ISI Reprint Series
ISI/RS-83-8
October 1983

12

University of Southern California

William C. Mann

Inquiry Semantics: A Functional Semantics of Natural Language Grammar

INFORMATION SCIENCES INSTITUTE

213/822-1511
4676 Admiralty Way/Marina del Rey/California 90292

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER ISI/RS-83-8	2. GOVT ACCESSION NO. A13515	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Inquiry Semantics: A Functional Semantics of Natural Language Grammar		5. TYPE OF REPORT & PERIOD COVERED Research Report 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) William C. Mann		8. CONTRACT OR GRANT NUMBER(s) F49620-79-C-0181
9. PERFORMING ORGANIZATION NAME AND ADDRESS USC/Information Sciences Institute 4676 Admiralty Way Marina del Rey, CA 90292		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research Building 410, Bolling Air Force Base Washington, DC 20332		12. REPORT DATE October 1983 13. NUMBER OF PAGES 21
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

This document is approved for public release; distribution is unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

.....

18. SUPPLEMENTARY NOTES

This report replaces ISI/RR-82-106, *Generating Text: The Grammar's Demands*, which was not issued.

This report is a reprint of a paper that appears in the proceedings of the First Annual Conference of the Association for Computational Linguistics, held in Pisa, Italy, in September 1983.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

artificial intelligence, choice experts, computer generation of text, inquiry semantics, knowledge representation methods, natural language, Nigel, Penman, semantics of English, systemic grammar, text generation

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

(OVER)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. ABSTRACT (continued)

Programming a computer to operate to a significant degree as an author is a challenging research task. The creation of fluent multiparagraph text is a complex process because knowledge must be expressed in linguistic forms at several levels of organization, including paragraphs, sentences and words, each of which involves its own kinds of complexity. Accommodating this natural complexity is a difficult design problem. To solve it we must separate the various relevant kinds of knowledge into nearly independent collections, factoring the problem.

Inquiry semantics is a new factoring of the text generation problem. It is novel in that it provides a distinct semantics for the grammar, independent of world knowledge, discourse knowledge, text plans and the lexicon, but appropriately linked to each. It has been implemented as part of the Nigel text generation grammar of English.

This paper characterizes inquiry semantics, shows how it factors text generation, and describes its exemplification in Nigel. The resulting description of inquiries for English has three dimensions: the varieties of operations on information, the varieties of information operated upon, and the subject matter of the operations. The definition framework for inquiries involves both traditional and nontraditional linguistic abstractions, spanning the knowledge to be represented and the plans required for presenting it.

ISI Reprint Series
ISI/RS-83-8
October 1983

University of Southern California

William C. Mann

Inquiry Semantics: A Functional Semantics of Natural Language Grammar

INFORMATION SCIENCES INSTITUTE

213/822-1511
4676 Admiralty Way/Marina del Rey/California 90292

This research was supported by the Air Force Office of Scientific Research contract No. F49620-79-C-0181. The views and conclusions contained in this document are those of the author and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Air Force Office of Scientific Research of the U.S. Government.

ISI Reprint Series

This report is one in a series of reprints of articles written by ISI research staff and published in professional journals and conference proceedings. For a complete list of ISI reports, write to

Document Distribution
USC/Information Sciences Institute
4676 Admiralty Way
Marina del Rey, CA 90291
USA

Contents

1. Introduction

Text generation is ***the generation of language to conform to an a priori intention and plan to communicate.***[1] The problem of text generation is naturally complex, requiring the active coordination of many kinds of knowledge having independent origins and character. A significant part of this complexity is in grammatical knowledge. It is important for the grammar of a text generator to have its own integrity, yet without being operationally autonomous.

The methods of generating text presented here grew out of a concern to maintain the integrity and definitional independence of particular existing fragments of grammar. These methods employ the grammar in ways which do not make any strong assumptions about the nongrammatical kinds of knowledge in the text generator. They control the use of the grammar in generation.

We first describe the methods, showing how they make grammatical generation possible. Then we show how they factor the problem of text generation and clarify the role of knowledge representations. Finally we characterize inquiry semantics and the notion of meaning.

