkin, pp. 132-133]

Because it is a form of life, and because it is an instrument to be

used in the participation on human activities, language is also largely con-

ventional. This is similar to what Saussure meant when he wrote that lan-

guage is a legacy. But Wittgenstein understood the "natural history" of our

human activities, our forms of life, as something far richer than the word

"conventional" implies. Forms of life are the embodiment of our lives as

social beings; they are more than just a legacy, they are the inherited

background of intensely human activities that are common to people from one

generation to the next." . . . now we are thinking of convention not as the

arrangements of a particular culture has found convenient . . . Here the

array of 'conventions' are [sic] not patterns of life which differentiate

men from one another, but those exigencies of conduct which all men share."

[Cavell, p. 98] Thus, language is conventional because the activities in

which it embedded are repeatible and, thus, conventional.

The forms of life are the activities we engage in every day: eating,

drinking, walking, guessing, explaining, hinting, describing, joking, chat-

ting, searching, categorizing, advising, questioning, keeping informed,

evaluating, etc. We can talk about these activities in a general, familiar

way; we can outline the loose rules of these activities; and we can use them

to provide a background for the finer analysis of how we use particular words

in specific language games. But we cannot definitively explain or give rea-

sons for these forms of life themselves. As Wittgenstein [1953] put it,

"What has to be accepted, the given, is--so one could say--forms of life."

[p. 226] "If I have exhausted the justifications I have reached bedrock,

and my spade is turned. Then I am inclined to say: 'This is simply what I do,'" [1953, para. 217]. The forms of life are what we do.

Recapitulation

It is important now to outline the general structure of this pragmatic theory of language. In the first, and most fundamental, place we have the fabric of human activities (forms of life) which comprise "what we do." In terms of inquiry these activities consist of processes such as: keeping informed, giving or receiving advice, observing, evaluating information, comparing texts, describing works, exploiting collections of information, reading, finding the best textual means to some end, formulating a query, articulating an information need, doing research, discussing issues, determining the relevance of information, determining the usefulness of information, setting up an information system, buying books, subscribing to journals, attending a conference, attending a class, recommending a lecture to a colleague, etc. This list, like Wittgenstein's list of language games, could be extended indefinitely. It is also important to note that these activities may overlap in many areas, and this is not a fault. What is significant about these activities is that as long as we desire to participate in them (and become proficient in them) they comprise a set of tasks to be accomplished. We are interested in these activities insofar as we must use written or spoken language in order to participate in them.

The next question to be asked must be aimed at clarifying how we use written or spoken language to pursue an activity; and this question implies an understanding of the language games involves in the process of inquiry. In order for us to use certain linguistic tools and in order to use these tools appropriately we must understand what role these tools play in relevant

language games (such as describing the subjects or references of written texts).

It is important to distinguish as best we can between language games and forms of life. The language game "describing written texts" could also be seen as form of life insofar as it constitutes a human activity which is commonly engaged in (at least by a certain group of people). In this sense, it is more than just a linguistic process; it is a human activity that includes much that is not linguistic, such as: motivations, expectations, physical skills, satisfaction, evaluation, etc. The notion of textual description as a language game has a more limited scope than its notion as a form of life. As a language game, textual description comprises a set of patterns or regular ways of using language. The language game of textual description should make clear the role of the words or phrases used in this process. It will provide a framework in which the words or phrases used will have a role. Here, a familiarity with the language game of textual description will (hopefully) enable a person to understand how such descriptions are used and what would count as a textual description. The broader context of the form of life of textual description would provide the background for descriptions in which we might see why textual descriptions are used and how they contribute to the accomplishment of a certain task. In a more formal sense, the language game provides a kind of grammar for the use of expressions; it enables us to tell whether our use of an expression in a particular linguistic context is "well-formed" or not. The activity which the language game supports tells us whether the use of these expressions is appropriate for, or useful in the pursuit of this activity. We can see that while a given expression may be used correctly (its role in the language

game is understood), it may nevertheless be inappropriate for the task at hand. It is this "dissonance" between correct uses of an expression (in a language game) and the appropriateness of this correct usage for the support of a particular activity, that forms the primary problem area in the description of texts.

In order to complete this recapitulation, it is helpful to look at how the previously discussed aspects of our pragmatic theory of language fit into the general framework of activities and language games. We already discussed the reciprocal economy of tools and jobs in this context: namely, that language games are tools with which we learn the use of expressions, which, in turn, we use to participate in certain activities. These activities, in their part, require the use of certain expressions, which, in turn, require the mastery of specific language games in order to understand the usage of these expressions.

