American Journal of Computational Linguistics *Microfiche 38*

VERB PARADIGMS FOR

S E N T E N C E R E C O G N I T I O N

MARIANNE CELCE-MURCIA

DEPARTMENT OF ENGLISH
UNIVERSITY OF CALIFORNIA
LOS ANGELES 90024

Summary of "Verb Paradigms for Sentence Recognition"

by

Marianne Celce-Murcia

Department of English, UCLA

This paper describes a linguistically-based recognition grammar model, which was developed as part of a Computer-Aided Instruction Project, to the tasks of recognizing and analyzing a variety of basic sentence types in English. Ways of extending the model to the analysis of complex sentences are also suggested. The procedures and the model described herein are original; however, they owe much to insights found in the work of two linguists, Gruber and Fillmore.

The general problem of grammar recognition is that of going from a surface string of words to a deep representation that permits semantic interpretation. More specifically, our grammar recognition procedure depends on the identification of the precise function or semantic role that each noun phrase actant occurring in a given sentence exhibits with respect to the verb of that sentence.

By assigning verbs--or, to be more precise, verb senses--to one or more paradigms (i.e., perceptually and functionally defined surface configurations), it becomes possible to determine algorithmically for every sentence the functional relation (e.g., theme, causal actant, goal, source, locus) that each noun phrase in the sentence bears to the verb, thereby assisting greatly in

arriving at a representation of the meaning of each sentence.

A number of verb paradigms such as intransitive, transitive and ergative are defined. Verbs belonging to the intransitive paradigm such as die, fall, go, etc. always have subjects that function as themes. Verbs belonging to the transitive paradigm such as kill, read, eat, etc. have subjects that function as causal actants and objects that function as themes. The ergative paradigm, which is more complex, consists of change-of-state verbs such as open, melt, increase, etc. If an ergative paradigm verb has both a subject and an object, the subject is a causal actant and the object is a theme; however; if such a verb takes only a subject, then the subject functions as a theme. The paradigm membership of each verb sense in the data base is determined and is recorded as a lexical feature of that verb.

The number of verb paradigms would proliferate almost indefinitely were it not for several devices, built into the grammar. One of these devices is the reversal of transformations such as passive and interrogative so that subject and object functional relations remain the same as in active, declarative sentences Another such device is the recovery of noun phrase deletions such as the one occurring in a sentence like "John ate." In this case, an indefinite object is reconstructed and the transitive paradigm feature of the verb eat remains intact. A third device of this nature is use of a notion called incorporation. For example, a sentence such as "It is raining." will be analyzed as having incorporated the theme subject rain' into

the verb with the result that a more abstract structure resembling "Rain is falling" gets reconstructed and processed as an intransitive verb paradigm item.

To overview the entire procedure, we start by parsing the surface structure of any given sentence. The major constituents and parts of speech are identified. Next, we determine the type of sentence involved (i.e., declarative, interrogative, imperative) and transform the word order where necessary. Following that the form of the verb (i.e., voice and tense) is identified and again the word order is transformed if need be. A surface role (i.e., subject, object) is then assigned to each noun phrase not preceded by a preposition. At this time the lexical entry of the verb is consulted for features of movement, deletion, incorporation, etc. so that any necessary final adjustments can be made. The paradigm membership of the verb is then identified thereby permitting a predetermined heuristic to assign a relation (i.e., theme, causal actant, locative, etc.) to every noun phrase in the sentence. Finally, by applying the above information as well as other lexical and grammatical information the surface structure is transformed into an appropriate abstract structure that permits semantic interpretation

TABLE OF CONTENTS

Verb Paradigms for Sentence Recognition*

Preface

For a number of years in the early 1970's the U.S. Air Force sponsored research at System Development Corporation in the application of natural language processing techniques to computer-aided instruction (CAI). The purpose of this research was to increase the overall effectiveness of CAI as an instructional method, with particular emphasis on extending the student computer interface to permit student-generated free form responses and queries. The research included experiments with a deductive question-answering system designed for use in CAI, the modeling of the behavior of a hypothetical tutor, and lastly the computer evaluation of constructed student responses, and a question-answering system driven by a dynamic model of the CAI lesson content. The subject matter used in the project was introductory meteorology--specifically concepts relating to the nature of precipitation.

---

*A slightly shorter version of this paper originally appeared in April, 1972, as one of several reports generated under Contract F33615-70-C-1726, which the Human Resources Laboratory of the U.S. Air Force entered into with System Development Corporation, Santa Monica, California. The views expressed herein are those of the author and do not necessarily reflect the views of the U.S Air Force or the Department of Defense.

One of the goals of this Computer-Aided Instruction Project was the development of an efficient recognition grammar that is as simple and realistic as possible without sacrificing information critical to semantic interpretation The primary purpose of this grammar was the analysis of student responses phrased in natural English- a necessary step in the answer-evaluation process. This paper describes in a general way the grammar recognition model developed as part of the CAI Project.

The model developed was based on the assumption that recognition grammars should be constructed quite differently from production grammars (i.e., they are something other than mere inverse algorithms of production grammars and involve certain questions of heuristics that are not of concern in the construction of production grammars); therefore an attempt was made to develop a unique model specifically oriented to recognition tasks.[1]

---

[1]The author is not a computer programmer but an applied linguist, who collaborated on a grammar recognition model for the computer-aided instruction project with a team of computer programmers and specialists in artificial intelligence that included William J. Schoene, John S. Brown and Robert F. Simmons, and the author is greatly indebted to these colleagues for their assistance. Questions concerning the details of actual programs and routines implemented should therefore be addressed to one of these specialists rather than the author, who served the project in a non-technical capacity as a linguist.

At the time this CAI project was undertaken two of the most successful recognition grammars that had been constructed were the IBM "Automated Recognition Grammar for English," of Culicover, et al. (1969), and the work done at MIT on Project MAC, which is described in Winograd (1971). The Culicover recognitio grammar is based largely on the generative-transformational work of Chomsky (1965) and others, and is a highly complex mechanism that attempts to capture the insights of Chomsky's generative-transformational model from the recognition grammar point of view. Winograd's highly pragmatic recognition procedures, on the other hand, make use of a systemic-type of grammar a la Halliday (1961, 1966, 1967); however, the power of Winograd's system is much more in his limited world-model and semantic networks than in his grammar.

Rather than produce an imitation of either the IBM Recognition Grammar or the MIT Project--it is, in fact, possible that neither of these approaches will ultimately be the most useful one for recognition purposes--we tried to reexamine recent insights, especially Fillmore's Deep Case Hypothesis (1968b), Chafe's semantically-based grammar (1970), and Gruber's work on the Lexicon (1965, 1967) in order to see how the functions and co-occurrence relations they describe might best be incorporated into a recognition grammar without sacrificing the positional insights that much of Chomsky's work has made explicit. In retrospect, our greatest debt was to Gruber (1965) This research was, of necessity, carried out somewhat independent of the

linguistic descriptions presented by these authors since they typically endeavor to remain neutral with respect to questions of production and recognition in their work.

INTRODUCTION

The problem we faced is, in general terms, the following: Given (more or less) accurate parsings of students' responses, how do we, as simply as possible, automate the reconstruction of deeper, more abstract message representations that permit us to make reasonable semantic interpretations? This paper outlines the essentials of our solution to this problem. The solution is described in terms of the surface structures and deep representations of simple sentences because it is easiest to understand and evaluate the solution at this level. The final section will discuss some complex sentence structures and will demonstrate that the paradigmatic technique discussed in this paper can be extended quite naturally to handle complex sentences as well as simple sentences.

Given an utterance that is a simple English sentence, or part of a simple English sentence, the grammar driving the parser must, first of all, be capable of recognizing constituents such as noun phrases, prepositional phrases, finite verbs or verb phrases,

---

[2] By 'verb phrase' we mean not the full 'predicate' but only the root verb with its tense marker and optional modal and aspectual auxiliaries, etc.

and adverbs or adverb phrases in this surface string of words. The noun phrase preceding the verb phrase is referred to as the surface subject; and if there is one noun phrase following the verb, it is identified as the surface object. If two noun phrases follow the verb phrase, there are two objects, identified as surface object$_1$ and surface object$_2$ respectively. The verb phrase, plus whatever follows it in a simple sentence, has traditionally been referred to as the predicate; we shall be making use of the term in this sense too; however, the notions 'surface subject,' 'surface verb phrase,' and 'surface object(s)' are the notions that are most critical in the development of our approach.