2. Grammar and Control

People often anticipate that a text generator will plan the operations of the grammar in full detail and then execute such plans. In fact, such a mode of operation has serious difficulties, and so it is worthwhile to consider other approaches. Even given the definition of a grammar and a particular way of manipulating it to produce text, there is an issue of where the initiative should be exercised in generation. Should the responsibility for conformity of the result to the given intention and plan lie within the grammar manipulator, i.e., be part of its process of employing the grammar, or are the details of grammar use preplanned? It is an issue of control.

To see the problem more clearly we can compare controlling the grammar to steering a car.

> If we intend to drive to a nearby store, we can imagine planning the trip (in terms of steering motions) in total detail, deciding just where to turn, change lanes, and so forth, with sufficient precision to insure success. This detailed plan could in principle then be used to steer the car to the store. Such methods of ***imposed control*** are practical only in very simple cases.
>
> Alternatively, we can make the decisions about steering ***at the point of need,*** on demand. Unanticipated conditions are thus allowed for, and the complexity of the task is reduced. (There is no need to compensate in the plan for tire pressures, for example.) At each significant point along the way, the driver chooses a direction that conforms to the goal of reaching the destination. This is an ***active conformity*** approach, in which decisions about direction are made while the trip is in progress.

---

[1] This role of intention in the use of language is one of the reasons for calling the semantics in this paper a *functional* semantics. Another is our use of one of the "functional" linguistic traditions.

This report replaces ISI/RR-82-106, *Generating Text: The Grammar's Demands*, which was not issued.
This report is a reprint of a paper that appears in the proceedings of the First Annual Conference of the Association for Computational Linguistics, held in Pisa, Italy, in September 1983.

With imposed control, information about how to satisfy the intention and plan is needed before the process is started. With active conformity, information is needed as the process proceeds.

The design of our generation methods is based on active conformity. The grammar demands the information it needs about the plan as generation proceeds.

What does a purposefully generating grammar need to know? As part of the development of the Penman text-generation program, we have created a large systemic grammar of English [Mann 83]. Penman is designed to create a text plan and then execute it by giving it, one sentential element at a time, to the grammar. The grammar, which is called Nigel, operates on its own initiative, requesting information about the planned text as it is needed. The central organizing concept in the grammar is ***choice***. The language offers a variety of grammatical options that can be represented as sets of alternatives, and means for producing surface forms from particular combinations of choices made among the alternatives. All syntactic options are expressed in the sets of alternatives. In any one set, choosing one option excludes all of the others. Nigel contains over 200 systems (collections of alternatives in systemic notation), along with provisions for realizing choices as structures, an experimental lexicon used to give the structures surface forms, and extensive provisions for experimental control.[2]

Given this orientation toward choice, the problem of conformity to the text plan is simply the problem of making appropriate choices. Each set of alternatives (each ***"system" in its systemic representation*** has an associated ***chooser*** or ***choice expert***, a process that embodies a method for choosing appropriately in any particular circumstance.

The choice experts require certain information as they proceed with text generation. Nigel's choice experts request this information by presenting ***inquiries*** to the ***environment*** (the place outside of the grammar where intentions and plans to communicate are found.) For this purpose, Nigel employs a formal ***inquiry language*** in which an inquiry is an expression containing an ***inquiry operator*** and a sequence of ***operands***. A single ***interface*** is provided for all interactions between Nigel and the environment; all interactions at the interface are in the inquiry language. This way of using such an interface is called ***inquiry semantics***.

In this framework, we can understand the demands of the grammar by understanding the inquiry operators.

3. Varieties of Demands

This section characterizes the demands for information that Nigel can make in generating sentences. Since Nigel demands information only by presenting inquiries, we first characterize the things that Nigel can inquire about (the operands of inquiries), then characterize in two different ways the questions that Nigel can ask.

---

[2]The grammar is written in an extended systemic notation and draws extensively on precedents in the work of Halliday and others [Berry 75, Berry 77, Halliday & Hasan 76, Halliday 76, Hudson 76, Halliday & Martin 81, de Joia & Stenton 80, Fawcett 80]. We gratefully acknowledge the participation of Michael Halliday and Christian Matthiessen in the work.

3.1 Categories of Operands of Inquiries

Nigel has four related information forms:

1. Concept symbols
2. Presentation specifications
3. Term sets
4. Terms

Concept symbols are names assigned by the environment to particular elements of its knowledge, either in the text plan for the text being formed or in the environment's knowledge base. A concept symbol represents an entity that may be simple or complex, decomposable or not; the symbols themselves are not decomposable. A concept symbol does not have to bear any particular relationship to any kind of *linguistic* entity.