Primary and secondary uses of expressions (loosely, denotations and connotations) must be evaluated in terms of the activity where they are used. Both kinds of usage are necessarily well-formed, so the language game in which they find a role will offer little to distinguish them. The distinction between primary and secondary uses of an expression can only be made in terms of a task to be accomplished or an activity in which to engage, and in relation to the expectations that participants in that activity have.

Finally, the significance of unlimited semiosis in the evaluation of the meaning of expressions cannot be over emphasized. The number of definitions or meanings which can be given to any expression are, both practically and theoretically, unlimited (Wittgenstein's observation that there are "uncountable" language games is a recognition of a similar problem). This

means that necessary and sufficient definitions of expressions cannot be given. Any such definitions, no matter how comprehensive, will be incomplete. The same is true for uses of expressions, too. There are unlimited ways in which an expression can be used. But if this is the case, why did Wittgenstein admonish us to look at how an expression is used if we want to know its meaning? He did so because he recognized that while there are unlimited ways an expression can be used, by looking at the language games and forms of life it appears in, we can see the conventional ways it is used. Wittgenstein recognized that in ordinary usage when we ask for the definition of an expression we are primarily interested in how that expression is used in a particular situation or activity. We are not interested in every possible usage of the expression. The number of possible meanings or usages of an expression will always be unlimited, but by narrowing our examination of usage to particular activities we will focus our attention on those usages most important to the situation at hand.

Now we can see more clearly what is meant by understanding textual descriptions by looking at their use. We are not, as Wittgenstein warned us, equating use with meaning. We look at use in order to orient ourselves in the right direction, for if we look at how textual descriptions are (or can be) used we must examine the activities in which they occur: The activities of authorship, research, publishing, providing access to information, giving advice, etc. We stated that "What is important . . . is not just the uses of an expression, but the conventional uses of that expression in relation to some situation or task at hand." What makes the use of expressions conventional and, hence, understandable, is not its repeated or frequent usage, but its relationship to an activity that evinces recurrent or

repeatable patterns. The use of a textual description is conventional, and understandable as conventional, only insofar as the activity in which it is embedded is conventional. For example, an author's name can only be used to connote quality research if, and only if, there exists certain predictable patterns in his work and his field of study which make such a usage of his name possible. Wittgenstein's notion of "forms of life" should help us to understand the sense in which language may be said to be conventional.

It is important to point out that a conventional usage of an expression does not necessarily imply that the expression has been used that way before. It is entirely consistent with the view of language we have developed to have a usage of an expression which is new, yet conventional and understandable without explanation. This is possible because what needs to be understood are the conditions under which this type of expression is used (the "language game"), and the activity in which the expression is used or refers to (the "form of life"). Thus, an inquirer can ask a colleague for information, relevant to a particular problem, and his colleague could answer, "The early conference proceedings of the XYZ Society," or "So-and-so's work," or "The large red book in the New Acquisitions shelf on the library." What permits these expressions to be understandable is not that they are frequently uttered in such circumstances, but that there are regular and predictable patterns in the activities that generate or use the desired information that make such uses of language understandable (of course, the inquirer's colleague could be wrong or misleading, but here we will assume he is not). Such regularities or patterns make these expressions reasonable responses to the inquirer's request. Consider the inquirer's puzzlement if he is asked where he might find certain textual information and his

colleague told him to look under a rock in the garden.

Language, Meaning and Full-Text Retrieval

Meaning in natural language, as we have seen, is based on how that language is used, and language use, in turn, is embedded within the context of forms of life--conventional, repeated human activities. Thus, the "meaning" of a word or phrase cannot be clarified without referring to the activity in which it plays a role. As Wittgenstein remarked, " . . . we don't start from certain words, but from certain occasions or activities." [1972, p. 3] We have a naive prejudice that words have an underlying, essential meaning which they carry with them in all contexts and through all the situations in which they are used. Wittgenstein showed that this is not the case. " . . . understanding language is not a matter of grasping some inner essence of meaning, but, rather, of knowing how to do certain things," and, " . . . of course, what confuses us is the uniform appearance of words when we hear them spoken or meet them in script and print." [1953, p. 36 and para. 11, resp.] A word does not have an underlying essential meaning, but, rather, a "family" of more or less similar meanings which can be distin- guished or clarified by examining the different activities ("forms of life") in which the word is used. Traditionally, linguistics has insisted that to understand the usage of a particular word in an activity, one must first understand the "essential meaning" of the word and then look to see how it is used. Wittgenstein reversed this, saying that the activity must be understood first before the "meaning" of the word could be understood. Consider the word "pitch." What "essential meaning" of "pitch" could possibly tell you what someone meant by the following sentences?