Surface Subjects, Surface Objects, and Marked Vs. Unmarked Noun Phrases

All noun phrases occurring in English sentences will be viewed as actants that bear a specific functional relation to the verb or element of predicatio in the sentence. These functional relations (e.g., theme, causal actant, locus, etc.) will be discussed in the next section of the paper. We view the identification of the precise functions that all noun phrase actants in a given sentence exhibit to their verb as the basic problem of sentence recognition. English, more so than languages such as German or Russian, tends to give relatively little direct indication in surface structure as to what the function of a

given noun phrase in a sentence is. This is because noun phrases occurring as surface subjects and surface objects--with the exception of some pronominal and interrogative or relative forms--are completely unmarked in English (i.e., they bear no inflection that would exclude or suggest a particular function). Thus we may have a sentence with an unmarked noun phrase as the surface subject:

(1) The object descended.

Surface Subj.

The same unmarked noun phrase may occur as a single unmarked surface object (2) or as one of two unmarked surface objects (3):

(2) John saw the object

Surface Surface

Subj. Obj.

(3) John gave Mary the object.

Surface Surface Surface

Subj. Obj. 1 Obj. 2

Thus in sentences such as (1) through (3), the only information we can use if we want to identify the function of the noun phrases is:

(4) a) The order of the NP's with respect to the verb (i.e., subject and object (if present)).

b) When two objects are present, the order of objects with respect to each other (i.e., obj 1 or obj

c) The semantic class and syntactic voice (i.e., active/

passive) of the verb.

In cases such as these, where the only information we have about the noun phrases in a sentence has to do with their serial order, we say that the noun phrases are unmarked.

The surface subject of any English sentence is, by definition always unmarked (i.e., it is not preceded by a preposition); however, it is possible for the other noun phrase(s) in a sentence all to be marked (i.e., preceded by a preposition) (5), or for the non-subject noun phrase(s) to be a combination of marked and unmarked noun phrases (6):

(5) The object descended to the ground.

Surface Subj.	Marked NP
(by definition: unmarked)	

(6) John gave the object to Mary

Surface Subj	Surface Object	Marked NP
(by definition: unmarked)	(by definition: unmarked)	

From this point of view, most prepositional phrases are analyzed as marked noun phrases in our grammar, and the lexical shape of a preposition preceding the noun phrase usually gives us some information as to what function the following noun phrase is fulfilling in the sentence. This information is of a different order from the information stated in (4); however, we believe

that it must be fully integrated with such information if sentences are to be meaningfully recognized.

At the surface level, then, we are interested in recognizing unmarked noun phrases and marked noun phrases (i.e., noun phrases preceded by prepositions); within the class of unmarked noun phrases we also distinguish surface subjects from surface objects in terms of their position with respect to the verb· and where two surface objects occur, we number them $object_1$ and $object_2$ on a left-to-right basis. When marked noun phrases occur, special attention is given to the preposition that precedes the noun phrase. The voice of the verb phrase of the sentence is also recognized at the surface level. None of this information, however, will directly constitute the deep representation of a sentence. Exactly how this surface-structure information will be used to arrive at accurate deep representations will become clear in subsequent sections of the paper.

Functional Relations: Elementary and Dérived

The elementary functional relations in our grammar are the following:

theme

causal actant (C.A.)

locus

goal

source

The theme[3] is the most neutral actant and the one obligatory actant in a simple sentence. It is the person thing, fact, or state of affairs about which something is being predicated such as movement, location, ownership (change of location or state of ownership), class membership, etc.

The functional relation causal actant refers to the person, object, natural force, fact, or state of affairs etc., that functions to cause or bring about the action, change, or situation described or implied in the predicate. In addition, all nouns that may function as the causal actant in a sentence are marked in the lexicon with the feature (+potent)--directly or by redundancy rules, following Chafe (1970). Not all sentences will have a causal actant.

Where present, the locus[4] is the object or being that the

---

[3]Theme is a syntactically and semantically defined functional relation introduced by Gruber (1965). We follow Gruber's choice of terminology and note that a similar functional relation was referred to as the 'objective case' by Fillmore (1968b) and the 'neutral case' by the UCLA English Syntax Project (1968)

[4]The functional relation locus includes under one category the two functional relations that Fillmore and others have called dative and locative. This is again more closely in line with Gruber's analysis than with other available analyses

theme is in, at, on, near, etc. Its use also implies that there is no motion of the theme being expressed in the sentence (i.e., the functional relation 'locus' typically occurs with stative predicates).

The two other elementary functional relations that may interact with themes, causal actants and loci are the directional functional relations 'source' and 'goal'[5]--the notion of source often being marked by the use of the prepositions '(away) from' 'out of', and goal being expressed by the prepositions 'to(wards) 'in(to)', and 'on(to)'

The following six sentences give several examples showing how marked and unmarked noun phrases can be reinterpreted as one of the five elementary functional relations defined above.

(7) The object descended.
Theme

(8) John gave the object to Mary.
Source Theme Goal

(9) Mary has the object.
Locus Theme

---

[5]These functional relations have been used with great success by Gruber in his lexical investigations (1965). We see the need to employ these notions as primitive functional relations in the type of grammar we have built. It is possible, however, that they should be treated as motional variants of locus. This further possible generalization is not utilized here.

(10) Mary broke the object

Causal Actant — Theme

(11) The object is on the table.

Theme — Locus

(12) John carried the object from New York to Los Angeles

Causal Actant — Theme — Source — Goal

The basic types of sentences that we will discuss shortly in terms of 'paradigms' have, as a minimum requirement, one noun phrase functioning as a theme; and at most, one occurrenc of any of the five functional relations described above. In other words, a simple or basic sentence type will not have two noun phrases functioning as themes, or two noun phrases functioning as causal actants, etc.

However, English has evolved a number of derived sentence types where two occurrences of a given functional relation are possible. Consider the following sentences:

(13) Hank broke the window.

C.A. — Theme

(14) A rock broke the window.

C.A. — Theme

(15) The noise broke the window.

C.A. — Theme

(16) Hank broke the window with a rock.

C.A. Direct — Theme — C.A. Indirect

In sentences (13) through (15), the surface subjects, regardless of semantic differences, are functioning as causal actants in our analysis. It is only in sentence (16) where two causal actants occur, that it becomes necessary to distinguish two kinds of causal actants: direct and indirect.[6]

Whenever two causal actants occur in an active sentence, the unmarked one in subject position is the obligatorily animate, direct causal actant and the one marked by 'with' that occurs later in the sentence is the inanimate indirect causal actant. In Fillmore (1968b) and other case-type grammars the preposition 'with' has often been analyzed as indicating the instrumental case--which we have reanalyzed as marking an 'indirect causal actant.' We feel that 'with' does not indicate any given functional relation consistently; rather one of the things 'with' indicates is a derived sentence type having more than one occurrence of a given functional relation. In (16) 'with' indicated an indirect causal actant. In sentence (17),

---

[6]Fillmore in his case grammar system distinguishes between agentive and instrumental causal actants. For our purposes, the functional relation causal actant is sufficient--'agent' and 'instrument' being something close to, though not identical to, what we describe as direct and indirect causal actants respectively A given sentence may contain both types of causal actants--in which case the direct causal actant is subject. If only one is present, its direct or indirect status is irrelevant; it is simply the causal actant subject of the sentence.