Presentation specifications are formal descriptions of the information that should be expressed in a particular reference, description, or predication. Through presentation specifications the environment designates the content to be conveyed in each particular constituent, (but not how the content is to be expressed.)

For nominal groups (NP's), for example, presentation specifications represent the identification of the content to present about the particular object, process, or relation which the nominal group represents. The collection of devices that express nominal group content include head terms (nouns, pronouns, substitute "one"), modifying nominals, adjectives and adjective groups, quantifiers, numerals, determiners, prepositional phrases, restrictive and nonrestrictive relative clauses. Normally the grammar will use some combination of these devices in the nominal group to express all of the content of the presentation specification.

As a minimal example, the grammar's decision on whether a pronoun is adequate as a referring phrase can be made on the basis of the presentation specification, since the specification tells what constitutes adequate reference at the point of referring. (If the presentation specification indicates that nothing beyond gender and number needs to be expressed, a pronoun is used.)

The presentation specification is thus a unifying device for all of the conceptual elements of an intention to refer. It is essential to the generation task because the various syntactic devices effectively compete for the content which the nominal group expresses in referring.

At the clause level, presentation specifications operate comparably, unifying the effects of adverbial, conjunctive, and clausal modifiers. The specifications are constructed units, not frames or delimited regions of knowledge.

Term sets are collections of lexical items created in a special way which insures that they are appropriate, in denotation, connotation, and information content, for their intended use. (The process which creates term sets does not restrict them syntactically; that is done later by the grammar.) The individual terms in a term set need not be so restrictive that they fully express the intent of the unit being constructed, since they are used with modifiers. Term sets are not like sets of synonyms since they do not have any uniformity of semantic content.

Term sets are used as collections of alternatives, from which one term will be picked for the

final syntactic unit. The best example is a term set giving alternatives that can serve as the head term of a nominal group.

A *Term* is a single lexical item selected from a term set. It identifies the particular lexical item to appear in the generated text. Currently Nigel is deliberately underdeveloped in its treatment of lexical items, having no morphological component at all. Hence terms are simply lexical items which bear lexical features that the grammar can employ for selectivity.

To see how these forms are used, consider the sentence:

The leader is John.

It refers to John twice. In generating this sentence, the same concept symbol, say JLDR, would be used to generate both of the references. However, two different presentation specifications for referring to JLDR would be created. The first might specify that the resulting expression should convey the fact that the individual holds the role of leader. The second could merely specify that the resulting expression should convey the person's name.

Two different term sets would also be created. Initially, each would contain conceptually and denotationally appropriate terms, possibly including "leader," "man," and "person," in one of the term sets, and "John," and "Mr. Jones" in the other. Under guidance from various inquiries, the grammar applies different selectivity to one term set than to the other, so that the terms "leader" and "John" are finally selected.

How do these operands of inquiries compare with conventional linguistic abstractions?

Concept symbols have many precedents, and terms are familiar. Both presentation specifications and term sets are new. As we will see, both presentation specifications and term sets are widely and frequently used in the grammar. Their central role in generation suggests that they are worthy of linguistic attention.

Presentation specifications are novel in that they represent the content of particular units without its allocation to constituent units. This permits the investigation of how the allocation works, and in particular how differing ranks compete for representational roles. Competition among the possible constituents of a nominal group for representation of posession seems to be a typical case. We would like to know, for example how the decision between using the determiner "his," the prepositional phrase "of his," and the clause "which he has" is made. A presentation specification can say in a syntactically neutral way that possession is to be expressed. Using them facilitates study of the alternation.

Nigel uses subtractive operations on presentation specifications to account for the fact that repeated expression of content in a nominal group is marked, but single expression is not. So it can account for the perception that "his car, which he owns" is marked in a way in which "his car, which he hates" is not.

Term sets are novel in that they represent the alternations and competition among lexical items. The sets of terms which compete as candidates, e.g. for the main verb of a clause or head term of a nominal group, are highly variable and dependent on the subject matter of the communication.

Hence they are not susceptible to static analysis as part of the grammar, and they are not easy to represent in systemic systems.