"They pitched their tent."
"They all pitched in."
"He covered his roof with pitch."
"The armies fought a pitched battle."
"He sang in a much higher pitch than before."
"Nicklaus pitched the ball within two feet of the cup."
"The sales pitch was very convincing."
"The falcon soared to a high pitch."
"The roof was sharply pitched."
"They worked at a feverous pitch."
"The game was marked by a number of successful pitch-outs."

Clearly, the activities of camping, working together, roof repair, warfare, singing, golf, etc., must be understood before the meaning of the word "pitch" can be understood in these sentences. No uniform, essential meaning of the word "pitch" runs through all the uses of the word shown above, and some uses of the word are so disparate that there appears to be no similarity at all. Consequently, we can say that the number of distinguishable meanings/uses of a word is equal to the number of different activities it is used in (or used to describe). This brings us to the major problem of full-text retrieval, namely, that the full-text of documents stored on-line, though they often contain similar sets of content-words, are often written to support largely different activities. As a consequence, the same words or phrases may have radically different uses or meanings in documents which deal with different activities. But even documents which purport to deal with, or result from the same activity and discuss the same topic may not do so using the same vocabulary. Natural language is an incredibly rich and diverse medium of expression which permits individuals to use an unpredictable variety of words to express the same ideas or discuss the same topic. These two phenomena militate against effective retrieval using a full-text system. Examples of how these characteristics of natural language reduce retrieval effectiveness were evident in this experiment.

Problems in Formulating Queries

One particular issue that was important to the lawyers who used the data base concerned an accident that had occurred and was now an object of litigation. The lawyers wanted all the reports, correspondence, memoranda, and minutes of meetings which discussed this accident. Formal queries were constructed which contained the word "accident(s)" along with several relevant proper nouns. Later in our search for unretrieved relevant documents we found that the accident was not always referred to as an "accident", but as an "incident", "event", "situation", "problem", or "difficulty" often without mentioning any of the proper names involved (because they were obvious to those discussing the issue). The manner in which an individual referred to the accident was frequently dependent on his point of view (and this, of course, reflects the kind of activity the individual is involved in). Those who discussed the event in a critical or accusatory way referred to it quite directly--they called it an "accident". But those individuals who were personally involved in the event (and, perhaps, culpable) tended to refer to it obliquely or euphamistically. It was they who referred to the accident as, inter alia, an "unfortunate situation", or a "difficulty". But these were not all the terms which were used on relevant unretrieved documents. Sometimes the accident was referred to obliquely as "the subject of your last letter", or "what happened last week was an unfortunate . . .", or, as the opening lines of the minutes of a meeting discussing the issue began "Mr. A: We all know why we're here . . .". Sometimes relevant documents dealt with the problem by only actually mentioning the technical aspects of why the accient occurred, and not mentioning the accident itself or the proper names involved. In addition, much relevant information

discussed the situation <u>prior</u> to the accident, and, naturally, contained no reference to the accident itself.

Another information request identified three key terms or phrases that were used to retrieve relevant information, but later we were able to find 26 other words and phrases which retrieved additional relevant documents. The three original key terms could not have been used individually because they would have retrieved 420 documents, or approximately 4,000 pages of hard copy, an unreasonably large retrieved set most of which contained irrelevant information. Another information request identified four key terms/phrases that were used to retrieve relevant documents, but later we were able to use 44 additional terms and combinations of terms to retrieve relevant documents that had been originally missed.

Sometimes we could follow a trail of linguistic creativity through the data base. In one example, one of the key phrases was "trap correction". This, of course, was used to retrieve relevant documents, but later we dis-covered that relevant, unretrieved documents had discussed the same issue but referred to it as the "wire warp". We continued our search and found that in other documents this same thing was referred to in a third way: The "shunt correction system". Further, we discovered that the inventor of this system was a man called "Coxwell". This directed us to some documents he had authored discussing this system, only he referred to it as the "Roman circle method". Using this phrase as a formal query we discovered still more relevant unretrieved documents. But this wasn't the end. Further searching revealed that this system had been tested in another city, and all documents germane to those tests referred to the system as the "air truck". At this point our search ended (having taken over an entire 40 hour week of

on-line searching), but there is no reason to believe that we had reached the end of the trail. We simply ran out of time.