'with' indicates a co-theme, and in (18) 'with' indicates a co-causal actant:

(17) John went to New York with Peter.
Theme Primary — Goal — Co-theme

(18) Bill played tennis with Mary
Causal Actant Primary — Theme — Co-causal Actant

Note that the co-theme in sentences such as (17) and the co-causal actant in (18) must be semantically parallel to the primary theme or causal actant with respect to humanness or animacy. In sentences such as (19) where 'with marks a semantically non-parallel theme, two sentences are always involved--the surface object functioning simultaneously as the theme of S_1 and the locus of S_2:

(19) Mary saw the man with the cane
Locus S_1 — Theme S_1 — Theme S_2
+ Locus S_2

The fact that two sentences are involved in (19) is reinforced by a two-sentence paraphrase of sentences such as (20) first suggested by Lees (1960):

(20) Mary saw the man who has the cane.

The error of assigning one semantic function to 'with' (or to most other prepositions for that matter) is further emphasized by the ambiguity inherent in sentences such as the following, which have been discussed in another context by Hall (1965):

(21) The detective broke the window with the burglar.

C.A. Primary — Theme — Co-causal/Secondary C.A. Actant

In one interpretation, (21) means the detective used the burglar's body to break the window,' in which case, the 'burglar' is functioning as an indirect causal actant. In the other interpretation, (21) means 'the detective and the burglar broke the window together,' in which case 'the burglar' is a co-causal actant.

Obviously, a good deal of semantic information about the verb as well as semantic information about the noun phrases involved is needed in order to specify the exact role of a noun phrase following 'with.

Before moving on to the next section of the paper, some remarks are in order as to why we use only one category 'locus' for what others may distinguish as 'dative' and 'locative' on grounds of differences in animacy. Consider the following sentences:

(22) John has a new car.

Locus — Theme

(23) New York has skyscrapers.

Locus — Theme

(24) John knows the answer.

Locus-Primary — Theme

(25) *The garden knows the answer.

(*Primary Locus) — Theme

Sentences like (22) and (23) show us that there is a deep functional similarity between so-called 'datives' and 'locatives which justifies treating them uniformly as *Loci* in precisely the way that 'agents' and 'instruments' were seen to function similarly and were uniformly labeled as causal actants. Sentences (24) and (25) illustrate that some verbs require a primary (i.e., mental) locus as the subject. This suggests that there is one functional category 'locus'--but that on occasion it is necessary to distinguish primary (mental) loci from secondary (physical) loci.

In this presentation of our basic and derived functional relations, we have tried to point out that there is no need to posit many different functional relations in order to distinguish among the kinds of causal actants, themes, loci, etc. When and where features of animacy, purposiveness, etc., raise the need for distinguishing either primary loci (animate) from other loci or direct causal actants (agents) from other causal actants, the apparatus is available in the form of lexical features of nouns.

Classification of Verbs into Paradigms

In several of Fillmore's papers (see 1968a and 1968b) it is suggested that each English verb be assigned a "case frame" that would indicate what functional actants occurred obligatorily and optionally with a given verb. Translating from Fillmore's cases

into our functional relations, this would mean that the lexical entry of verbs like <u>die</u>, <u>open</u>, and <u>kill</u>, would contain information such as is represented in terms of the case frames shown in (26), (27), and (28) respectively.

(26) <u>die</u> +[_____] theme

(27) <u>open</u> +[_____] theme { (C.A.) / (direct C.A. (indirect C A.)) }

(28) <u>kill</u> +[_____] theme { C.A / direct + (indirect C.A.) }

Notation: { } = disjunction
( ) = optionality
[ ] = case frame
C.A. = causal actant

What these lexical entries mean is that <u>die</u> occurs only with one obligatory actant 'theme' (29). The verb <u>open</u>, like <u>die</u>, has an obligatory theme (30a) but also an optional causal actant (30b), which may even be realized as direct causal actant plus an indirect causal actant (30c). The verb <u>kill</u> again has the obligatory theme, but, in this case, the causal actant is also obligatory (31a); there is also the possibility of expressing the required causal actant as a direct causal plus an indirect causal actant (31b).

(29) <u>John</u> died.
Theme

(30) a. <u>The door</u> opened.
Theme

b. John opened the door

C.A. Theme

c. John opened the door with a stick.

direct C.A. theme indirect C.A.

(31) a. John killed Roger

C.A. theme

b. John killed Roger with a knife.

direct C.A. theme indirect C.A.

While agreeing in spirit with Fillmore, we propose to take the case frame suggestion one step further and make it more useful in terms of grammar recognition. Since there are numerous verbs that behave exactly like die, many others that behave exactly like open, and a good number of others that behave like kill, we propose to establish paradigms that summarize or recapitulate the functional relations and syntactic and semantic features of large classes of verbs. Then the only feature needed in the lexicon would be a specification of which paradigm(s) a given verb sense belonged to.

In other words, for verbs that are like die (i.e., require a theme which occurs in surface subject position and do not permit a causal actant), we establish the intransitive paradigm. For recognition purposes, the paradigmatic feature intransitive tells us that the surface subject is functioning as a theme, that

a causal actant is impossible, and that any other actant must be marked by an appropriate preposition.

The verb kill, however, is a member of the transitive[7] paradigm which requires both a theme and a causal actant.

---

[7]Note the difference between our use of the terms transitive and intransitive and the traditional and generative uses. In traditional terms, 'transitive' means merely that a verb takes both a surface subject and a surface object; in generative terms, 'transitive' has meant that a verb is capable of being passivized. In our system, transitive means that a verb takes a causal actant surface subject and theme surface object in the active voice. We stress this distinction because many of the verbs that allow the passive transformation do not have causal actant subjects in the active voice (e.g., Mary saw the men/the men were seen by Mary.). This is why we stress that our use of the term transitive to describe the behavior of a particular verb paradigm is more specific and functionally oriented than other usages of the term. While it is true that all of our [+transitive] verbs may occur in the passive as well as the active voice, this fact is not a defining property of [+transitive] verbs in our system but a redundancy rule of sorts. Likewise, intransitive means not only that the verb takes a surface subject and no unmarked surface object but that this surface subject is functioning as a theme.

Furthermore, from the recognition point of view, we know that for all verbs marked [+transitive] the surface subject is the causal actant and the surface object is the theme if the verb is in the active voice.

A large class of verbs like open belong to what we refer to as the ergative paradigm. 'Ergative'[8] is a surface structure typological term long used by linguists to characterize the grammar of those languages that seem to assign the same syntactic role or case to both the subject NP of an intransitive verb and the object NP of a transitive verb. (Two languages typically described as ergative are Basque and Eskimo.) In the light of this definition, the behavior of open and the other two verbs in the examples below appear to be 'ergative':

(32) a. John opened the door.

b. The door opened.

(33) a. The heat melted the ice.

b. The ice melted.

(34) a. The pressure increased the temperature of the air.

b. The temperature of the air increased.

---

[8] We prefer this traditional term to Lakoff's rather unconventional use of the term 'inchoative' in his dissertation, The Nature of Syntactic Irregularity (1965), and elsewhere to describe similar phenomena.

What happens in these sentences is that if the causal actant subject of the (a) examples is present, it is the surface subject; and the surface object, which must also occur, is the theme. If no causal actant is present, the surface subject is the theme and there is no surface object. Also the verb in an active, declarative English sentence must, as a rule, come second. Thus if there is no causal actant present filling the subject slot, the theme (or a noun phrase bearing some other functional relation) must fill the subject slot.

There are many ergative paradigm[9] verbs in English. In particular, the vocabulary of the meteorology lessons that our project was centrally concerned with contains many verbs belonging to this ergative paradigm (e.g., move, (re)distribute, reduce, relate, replace, skew, slant, spread (throughout), start, stop, decrease, transfer (in)to, weaken, break(up), boil, begin, dry (out), dissolve, decompose, divide (into), condense, form (out of), etc.).

---

[9]Some readers will question the need for an ergative paradigm and suggest that each such verb be assigned to both the intransitive and transitive paradigms. However, both the economy gained by stating just one paradigmatic feature for verbs like open and the apparatus which then permits us to explain the relationship between raise/rose, kill/die as suppletive ergative forms convinces us that this paradigm feature is extremely useful if not necessary.