Consider, for example, the word "attention" at the end of the third paragraph back. Other candidates for use in the same setting would include words such as "research," "curiosity," "work," "perusal," and "funds." These terms (as well as "attention") would all be in the term set for generating that nominal group. However, they are from different lexical fields, fields which are ordinarily not in alternation. Since they are not the basis of a stable alternation, many sorts of static representations of them (including representation in systems in a systemic lexico-grammar) seem inappropriate. The situation is much more complex and dynamic, worthy of linguistic attention.

Notice that in both cases, addition of a new formal construct will facilitate study of how particular expressions are related to closely related alternatives in ways which are not in opposition in a conventional systemic account. Studies of functional alternation have long been a highly valued activity among systemicists.

Notice also that these constructs arise easily, almost inevitably, in studies of text construction, but are not inevitable at all in descriptive studies of text. Given a particular text to study, it is not at all clear what the rejected head term candidates were, nor what the alternate allocations of content to syntactic units might have been. In systemic terms, part of the meaning of a nominal group is derived from the particular choice of the head term, but, working descriptively, the alternation is hard to characterize. Study of text generation (and related work on constructive characterization) thus complements other methodologies in that it makes certain difficult tasks easier.

3.2 Abstract Categories of Inquiry Operators

The inquiries of the grammar can be differentiated according to categories of purposes they serve. Five such categories are described below. The first two kinds of inquiries are used for control, and the last three extract symbols from the environment -- either lexical items or symbols that can be included as subject matter in subsequent inquiries. Inquiries of the first two kinds have predetermined closed sets of possible responses; the last three kinds allow an unlimited number of responses.

1. ***information availability***
2. ***information characterization***
3. ***decomposition***
4. ***linking (identification of related information)***
5. ***mapping***

Some inquiries determine whether information of a certain character is ***available***, such as the location or duration of an event. These inquiries generally precede others used to characterize information.

The operators used for ***information characterization*** form the largest collection of operators among the five kinds. They are used to subcategorize and also to discover relations of inclusion, identity, precedence, adjacency, and attributes of manner, number, completeness, intended emphasis, identifiability to the reader, decomposability, gender, hypotheticality, extensionality, and many other sorts.

When the grammar has determined that some of the available information is decomposable into parts in a syntactically significant way (usually through information availability inquiries), ***information decomposition*** inquiries are used to obtain access to the parts. This is the largest category of inquiries for which an unlimited diversity of responses is allowed. These inquiries offer access to actors, affected objects, processes, causers, polarities, locations, time periods, extents, manners, and various kinds of participants or conditioners of processes.

The ***linking inquiries*** are a small collection of inquiries which resemble the information decomposition inquiries. They obtain information related in a particular way to known information, but not part of it. For example, given an event whose time must be expressed, there is an inquiry that obtains the identity of the time relative to which the event's time of occurrence should be expressed.

In terms of the four forms of information presented in section 3.1 above, exploration always proceeds from concepts to presentation specifications and term sets, and from term sets to terms, as shown in Figure 3-1.

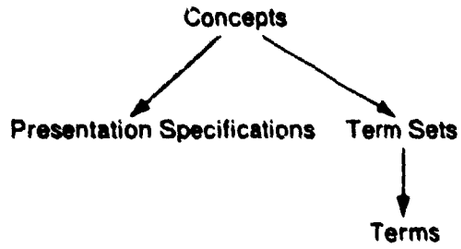

Figure 3-1: Information flow through mapping inquiries

A small collection of ***Mapping*** inquiries participate in this exploration at the points where information forms change. Several create specialized presentation specifications for concepts, and others create term sets and terms.

Since operators can request presentation specifications, they can in effect demand that the environment work out what information to include in a new reference to an entity. The environment must then use the knowledge of past mentions, a model of the hearer's attention and of possible confusion candidates, and also the knowledge of denotationally appropriate lexical items; these elements of knowledge are all outside the boundary of the grammar. The mapping from concepts to presentation specifications is thus dependent on the particular circumstances.

In a similar way, the mappings from concepts to term sets and from term sets to terms also vary depending on the communication situation.