Since the database included many items of personal correspondence and the verbatim minutes of meetings, the use of slang frequently changed the way in which one would "normally" talk about a subject. Disabled or malfunctioning mechanisms with which the lawsuit was concerned were sometimes referred to as "sick" or "dead", and a burned-out circuit was referred to as being "fried". A critical issue was sometimes referred to as the "smoking gun". Even misspellings proved an obstacle to effective retrieval. Key search terms (which were essential parts of phrases) such as "flattening", "gauge", "memos", and "correspondence", were used in formal queries to retrieve relevant documents. But we were also able to retrieve relevant documents using the same phrases but with the search terms spelled "flatening", "guage", "gage", "memoes", and "correspondance", respectively. Such misspellings are tolerable in normal everyday correspondence, but when included in a computerized database they become literal traps for inquirers who must not only anticipate the key words and phrases which might be used to discuss an issue, but also all the possible misspellings, letter transpositions, and typographical errors which might be made in using those key words and phrases (and we make no claim to having anticipated all the possible errors).

Some of the information requests placed almost impossible demands on the ingenuity of the individual who constructed the formal query. In one situation, the lawyer wanted "Company A's comments concerning . . .". Just looking at the documents authored by Company A was not enough. Many relevant documents were not retrieved initially because these comments were

embedded in the minutes of meetings or recorded second-hand in the documents authored by others. Merely retrieving all the documents in which Company A was mentioned was too broad a search. It retrieved over 5,000 documents (about 40,000+ pages of hard copy). But predicting the exact phraseology of the text in which Company A commented on the issue was almost impossible. Examples which occurred in unretrieved relevant documents included "Co. A agreed to consider", "Co. A. said", or "Co. A pointed out that". Sometimes Company A was not even mentioned, it was merely noted that So-and-so (who represented Company A) "said/considered/remarked/ pointed out/commented/noted/explained/discussed", etc.

In some information requests the most important terms and phrases were not used at all on relevant documents. For example, "steel quantity" was a key phrase used to retrieve important relevant documents germane to an actionable issue. But unretrieved relevant documents were found which did not report steel quantity at all, but merely recorded the number of such things are "girders", "beams", "frames", "bracings", etc. In another request it was important to find documents which discussed "non-expendable components". Here, relevant unretrieved documents merely listed the names of the components (of which there were hundreds) and made no mention of the broader generic description of these items as "non-expendable".

These examples are only a few of the myriad linguistic problems which confronted the inquirers who had to use STAIRS to search for relevant textual information. The task was an impossibly difficult one, due to the unlimited and unpredictable way in which individuals can talk about a particular subject.

Discussion

The results of the STAIRS evaluation demonstrate that the number of different words and phrases which can be used to talk about a particular topic is unpredictably varied, and our discussion of theories of meaning in language argue convincingly that this variety and flexibility in natural language will be found in any significantly large natural language text. Two phenomena contribute to this variety of expression: 1. The inherent flexibility and creativity of linguistic expression; and, 2. the variety of activities in which authors of documents are engaged and from which those authors' expressions and descriptions are derived.

How can the effects of these two linguistic phenomena be lessened in the retrieval of documents? First, the variety of expression in language must be artificially limited by the introduction of a normative descriptive language that can be used to describe documents for retrieval. Second, the number of activities which underly the linguistic expressions used for searching must be minimized or reduced.

For information retrieval, the most direct way in which to limit the variety of expression in the searching vocabulary is to replace the full-text searching with a retrieval system based on manual indexing using a controlled vocabulary. A controlled vocabulary would mandate that only one word or phrase would be used to represent a particular subject. For example, all documents which indexers believed to be concerned with the accident which we discussed in our previous example could be re presented on the data base using the term "accident", and no other. So, regardless of whether the author of a relevant document referred to this topic as, inter alia, an "incident", "event", "situation", "problem", or "difficulty", the

document would be retrievable by submitting a search query with only the term "accident."

The second need (to reduce the number of activities which underlie the linguistic expressions used for searching) can be effected by replacing the full-text retrieval procedures with a manually-indexed retrieval system. By having a group of indexers describe the subject of the documents in the data base we insure that the meaning of the searching vocabulary will be derived from only one activity--the process of indexing (i.e., describing the subject content of the documents). Since all the indexers would be engaged in the same activity, it insures that the same subject descriptions would be used in a minimal variety of ways.