The above ergative paradigm verbs retain the same lexical shape whether or not a causal actant is present in subject position. However, there are also a large number of pairs of transitive-intransitive verbs that, by virtue of their semantic (and sometimes phonological or orthographical) similarity, should perhaps be classified as suppletive members of an ergative verb pair. The transitive member of the pair occurs when a causal actant is present; the intransitive member occurs when no causal actant is present. For example:

(35) a. The pressure raised the temperature of the air.

b. The temperature of the air rose.

(36) a. John killed the rabbit.

b. The rabbit died.

The full ergative paradigm of some verbs contains a stative level as well as the two nonstative levels we have been discussing:

(37) a. Henry opened the door. (+C.A., -state)

b. The door opened. (-C.A., -state)

c. The door is open. (-C.A., +state)

(38) a. John killed the rabbit. (+C.A., -state)

b. The rabbit died. (-C.A., -state)

c. The rabbit is dead. (-C.A., +state)

The functional relation of the subject in the (c) sentences above is the same as in the (b) sentences--the difference between

them being the stative vs. the nonstative nature of the verb phrase. This difference sometimes has no effect on the surface form of the verb in irregular cases such as (37c), but it is more generally the case that a different surface form occurs (e.g., dead of (38c))-the regularly different form being not a lexically related adjective like dead but a past participle (e.g., the door is closed). The (a) sentences above contain what Chafe (1970) calls activity predicates, the (b) sentences process predicates, and the (c) sentences stative predicates.

Grüber (1965) has discussed transfer verbs at length, and based on his discussion we have found it desirable to establish three transfer paradigms. First of all, there are the two-way transfer verbs which allow overt expression of both a source and a goal:

(39) John (C.A.) {carried, brought, took, transferred} the package (Theme) from the living room (Source) to the kitchen (Goal)

Secondly, there are source-subject transfer paradigms which permit goals but not sources to be overtly stated in the predicate since the subject is the source of the transfer.

(40) John (Source and C.A.) {gave, sold, sent, ...} the car (Theme) to Mary. (Goal)

Thirdly, we also have the goal-subject transfer paradigm which permits sources but not goals to be directly stated in the predicate, and the subject is the goal of the transfer since a from phrase is implicitly or explicitly expressed.

(41) Peter (Goal) { bought / took / received / got / . . . } a watch (Theme) from Jane. (Source)

Somewhat similar to the transfer paradigm is the one-way put/take paradigm that includes additive verbs like put in(to)/on(to) (42) and privative verbs such as take out of/off (43)

(42) Sam (C.A) { put / placed / threw } the book (Theme) { in(to) the box / on(to) the table (Goal) }

(43) Helen (C.A) { took / pushed / pulled } the dog (Theme) { out of the house / off of the sofa (Source) }

In this paradigm virtually all of the verbs can be used in conjunction with either goals or sources--denoting the theme's coming to or going from a locus respectively. The preposition makes clear which case is intended. A few of these verbs even permit the expression of both source and goal in the predicate, and these verbs form a separate two-way put/take paradigm (44):

(44) Stanley (C.A.) { led / ushered / pushed } Mavis (Theme) out of the house (Source) into the garden. (Goal)

There is another paradigm for verbs like 'have' which require a locus as surface subject and a theme in object position (45):

(45) $\underset{\text{Locus}}{\text{Clyde}}$ {has / owns / possesses} $\underset{\text{Theme}}{\text{a fortune}}$

A paradigm for the converse of (45) includes verbs such as 'occupy which have surface subject themes and take loci in surface object position (46):

(46) $\underset{\text{Theme}}{\text{Joe}}$ {occupies / inhabits} $\underset{\text{Locus}}{\text{this room.}}$

There are several other possible paradigms which will not be mentioned and illustrated here. The point we want to make is that large numbers of verb senses can be classified in a way that facilitates recognition: the semantic function of surface subjects and objects of verbs can be easily ascertained by reference to a paradigmatic feature on the verb.

Paradigms and Transformations

The paradigms discussed above interact with a number of movement and deletion transformations that should also be discussed.

Firstly, a number of English verbs permit a transformational deletion in the surface structure of a lexically unspecified yet semantically delimited surface object theme. For example, a verb such as eat may occur either with a lexically specified object theme (47) or with a lexically unspecified one (48)

(47) Bill (C.A.) ate a sandwich. (Theme)

(48) Bill (C.A.) ate. = Bill (C.A.) ate _____ (Theme = food/something edible)

We describe sentences such as (48) as having undergone an 'accusative'[10] deletion. The deleted indefinite theme must be

---

[10]The term 'accusative' has been used by language typologists to refer to languages that treat the subject of an intransitive verb and the subject of a transitive verb with the same inflection. Also, in languages commonly described in terms of case grammar the accusative case generally corresponds closely to our notion of a surface object in English. We have borrowed this term and are using it to specify a particular type of deletion that occurs in English.

reconstructed if all the functional relations expressed in the sentence are to be explicitly stated in the deep structure. Other examples of verbs with potential accusative deletion in the data are: conduct, evaluate, gain, lose, observe, read, recall, understand, consume, etc.

Another type of transformational deletion occurs with virtually all motional verbs such as run, move, stir, walk, jump, etc., whenever the surface subject of such verb is [+ animate]. Consider the following sentences:

(49) a. The machine ran.
Theme

b. The door moved.
Theme

(50) a. John ran.
Theme and C.A.

b. The cat moved
Theme and C.A

The sentences in (49) which have inanimate surface subjects conform perfectly to the intransitive variant of the ergative paradigm discussed previously (i.e., the surface subject functions as the Theme) The sentences in (50), however, are somehow different in that the surface subject is not only functioning as a theme but also as a causal actant since no external causal actant has been explicitly mentioned. We can roughly paraphrase this type of sentence as follows:

(51) a. John caused himself to run.

b. The cat caused itself to move.

Since such paraphrases are not possible for the sentences in

(49), we assume that a reflexive' deletion takes place in sentences with transitive-paradigm verbs of motion or verbs of self-oriented habit (e.g., wash, dress, shave, etc.) When an animate surface subject functions both as the causal actant and the theme of the [+transitive] variant of an ergative verb; then the coreferential theme/causal actant is not stated twice in the surface structure if the verb may undergo reflexive deletion.[11]

The accusative and reflexive deletions discussed above help to explain an interesting ambiguity that occurs in sentences such as (52):

(52) Mary washed.

which may have either of the interpretations in (52'):

(52') a. Mary washed (the laundry)

b. Mary washed (herself).

---

[11]This is different, of course, from what happens to verbs that permit overt reflexivization. However, these non-deletable reflexives tend not to include verbs of motion or verbs of self-oriented habit.

(i) John cut himself.

(ii) Mary wore herself out.

(iii) The dog protected himself.

Since the transitive paradigm verb wash will be marked in the lexicon as additionally permitting either the accusative deletion or the reflexive deletion, the two ambiguous readings of (51) given in (52) will be automatically predicted by the recognition grammar since the accusative deletion rule will reconstruct (52a) and the reflexive deletion rule (52b).

In addition to deletions, the grammar recognizes and effectively 'reverses' various movement transformations or permutations that may have operated to produce the surface structure of certain sentences.

A well-known movement transformation is the passive transformation which operates on a sentence such as (53) and produces (54):

(53) Dick purchased the car yesterday.

(54) The car was purchased yesterday by Dick.

A subsequent transformation may optionally delete the 'by' phrase in (54) yielding (55).

(55) The car was purchased yesterday.

Given a sentence such as (54) our rules will 'reverse' the passive transformation and yield (53). Given a sentence such as (55) our rules will reverse the passive transformation and produce an active structure similar to (56).

(56) Someone purchased the car yesterday.