3.3 Categories of Subject Matter

Recurrent topics and categories of subject matter in the inquiries reflect the syntactically encoded categories of knowledge in English. The subject matter categories form two groups:

1. Elements of knowledge that typically exist prior to the intention or plan to communicate (described in section 3.3.0.1 below), and

2. Elements of knowledge created as part of pursuing the intention or plan to communicate (described in section 3.3.2 below.)

These are called the ***Knowledge Base*** and the ***Text Plan***, respectively.

Surprisingly, we do not see any sharing of inquiries between these two kinds of knowledge. In Nigel, we find that each inquiry operator addresses solely one body of knowledge or the other. A few of the categories of operations address both kinds of knowledge, notably inquiries about availability of information. Within the categories, however, each individual inquiry is specialized to a single kind of knowledge.

3.3.1 Subject Matter of Inquiries Concerning Prior Knowledge

In addition to inquiring about availability of information, the grammar asks about abstract characteristics of processes, about number and discreteness, and about time and space. Also, there is a substantial collection of inquiries about logical relations such as set membership, interval inclusion, identity of two entities, extensionality, definiteness of existence, hypotheticality, polarity and conditionality.

3.3.2 Subject Matter of Inquiries for Communication

Among the inquiry operators that refer to information created in pursuit of an intention or plan to communicate, there are inquiries about speech acts and about controlling the hearer's attention. The latter are used in controlling thematicity, various kinds of marking, and the foregrounding or backgrounding of information.

3.4 Support Processes in the Environment

The organization of inquiry requires that various kinds of processes be available in the environment for responding to inquiries. At a detailed level, there must be a capability for the environment to recognize each inquiry operator and to respond to each one appropriately. In computational terms, for a particular domain of expressive problems, all of the inquiry operators which are called upon to serve that domain must be implemented. (For simple expressive problems this can be far fewer than the total for the grammar.)

At a more comprehensive level, we can identify certain recurrent activities which must underlie the operations of the inquiry operator implementations. These include searching for an appropriate set of lexical items (such as candidate head nouns for a nominal group), creating a presentation specification for expressing a particular idea, and choosing among a set of terms which the grammar has approved as appropriate for a certain use.

At an even more comprehensive level, the grammar relies on the prior activity of processes which plan the text.

4. Inquiries in Action: An Example

The following list summarizes Nigel's activity in developing a particular nominal group: "her appointment on Wednesday morning with us." The starting point is identification of a need to refer to an object represented by concept APPOINTMENT. At the end of the activity shown, there is a structure containing the word "appointment" as the head term, the word "her" as its determiner, and elements that could be further developed into the phrases "on Wednesday morning" and "with us." The category of each inquiry operator is indicated in <brackets>. The order of presentation is the order actually used in the program. It is somewhat disconnected, since the program often chooses in an arbitrary way between several things which it could do next. An inquiry appears more than once if it is used by more than one choice expert.

1. Obtain a presentation specification for APPOINTMENT <mapping>

developing the head term of the group

2. Obtain a set of candidate head terms <mapping>
3. Establish that APPOINTMENT is countable <characterization>
4. Classify APPOINTMENT as extensional <characterization>
5. Classify APPOINTMENT as unitary <characterization>
6. Classify APPOINTMENT as not a question variable <characterization>
7. Classify APPOINTMENT as extensional (as part of pronoun control) <characterization>
8. Classify APPOINTMENT as unitary (as part of pronoun control) <characterization>
9. Establish that the gender of APPOINTMENT is known <availability>
10. Establish that in the presentation specification of APPOINTMENT, there is more to be expressed than gender and number <characterization>
11. Determine that it is preferable to exclude proper nouns from the term set, rather than exclude the remainder <preference>

begin developing the determiner

12. Establish that APPOINTMENT is extensional (for determiner control) <characterization>
13. Establish that APPOINTMENT is identifiable to the reader <characterization>

resume developing the head term

14. Have the environment select a term, here "appointment," from among the terms that survived syntactic selectivity <mapping>

developing the modifiers of the head term

15. Establish that the presentation specification for APPOINTMENT does not indicate that color, location, use, substance, size, place of origin or age should be expressed (7 inquiries) <characterization>

developing the accompaniment modifier

16. Establish that some kind of accompaniment of APPOINTMENT should be expressed <characterization>
17. Obtain a symbol (WITHUS) representing the accompaniment knowledge to be expressed <decomposition>