Bar-Hillel, Yehoshua. "Language," in Scientific Thought, Mouton/Unesco,
     Paris, 1972.

Bar-Hillel, Yehoshua. "Theoretical Aspects of the Mechanization of
     Literature Searching," in his Language and Information; Selected
     Essays on their Theory and Application, Addison-Wesley, London,
     1964.

Barthes, Roland. Mythologies, Hill and Wang, New York, 1972.

Barthes, Roland. Elements of Semiology, Hill and Wang, New York, 1968.

Black, Max. The Labyrinth of Language, Frederick A. Praeger, New York,
     1968.

Blair, David. "The Data Document Distinction in Information Retrieval."
     Communications of the ACM, in press.

Blair, David and M. E. Maron. "A Study of Retrieval Effectiveness for a
     Full-Text Document Retrieval System." unpublished.

Cavell, Stanley. The Claim to Rationality, unpublished dissertation,
     Harvard University.

Chomsky, Noam. Aspects of the Theory of Syntax, M. I. T. Press, Cambridge,
     Mass., 1965.

Chomsky, Noam. Syntactic Structures, Mouton, The Hague, 1957.

Eco, Umberto. A Theory of Semiotics, Indiana University Press, Bloomington,
     Indiana, 1976.

Galbraith, John Kenneth. The New Industrial State.

Guiraud, Pierre. Semiology, Routlage, London, 1975.

Hawkes, _____. Structuralism and Semiotics, University of California
     Press, Berkeley, 1977.

Hjelmslev. Prolegomena to a Theory of Language, University of Wisconsin Press,
     Madison, Wisconsin, 1969.

Jakobson, Roman. "Language in Relation to Other Communication Systems,"
     in his Selected Writings, Vol. II, Mouton, The Hague, 1971, pp. 697-708.

Katz, Jerold J., and Jerry A. Fodor. "The Structure of Semantic Theory,"
     Language: V. 39, 1963.

Lakoff, George. Irregularity in Syntax.

Locke, John. "An Essay Concerning Human Understanding," in British Empirical
     Philosophers, A. J. Ayer and Raymond Winch (eds.), Routledge and Kegan
     Paul, Ltd., London, 1965.

Luigi, Romeo. "Hericlitus and the Foundations of Semiotics," VS, 15,
    Sept.-Dec., 1976.

Malcolm, Norman. Ludwig Wittgenstein: a Memoir, Oxford University Press,
    London, 1972.

Peirce, Charles Sanders. "Logic as Semiotic: the Theory of Signs," in
    Philosophical Writings of Peirce, Justus Buchler (ed.), Dover
    Publications, New York, 1955.

Peirce, C.S. Collected Papers, Harvard University Press, Cambridge, 1931.

Pitkin, Hanna F. Wittgenstein and Justice, University of California Press,
    Berkeley, 1972.

Saussure, Ferdinand de. Cours de linguistique generale, Payot, Paris, 1916.
    English translation: Course in General Linguistics, Philosophical
    Library, New York, 1959.

Schneider, David. American Kinship: A Cultural Account, Prentice-Hall,
    New York, 1968.

Sebeok, Thomas. The Tell-Tale Sign-A Survey of Semiotics.

Wittgenstein, Ludwig. Blue and Brown Books, Harper and Brothers, New York,
    1958.

Wittgenstein, Ludwig. Lectures and Conversations on Aesthetics, Psychology,
    and Religious Belief, Cyril Barrett, ed., University of California Press,
    Berkeley, 1972.

Wittgenstein, Ludwig. Philosophical Investigations, The MacMillan Co., New
    York, 1953.

Wittgenstein, Ludwig. "Notes for Lectures on 'Private Experience' and 'Sense
    Data'," in Morick (ed.), Introduction to the Philosophy of Mind, Scott,
    Foresman and Co., Chicago, 1970.

Zabeeh, Farhang. "On Language Games and Forms of Life," in E. D. Kenke (ed.),
    Essays on Wittgenstein, University of Illinois Press, Chicago, 1971.

Ziff, Paul. Semantic Analysis, Cornell University Press, Ithica, 1960.

Zipf, George Kingsley. Human Behavior and the Principle of Least Effort, Hafner
    Publishing Company, New York, 1965.