The procedure seems simple enough yet there are many problems involved in the accurate recognition and reversal of passive sentences First of all, not all 'by' phrases co-occurring with a passive verb can be reanalyzed as the surface subject of the corresponding active sentence:

(57) The rations were increased { a. by John. / b. by the sea. / c. by noon. / d. by 50 }

Of the above four 'by' phrases, only (a) could function as surface subject in the active voice version of (57); semantic features are needed to determine whether the noun object in a 'by' phrase is a causal actant, a location, a time, or a measurement. In our lexicon John is [+potent], sea is [+stationary], noon is [+time] and percent is [+unit][12] Such features will allow our passive reversal rule to construct a reasonable active voice variant for all the passive sentences in (57).

Another problem in passive reversal concerns the verbs in the various transfer paradigms. For several of these verbs most speakers of English recognize two different passive permutations (i.e., (59) and (60)) of the same active sentence (58):

Active (58) John (Source) gave Mary (Goal) the book. (Theme)

---

[12]Any unit preceded by a cardinal number gets reanalyzed as a measure phrase'

Passive (59) Mary was given the book (by John).

(60) The book was given (to) Mary (by John).

In such cases, semantic information about the surface subjects and surface objects of the passive sentences is required if the 'theme' and the 'goal' are to be properly labeled in the deep structure. In such cases the following feature hierarchy--or something like it--seems to operate:

(61)

human

animate

stationary

concrete

abstract

The two nouns involved as 'goal' and 'theme' seldom are at the same level on the hierarchy in terms of their lexical features, and the one that is higher than the other is always the goal. Thus in (59) and (60)--irrespective of the surface order in the passive--'Mary' is analyzed as <u>goal</u> and 'the book as <u>theme</u>.

Some readers no doubt may wonder why we have bothered to

write rules that depassivize[13] sentences. The answer is that we wish to get optimum efficiency and accuracy out of our paradigms without indefinitely proliferating their number. The paradigms are set up to assign functional relations to the marked and unmarked noun phrases occurring in active declarative structures. By first reversing the results of deletion transformations, question transformations and other movement transformations (e.g. passive), we are able to use a minimum number of paradigms to assign accurate functional labels to all of the noun phrases in a sentence.

The paradigms for transfer verbs just discussed above with respect to the passive transformation are also involved in another movement transformation which we refer to as 'goal focus'

---

[13]We are for the moment ignoring potential ambiguities and changes of meaning caused by changes of quantifier ordering in active and passive sentences such as those which Chomsky has pointed out (1965).

e.g. (1) Everyone in the room speaks two languages.
(can be interpreted as being different languages for each person)

(2) Two languages are spoken by everyone in the room.
(tends to be interpreted as two specific languages that everyone speaks)

To explain this transformation we must first establish the difference between immediate sources and goals on the one hand, and remote sources and goals on the other. With verbs of transfer in the active voice, the surface subject is a source, goal, or causal actant. The theme is the surface object. Other aspects of the motion or transfer are usually implied. The following examples illustrate this:

(62) John sold the house. (Implied: to some 'Goal')
Source Theme

(63) Harry bought a dune buggy. (Implied: from some 'Source')
Goal Theme

(64) Sam brought the beer. (Implied: from some 'Source' and to some 'Goal')
Causal Actant Theme

Verbs like 'buy' (e.g., get, obtain, acquire, fetch, borrow, etc.) and verbs like 'sell' (e.g., teach, serve, donate, give, lend, etc.) have an immediate goal and an immediate source in subject position respectively. In addition, they may express a remote goal or a remote source, both of which get marked with the preposition "for " The following examples show that, if the subject embodies the function "source," the optional "for" phrase will also embody (remotely) the function of source. If the subject expresses the function "goal, then the optional "for" phrase may express (remotely) either the function "goal" or the function "source."

(65) John sold the house for Harry.
Immed source — Theme — Remote source

(66) Harry bought a dune buggy for his son
Immed. goal — Theme — Remote goal or Remote source

Transfer verbs like 'bring' that express a causal actant in subject position and imply both source and goal are also ambiguous when occurring with an optional "for" phrase unless it is clear from the context that the causal actant is functioning on behalf of a remote source or for the benefit of a remote goal.

(67) Sam brought the beer for us.
Causal actant — Theme — Remote goal or Remote source.

In the above sentence "for us" can mean "for our use or benefit," in which case "for us" represents a remote goal. It also can mean "at our request, acting on our behalf," in which case "for us" represents a remote source.

One way of disambiguating sentences like (67) is to apply 'Goal Focus' movement (68) whenever the remote goal sense of "for NP" is the sense intended.

(68) Sam brought us the beer.
Causal Actant — Goal object$_1$ — Theme object$_2$

'Goal Focus' can move either a goal (70) or a remote goal that immediately follows the theme (69) to a position of focus

between the verb and the theme.[14]

(69) a. John bought the house for Mary.
Immed goal — Theme — Remote goal or Source

b. John bought Mary the house.
Immed. goal — Remote goal — Theme

(70) a. John gave the book to Mary.
Source — Theme — Goal

b. John gave Mary the book.
Source — Goal — Theme

Whenever goal focus movement occurs, it seems that the remote goal loses its preposition and becomes a goal object more intimately associated with the verb than it had been as a

---

[14]One apparent condition on "Goal Focus" movement is that the theme may not be a pronoun (the goal may be either a noun or pronoun). In other words, it can not occur with sentences such as the following:

a. John gave it to {Bill / him}

b. *John gave {Bill / him} it.

However the possibility (in fact the grammaticality of utterances such as "gimme it!" (i.e., Give me it) renders this condition dubious.

prepositional object.[15]

Our grammar makes use of several other deletion and movement reversing transformations that will not be discussed here. What we have tried to illustrate in this section of the paper is that both paradigms and transformations are useful components in a recognition grammar.

---

[15]We have reason to believe that an objectivalization movement rule like 'Goal Focus' in English is also widely used in other languages. French, for example, appears to have a movement operation even more general than the goal-restricted objectivalization transformation of English. It is for this reason that speakers of French say sentences such as "Open me the door" when speaking English. English does not permit the movement of "me" to object position in this sentence because "me" is a remote causal actant or source and not a remote goal in "Open the door for me." In this sentence "(you)" is the immediate causal actant and source and the verb "open" belongs to the ergative paradigm and not to the movement-transfer class of verbs. Thus the rules of English do not permit "Goal Focus" movement to apply in such a case.

Paradigms and Aspect

In the course of the development of this recognition grammar we noted that completive aspect (i.e., inference of completion vs. no inference of completion) sometimes serves to distinguish highly similar verb paradigms. In such cases, both the lexical shape of the verbs and the prepositions, as well as the order of actants in the surface structure, play a role in signaling the presence of completive aspect. Recall the paradigm for verbs of putting and taking (71) discussed earlier in the paper:

(71) a. John put water in the pool.
C.A. [-completive] Theme Goal

b. Mary took the groceries out of the bag.
C.A. [-completive] Theme Source

Note that the verbs of the sentences in (71) are not marked as having completive aspect: one cannot infer that the pool in (71a) is full of water nor that 'the bag' in (71b) is empty of groceries. There is, however, another paradigm similar to the put-take' paradigm but different from it in that all of its verbs indicate completive aspect. We refer to this [+completive] paradigm as the 'join-separate' paradigm (72), which is the paradigm to which the verbs fill and empty, for example, belong.

(72) a. John filled the pool with water.
C.A. [+completive] Locus Theme

b. Mary emptied the bag of groceries.
C.A. [+completive] Locus Theme

The sentences in (72)--unlike those in (71)--do allow us to infer that 'the pool' in (72a) is full and that 'the bag' in (72b) is empty This is because verbs of joining--which mark their theme with the preposition 'with'--and verbs of separating-which mark their theme with 'of'--are always [+completive] Other examples of joining verbs are: fill, supply, provide, stock, cover, present, furnish, plant, smear, sprinkle, etc. Additional examples of separating verbs are: relieve, empty, rob, deprive, withhold, rid, clear, drain, deplete, etc.

Some English verbs--as well as being used to express the completive notions of joining and separation--may also be used to express non-completive activities such as putting or taking. This dual function of the verbs has been the source of much confusion and unsatisfactory analysis, and is a topic we should, therefore, like to pursue in some detail. In Hall (1965), for example, the following sentences were considered more or less equivalent and thus relatable via a transformational rule which considered the object paint in (73) as basic and the object 'wall' in (74) as derived:

(73) John smeared paint on the wall.