complete development of the head term

18. Determine that the speaker wants the hearer to pay more than minimal attention to APPOINTMENT (thus cutting off further investigation of a substitution of "one" for "appointment") <characterization>

developing the time period modifier

19. Establish that the presentation specification of APPOINTMENT indicates that a time constraint should be expressed <characterization>
20. Obtain a symbol (ONWEDNESDAYMORN) representing the time constraint to be expressed <decomposition>

resume developing the determiner

21. Establish that no information about the proximity of APPOINTMENT should be expressed <characterization>
22. Establish that information about the possessor of APPOINTMENT should be expressed <characterization>
23. Obtain a symbol (JANE) representing the possessor of APPOINTMENT <decomposition>
24. Establish that JANE is unitary <characterization>
25. Establish that JANE does not represent a question variable <characterization>
26. Obtain a symbol (SELF) representing the speaker <decomposition>
27. Obtain a symbol (PUBLIC) representing the hearer of the entire nominal group <decomposition>
28. Establish that SELF is not identical with or included in JANE <characterization>
29. Establish that PUBLIC is not identical with or included in JANE <characterization>
30. Establish that the gender of JANE is known <availability>
31. Establish that the gender of JANE is female <characterization>

finish developing the modifiers

32. Establish that there is no residue of unexpressed content in the presentation specification <characterization>

Using the answers to these inquiries, the grammar builds a structure consisting of four elements in an ordered sequence:

"her," "appointment," ONWEDNESDAYMORN, WITHUS,

the latter two representing conceptual elements to be further developed in subsequent applications of the grammar.

5. Relations between Operators

Some operators are closely related in ways not suggested above. In particular, some pairs of operators are used together in a characteristic way: First an availability operator asks if certain information is available, for example, whether the location of an event is known. If a positive response is given, a decomposition inquiry asks for a symbol to represent the available information, such as the location.

Almost all of the decomposition inquiries are paired with availability inquiries in this way.

However, a few are not. For these, the grammar assumes the existence and separability of the information it requests. The following are the exception cases:

1. the identity of the speaker.
2. the identity of the time of speaking, the "now" of tense.
3. given an event to express in an independent clause, the identity of the time of occurrence of the event.
4. given the need to generate a clause, the identity of the process portion (which will be realized in the main verb.)

In addition, none of the mapping operators and none of the linking operators are paired. We see that the decomposition operators have little intellectual content, but the other kinds all contribute significantly.

6. Demands on the Knowledge Representation

Reviewing the inquiries, we can find several kinds of operations that are particularly difficult to support in explicit knowledge representations such as those currently used in AI or logic.

One operator asks whether the existence of a particular entity is hypothetical. Knowledge gained from this inquiry is useful in controlling contrasts such as the following:

If they run to town, they will be sorry.

If they are running to town, they will be sorry.

Another operator asks about conjectural existence. It controls contrasts such as:

They will run to town.

They might run to town.

In the first case the running to town is treated as definite but occurring in the future.

Another asks whether an action to be expressed is habitual and recurrent rather than a particular instance. Another group of inquiries seeks to determine the manner of performance of an action. Others deal with partial specifications and "question variables" of the sort that are often realized by "wh" terms such as "what," "how," and "whether." Some operators control negation and quantification, which often cause representation problems.

In addition to all of these potential problem sources, associated with inquiries whose responses will be difficult to determine, there are also many difficulties which do not arise from difficulties of representation. For example, knowing what to thematize and what to mark, knowing causes and beneficiaries, knowing which of several lexical items to use (after passing all syntactic and semantic tests), knowing what relations can be expressed as possession, knowing whether the

reader is able to identify an entity in memory (for definite determination), discriminating near from far, all present difficulties without appearing to stress the capabilities of modern knowledge representations.

Thus the inquiries can be used as an indication of what sorts of expansion a knowledge representation needs and as a guide to the ways in which current knowledge of discourse is inadequate to support text generation programming.

7. Factoring the Text Generation Problem

Inquiry semantics separates the problem of designing a text generator into parts which seem much more approachable than the problem as a whole. The grammar is separated from the environment by a tight interface which does not allow the grammar to access any elements of the environment directly. The inquiries are defined in a syntactically neutral or pre-syntactic form; answering them never requires knowledge of the syntax of the language being generated. As a result, the environment and the grammar can develop independently. This is particularly important today, since the technologies of the environment are very unstable, and we would like to be able to use a grammar in conjunction with several styles of knowledge representation.