(74) John smeared the wall with paint.

Fillmore (1968b, p. 48) explained these sentences by suggesting that both 'wall' and 'paint' were originally supplied with prepositions reflecting the locative and instrumental cases

respectively. This yields 'on the wall' and 'with paint' in the deep structure. Fillmore then analyzed the verb smear as having the following property: whichever of the two deep structure elements concerned is chosen as the 'direct object', it must fall next to the verb and must lose its preposition.

While more or less agreeing with Hall's intuition that 'paint' is a basic 'object'--in our system 'theme'--in (73) and a derived'object' or displaced theme in (74), we disagree with her implicit assumption that sentences such as (73) and (74) are syntactically or semantically equivalent or the explicit proposal that they should be related by a transformational rule. We also disagree with Fillmore's decision to analyze 'with' in sentences such as (74) as an instance of the instrumental case. In our view, aspectual differences between (73) and (74) call for a different assignment of deep structure functional relations and necessitate the postulation of distinct paradigms. In our analysis the surface object of (73) is a theme; the verb "smear" expresses an activity similar to "putting" in (73) and 'the wall is the goal of the theme 'paint'. In (74) 'smear' is being used to express the completed joining of a displaced theme--i.e., 'paint'--to the locus--i.e., 'wall'. In other words, our solution will analyze verbs such as 'smear' as belonging to two different paradigms. In one of the paradigms, 'smear' behaves like 'put' and in the other it expresses joining of a location and a theme. In this latter sense, the displaced theme is

expressed optionally, though always implied as something present but unspecified if not overtly expressed:

(75) a. John smeared the wall with paint.

b. John smeared the wall. (Implicit: with something)

(76) a John filled the glass with water.

b. John filled the glass. (Implicit: with something)

In the non-completive sense of 'smear' where 'wall' functions as goal, this 'goal' element is obligatorily expressed and not deletable.

(77) a. John smeared paint { on his brother / on the wall / ETC. }

b. *John smeared paint.

(78) a. John put the book { in the box / (over) there / ETC. }

b *John put the book.

Verbs of joining and separating are not the only ones belonging to a paradigm that signals completive aspect. There are two paradigms for verbs of contact that also can be used as examples. First consider the non-completive contact paradigm (79):

(79) $\underset{\text{C.A.}}{\text{John}}$ { threw / hurled / kicked } $\underset{\text{Theme}}{\text{stones}}$ $\underset{\text{Goal}}{\text{at the wall}}$.

(-completive)

In sentences such as those in (79), we cannot infer that the theme made contact with the goal; we assume the theme moved toward the goal, but we do not know whether contact was made. Contrast these sentences with verbs occurring in the completive contact paradigm (80):

(80) John (C.A.) { hit / pelted / bombarded } (+completive) the wall (Locus) with stones. (Theme)

These sentences in (80) force us to infer that the theme has made contact with the locus; no other interpretation is possible.

In the two sets of paradigms discussed above, some unusual phenomena have taken place. In those cases where the verb signals completive action, the surface order of actants in active surfaces is as follows:

(81)

Surface Subject				Surface Object		Marked NP
NP	+	VP [+completive]	+	NP	+	Prep + NP
↓				↓		↓
C.A.				Locus		Theme

This is not the normal order of elements. The usual surface order for active sentences containing a causal actant, a theme, and a locus is the following:

(82) Surface Subject NP + VP + Surface Object NP + Prep + NP

Surface Subject NP → ; Surface Object NP → Theme ; Prep + NP → {Locus / Goal}

Thus is appears that we get completive aspect only when a locus occurs in the position where one would normally expect the theme to occur.

This hypothesis is confirmed by a completive paradigm with no unmarked surface objects that has a locus in subject position instead of a theme (83). This paradigm is similar to the non-completive paradigm which has the theme in subject position (84) and follows the normal order of constituents for sentences having only a theme and a locative actant:

(83) The garden (Locus) is swarming [+completive] with bees. (Theme)

(84) Bees (Theme) are swarming [-completive] in the garden. (Locus)

The meanings of these sentences have been debated by Fillmore (1968b) and Chomsky (1972), among others. Fillmore assigns the same case representation to (83) and (84) suggesting that there is a difference of focus with perhaps some corresponding difference in cognitive content Chomsky points out that the

sentences are not synonymous--that in (83) the bees are necessarily all over the garden but that in (84) the bees might be only around their hive. In light of our analyses of sentences (71)-(80) we suggest that both Fillmore's remarks and Chomsky's remarks are correct but incomplete. The functional relations exhibited by the noun phrases in (83) and (84) are the same, the verb aspect is different. This suggests that there are two paradigms involved: a completive paradigm accounting for (83) and (85), and a non-completive paradigm accounting for sentence (84) and similar sentences like (86):

Theme + verb [-completive] + Locus

(85) a. Passengers are riding in the bus.

b. Fish swim in the stream.

c. Groceries were in the bag.

Locus + verb [+completive] + Theme

(86) a. The bus is sagging with passengers.

b. The stream teems with fish.

c. The bag bulged with groceries.

With lexically different verbs and prepositions as in (85) and (86) the two paradigms are easily distinguishable. When there is lexical overlap in different verb senses as in (83) and (84), the prepositions as well as semantic features will serve to distinguish [+completive] paradigms from [-completive] paradigms

Paradigms and Incorporation

In the second section of Gruber (1965: 5-27) there is a discussion of the grammatical process of incorporation. The term incorporation has been used most frequently by linguists working on American Indian languages; however, Gruber's discussion of incorporation along with the following discussion of certain paradigms supports the position that the concept of incorporation is needed to describe in full generality certain lexical and grammatical facts about English.

Gruber discusses verbs that obligatorily or optionally incorporate certain adverbs or prepositions. For example, Gruber claims that the verb cross obligatorily incorporates the adverb across. He illustrates this with paraphrase relations in acceptable sentences such as the following:

(87) John crossed the street.

(88) John went across the street.

Note also that the sentence below is impossible:

(89) *John crossed across the street.

The lexical and grammatical facts cannot be explained with maximum generality unless the verb cross is analyzed as a motional verb similar to go yet also incorporating the adverb across.

Another example of incorporation that Gruber provides concerns verbs co-occurring with the preposition for. The verb want obligatorily requires and incorporates for, while the verb

yearn requires but never incorporates for. The verb wish, however, requires for and allows both possibilities (i.e., incorporation of for is optional). Gruber's examples illustrate this:

(90) a. John wants a book.

b. *John wants for a book.

c. *John yearns a book.

d. John yearns for a book.

e. John wishes a book.

d. John wished for a book.

These verb-preposition co-occurrences would be very difficult to explain economically without making reference to the notion of incorporation.

Likewise, in setting up the paradigms needed to account for a large number of verbs in English we realized that certain verbs could be best described as belonging to a particular paradigm but, in addition, incorporating a specific noun actant. For example, the paradigm describing source-subject transfer verbs like give, send, sell, etc. can also include verbs like help, aid, support, etc. if we recognize that such verbs incorporate the theme but are otherwise exactly like the source-subject transfer paradigm:

Normal Source Transfer

(91) John Source { gave / sold / sent } the car Theme to Mary Goal

Goal Focus

(92) John Source gave Mary Goal the car Theme

Theme-incorporating Source Transfer

(93) John {helped / aided / supported} Mary
Source +Theme Goal

The fact that these incorporating verbs can be paraphrased as follows further reinforces this analysis.

(94) John gave {help / aid / support} to Mary
Source Theme Goal

Goal Focus

(95) John gave Mary {help / aid / support}
Source Theme

The theme is not the only actant that may be incorporated. Another such example of incorporation is found in verbs such as map, chart, graph, stratify, etc. which resemble two-way transfer verbs like bring, transfer, take etc. except that the goal actant has been incorporatdd in the verb Thus we have the regular two-way transfer paradigm:

(96) John {brought / took / transferred} the parcel into the room
Causal Actant +Theme Theme Goal

and the goal-incorporating version of the two-way transfer paradigm:

(97) John {mapped / charted / graphed} the statistics.
Causal Actant +Goal Theme

This analysis is again reinforced by the possibility of paraphrasing the goal-incorporating two-way transfer verbs as follows:

(98) John (Causal Actant) transferred the statistics (Theme) onto the {map / chart / graph} (Goal)

Once it is recognized that part of the lexical description of a verb may include the fact that it belongs to a particular paradigm but also that it incorporates a particular functional relation more typically expressed by a surface noun, the paradigms will apply to many additional verbs, and thus many unusual phenomena can be explained more consistently and adequately.

Consider the verb 'surface' as it occurs in the two following sentences--which on a superficial level appear grammatically similar:

(99) Phil surfaced the treasure chest.

(100) The men surfaced the sidestreet.

In each of these sentences the verb has incorporated the noun 'surface' as an actant; however, in (99) the goal actant has been incorporated and in (100) the theme actant has. Thus the verb 'surface' is being used in two very different ways The following paraphrases of the above sentences will help clarify this.

(99') Phil (C.A.) brought the treasure chest (Theme) to the surface (Goal)

(100') The men (C.A.) {put / added} a surface (Theme) {on / to} the street (Goal)

Without the notion of incorporation such pairs of sentences would pose severe difficulties for our paradigmatic analysis of verbs. Given the notion of incorporation as a working hypothesis, however, virtually every verb in English will be explainable in terms of a finite set of paradigms.

One final example of incorporation that we would like to discuss has to do specifically with verbs of precipitation.

(101) It's { raining / snowing / drizzling / sleeting }

Such sentences are unusual because the surface subject is a lexically empty dummy element 'it', so it appears that this sentence-type has no real theme. This could be somewhat distressing since we agree with Grüber that a predicating element and a theme are the minimal and obligatory elements in every sentence. Again, the notion of incorporation proves to be useful. If we analyze the verb in the above sentences as being very much like the minimal verb "fall" yet incorporating the theme, a plausible analysis is achieved in terms of both paraphrasability and paradigm assignment.

(102) $\underset{\Delta}{\text{It}}$ is raining.

dummy subject +Theme

(103) Rain is falling
Theme

The related historical fact is that English sentences require the verb in second position; thus verbs that for some reason (e.g., incorporation) are without an overt nominative subject to function as theme have come to require an empty 'it' (referred to as the impersonal or expletive it) in subject position.

From the Surface Structure of a Simple Sentence to its Deep Functional Structure

In the recognition grammar that we have developed, the first step in making a correct deep structure analysis of a simple sentence is the surface grammar's parsing of the noun phrase(s), prepositional phrase(s) (i.e., marked noun phrases), the finite verb (phrase), and other possible surface structure constituents such as conjunctions and adjectives or adverb phrases.

At this point, it must be established whether the form of the verb (phrase) is active or passive. The next step consists of looking up the surface verb in the lexicon to determine (i) what paradigm class it belongs to and (ii) what movement and

deletion operations may apply to the verb.

Next, it is possible, and in fact very simple, to take each unmarked noun phrase and to ascertain whether its surface structure role is subject or object and to mark the noun phrases accordingly.

Finally, by using the lexical information, a set of heuristics is consulted to assign the proper deep structure functional relation to each noun phrase occurring in the sentence. Once this has been done, rules of interpretation or "understanding" may be applied to the deep structure which has been reconstructed by the recognition grammar.

We shall now apply all these rules and strategies (except for the rules of interpretation) in an admittedly oversimplified form to a sentence in order to demonstrate the sequence of the procedure. The simple sentence we shall consider is the following:

The heat slowly evaporated the water.

(104) Step I: Apply surface grammar rules and parse the surface structure into labeled constituents.

The heat	+	slowly	+	evaporated	+	the water
NP		ADV		V(P)		NP

Step II: Assign a surface role to each NP not preceded by preposition.

The heat	+	slowly	+	evaporated	+	the water
SUBJECT						OBJECT

Step III: Determine the form of the finite verb (phrase) active, (simple past)

Step IV: Look up the verb 'evaporate' in the lexicon for paradigm classification and its features of movement, deletion, and incorporation.
ergative

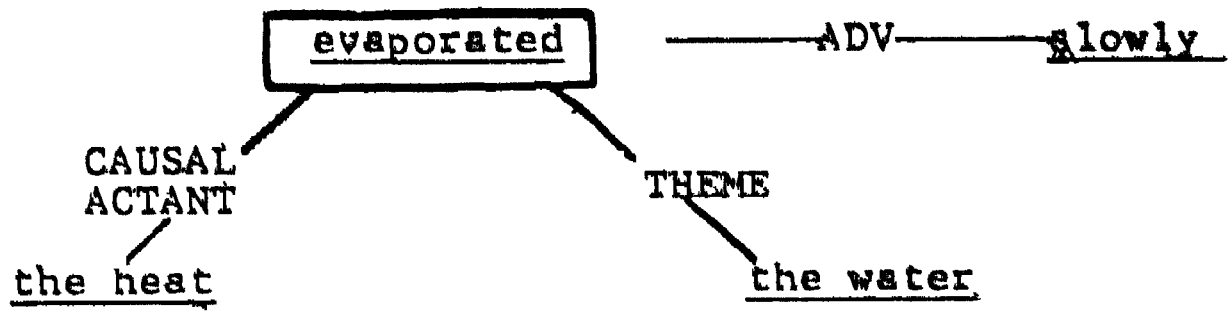
Step V: Look up the predetermined heuristics for surface sentences with active-ergative deletion verbs.
Heuristic: For sentences with active, ergative-paradigm verbs:
(i) If there is a subject and an object, the subject is a causal actant and the object is the theme.
(ii) If there is a subject but no object, the subject is the theme. (The stative form of an ergative-transitive verb always uses this strategy.)

Step VI: Apply appropriate heuristic of Step V to surface information determined during Steps II-IV and transform the surface structure into the appropriate deep structure.

(105) a. Surface structure:

The heat	+ slowly	+ evaporated	+ the water
NP	ADV	V(P)	NP
SUBJECT		ACTIVE	OBJECT
		[ERGATIVE +TRANSITIVE VARIANT]	

b. Deep Structure:



Thus the recognition system is conceptually complete as far as simple sentences are concerned, and complex sentences can also be handled, given certain modifications that will be discussed briefly in the following sections.

Extension of the System to Complex Sentences

It would be impossible to review all our procedures concerning complex sentence types in English; thus we shall merely exemplify our techniques focusing mainly on infinitival--and marginally on gerundive--sentential complements in the process.

Embedded sentences typically contain a number of optional and obligatory surface deletions which must be recovered. Reconstructions involve those subjectless predicates that are a result of deletion transformations taking place as part of the embedding process. For the sake of comparison, we have provided in (106) sentences which have undergone subject deletions or movements of this type, and sentences in (107) which have not.