The environment is divided into the Knowledge Base and Text Plan parts, an informal but potentially very useful distinction. It tends to facilitate independent development of discourse planning methods. Truth-functional issues seem to be related largely to the Knowledge Base.

The treatment of the lexicon separates a variety of lexical phenomena in separate, controlled ways: denotational appropriateness, syntactic features, and nonsyntactic nondenotational attributes such as frequency and register, each receive distinctive treatment in Nigel.

8. The Abstract Character of Inquiry Semantics

In this section we compare inquiry semantics to other kinds of semantics, and also identify the nature of meaning in this framework.

8.1 Comparative Semantics

The inquiry-based semantics presented here contrasts with other accounts also called "semantics" in many ways, but it does not particularly compete with them. This semantics, as a way of theorizing, is an answer to the question "How can we characterize the circumstances under which it is appropriate to make each particular grammatical choice of a language?"

It differs from other semantic approaches in that

1. its scope is confined to grammar, rather than addressing linguistic behavior as a whole;
2. it does not presume particular structures (deep or otherwise) in the environment;
3. it is not particularly limited to issues reducible to questions of truth value;

4. its scope includes nondeclarative, noninterrogative speech actions (including imperative, imprecation, and greeting functions) on a par with declarative and interrogative ones;
5. it includes other functions of language in addition to the representational ones (such as the attention-direction functions);
6. it is defined relative to generation rather than interpretation, but is not thereby "generative".

This semantics is potentially compatible with other sorts, since it makes very few theoretical assumptions about the nature of language and communication. By encompassing every kind of syntactic construction, it is more inclusive than most.

Nothing in inquiry semantics rules out any particular formal apparatus as the notation for the methods by which the environment responds to inquiries. Accounts of particular languages and grammars will give some informal guidance as to which sorts of methods will be perspicuous, and may rule out particular formalisms as response mechanisms for particular grammars. The topic is as yet unexplored.

8.2 The Nature of Meaning in Inquiry Semantics

We could assign meanings to any of several kinds of entities in this framework: grammatical features, collections of features, realizations of collections of features (i.e., structures), inquiry responses--or other possibilities. Our selection of a particular kind of entity as the locus of meaning depends on our intended use for that locus. We intend to use this notion of meaning to identify the ways in which minimal structurally-justified distinctives are responsive to their conditions of use. This selection does not preclude other selections for other purposes, and it certainly does not suggest that there are no other entities which are meaningful.

We associate meanings with grammatical features, in part because these are the controlling entities in the systemic framework. Given a systemic grammar, the syntactic structures which are produced depend entirely on the grammatical features which are chosen, and the opportunity to choose a grammatical feature also depends entirely on the grammatical features which are chosen, i.e., the entry conditions of the system in which the feature occurs. So it is convenient to associate meaning with features, and to derive meanings for any other entity by the determinate derivational methods which the systemic framework provides.

To state the meaning of a grammatical feature is to state the technical circumstances under which the feature is chosen. We identify these circumstances as the set of possible collections of inquiry responses which are sufficient to lead to the choice of the feature. The definitions of the systems of the grammar and their choice experts are thus sufficient to determine the meaning of every grammatical feature.[3,4] Ambiguity of a feature arises when there is more than one collection of relevant inquiry responses which leads to the choice of the feature.

---

[3] We do not state the method here, since that involves many systemic details, but it is normally a rather straightforward matter for the Nigel grammar. More detail can be found in [Mann 82, Mann & Matthiessen 83a, Mann & Matthiessen 83b].

[4] The meanings of the features are not sufficient to find the sets of meanings which correspond to particular structures, since that requires the realization mapping of features to structures. However, given the associations of features with realization operations, the structures for which a particular feature (or combination of features) is chosen can be identified, and so in principle the sets of techincal circumstances which can yield a particular string can be identified.

Differences of meaning reflect differences between collections of inquiry responses. In Nigel, for the features Singular and Plural, one of the collections of inquiry responses which leads to Singular contains a response "unitary" to MultiplicityQ, and a corresponding collection contains "multiple" as a response to MultiplicityQ, which leads to Plural. We can determine by inspection of the entire meanings that Singular and Plural exclude each other, and the determination could be made even if the features were not in direct opposition in the grammar.

Notice that this approach is compatible with approaches to grammar other than traditional systemic grammar, provided that their optionality is reexpressed as alternation of features, with choice experts defined to identify the circumstances under which each option is chosen.

Notice also that it is possible to have meanings in the grammar which are ruled out by the environment, for example, by consistency conditions. A change in the environment's epistemology could lead to changes in how the grammar is employed, without changes in meaning, the grammar being more neutral than its user.

Notice also that the collection of inquiry operators for a language is a claim concerning the semantic range of the grammar of that language, a characterization of what can be expressed syntactically.

Notice finally that, given a grammar and an inquiry semantics of each of two different languages, the question of whether a particular sentence of one language has the same meaning as a particular sentence of the other language is an addressable question, and that it is possible in principle to find cases for which the meanings are the same. One can also investigate the extent to which a particular opposition in one language is an exact translation of an opposition in another.

9. Conclusions

The inquiry language as a level of abstraction provides a useful factoring of the text generation problem, isolating the grammar-intensive part.

Development of inquiry language has led to the creation of new kinds of abstract elements that can be the operands of inquiries. Of these, presentation specifications and term sets have sufficiently novel scopes to suggest that they may be useful in defining relationships between grammar and language use.

We have identified three dimensions of characterization that yield a convenient abstract structure for understanding inquiry language collectively (by categories of operands, categories of operators and categories of subject matter.) These categorizations clarify the ways in which effective use of a grammar depends on processes and information outside of the grammar, including some ways which are not well controlled in available knowledge representations.

Inquiry semantics contrasts with other theoretical entities also called "semantics" in many ways. It is potentially compatible with some other forms, but tends to be broader than many in including non-representational functions and non-declarative speech actions in its scope.

References

[Berry 75] Berry, M., *Introduction to Systemic Linguistics: Structures and Systems*, B. T. Batsford, Ltd., London, 1975.

[Berry 77] Berry, M., *Introduction to Systemic Linguistics: Levels and Links*, B. T. Batsford, Ltd., London, 1977.

[de Joia & Stenton 80] de Joia, A., and A. Stenton, *Terms in Systemic Linguistics*, Batsford Academic and Educational, Ltd., London, 1980.

[Fawcett 80] Fawcett, R., *Cognitive Linguistics and Social Interaction*, Julius Groos Verlag and Exeter University Press, 1980.

[Halliday 76] Halliday, M. A. K., *System and Function in Language*, Oxford University Press, London, 1976.

[Halliday & Hasan 76] Halliday, M. A. K., and R. Hasan, *Cohesion in English*, Longman, London, 1976. English Language Series, Title No. 9.

[Halliday & Martin 81] Halliday, M.A.K., and J. R. Martin (eds.), *Readings in Systemic Linguistics*, Batsford, London, 1981.

[Hudson 76] Hudson, R. A., *Arguments for a Non-Transformational Grammar*, University of Chicago Press, Chicago, 1976.

[Mann 82] Mann, W. C., *The Anatomy of a Systemic Choice*, USC/Information Sciences Institute, Marina del Rey, CA, Technical Report RR-82-104, October 1982. To appear in Discourse Processes.

[Mann 83] Mann, William C., *An Overview of the Penman Text Generation System*, USC Information Sciences Institute, Marina del Rey, CA 90291., Technical Report RR-83-114, 1983. To appear in the 1983 AAAI Proceedings.

[Mann & Matthiessen 83a] Mann, W. C., and C. M. I. M. Matthiessen, *Nigel: A Systemic Grammar for Text Generation*, USC/Information Sciences Institute, RR-83-105, February 1983. The papers in this report will also appear in a forthcoming volume of the *Advances in Discourse Processes Series*, R. Freedle (ed.): *Systemic Perspectives on Discourse: Selected Theoretical Papers from the 9th International Systemic Workshop* to be published by Ablex.

[Mann & Matthiessen 83b] Mann, William C. and Christian M. I. M. Matthiessen, *An Overview of the Nigel Text Generation Grammar*, USC Information Sciences Institute, Marina del Rey, CA 90291., Technical Report RR-83-113, 1983.