(106) a. Sid wants to go there.
Subjectless predicate

b. To go there would be unwise.
Subjectless predicate

c. Camping is enjoyable.
Subjectless predicate

d. The children started playing.
Subjectless predicate

(107) a. Sid waited <u>for John</u> to come.

embedded subject

b. It would be unwise <u>for Eric</u> to go there

embedded subject

c. <u>Walter's</u> insulting us came as no surprise.

embedded subject

d. The mother regretted her son's <u>having stolen</u> the money.

embedded subject

In order to capture the fact that the infinitival and gerundive verb phrases in (106) have embedded subjects that have been deleted, moved or left lexically unspecified, our grammar makes use of the following highly general rules:

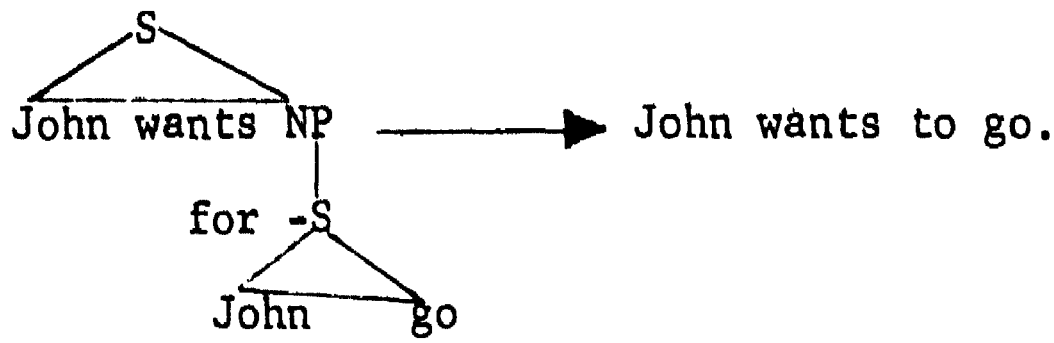
(108) a. Insert a 'for + Δ ' subject before any infinitival phrase not already preceded by a'(for) + NP' subject.

b. Insert a 'Δ' subject before any gerundive (i.e. -ing) verb phrase not already preceded by either 'NP' or 'NP' [+ poss]

Let us consider specifically four types of 'for -S' embedding situations that move or delete subjects.

Equi subject-subject deletion:

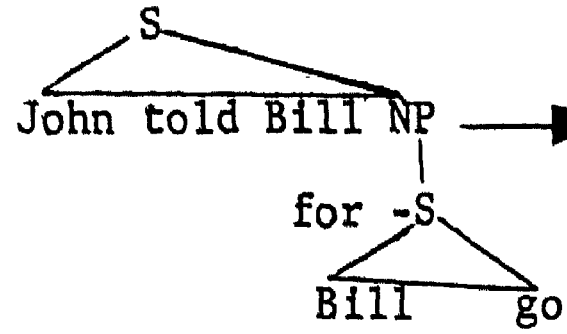
(109)



John wants to go.

Equi object-subject deletion:

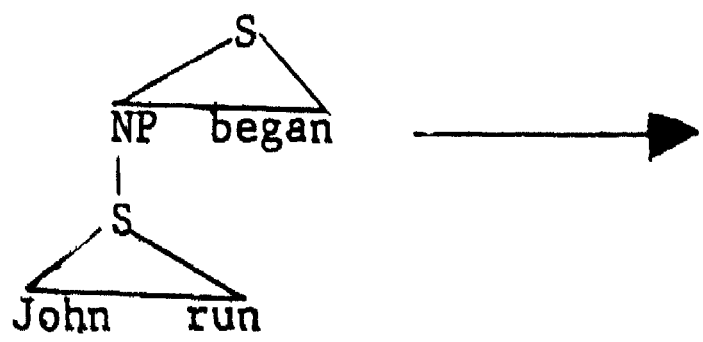
(110)



John told Bill to go.

Subject-to-subject raising:

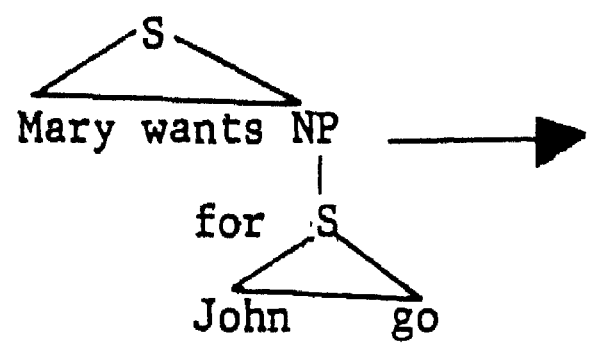
(111)



John began to run.

Subject-to-object raising:

(112)



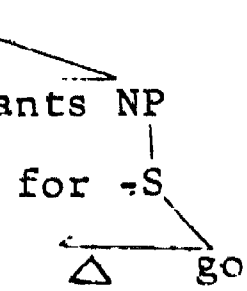
Mary wants John to go.

Note that in all of the above cases, when the lower sentence loses its subject through either movement or deletion, the lower predicate becomes infinitivalized.[16]

From the recognition point of view these movements and deletions must be reversed if we wish to reconstruct the subject and predicate of lower sentences accurately. The first step in our procedure is the application of the rule stated in (108b) since this will identify subjectless predicates, thus permitting the rules to build tentative deep structures.

(113) John wants to go: ⟶ John wants for △ to go. [17] ⟶

tentative deep syntactic structure:



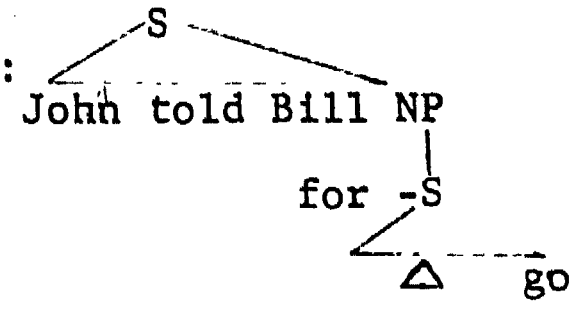
---

[16]It should also be pointed out that gerundivization as wel as infinitivalization can result from either of these operations (e.g., John began running; I like playing volleyball, etc.)

[17]We are tentatively inserting a deleted 'for △ ' in all such cases; however, this is a simplication which may eventually turn out to be infelicitous (i.e., in some cases it may not be necessary to insert a 'for').

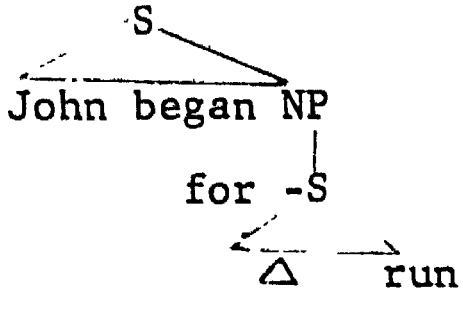
(114) John told Bill to go. → John told Bill for △ to go. →

tentative deep syntactic structure:



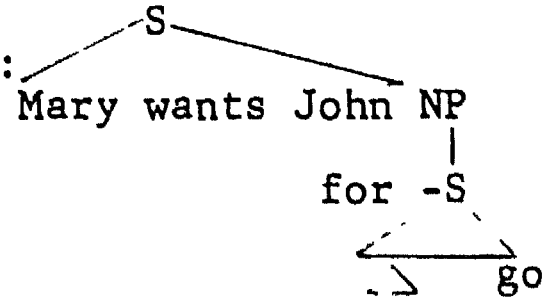
(115) John began to run. → John began for △ to run. →

tentative deep syntactic structure:



(116) Mary wants John to go. → Mary wants John for △ to go. →

tentative deep syntactic structure:



The next thing we must do is classify verbs in such a way that we can correctly fill the deltas in all of the above structures. Thus instead of describing a verb as a 'subject-subject deletion' verb (109), we will describe it as a 'copy subject' verb (113), so we know that the higher subject must be copied onto the delta in the embedded sentence in order for the syntactic deep structure to be complete and accurate. The following table shows how the features in our recognition grammar correspond to the features that might be used in a generative grammar.

Generative Feature	Recognition Feature	Sample Verbs
Subject-s		hope, ask...
Object-subject deletion	Copy Object	tell, order, command...
Subject-to-subject raising	Drop Subject	began, continue, tend, seem...
Subject-to-object raising	Drop Object	want, expect, wish...

Once these features have been assigned to the appropriate verbs, the tentative deep syntactic structures can be finalized.

There is, however, one transformation which must take place before the analysis of embeddings begins, and this is the 'passive-to-active' transformation. Thus a sentence such as (117) is first changed to (118) before the embedding procedures are effected.

(117) John was told to leave.

(118) ___ told John to leave.

If this transformation were not carried out, all features of object-embedding verbs would have to be re-analyzed in the passive as follows:

Active Voice	Passive Voice
Copy Object	Copy Subject
Drop Object	Drop Subject

While it would not be particularly difficult to change these specifications, it seems quite unnecessary in light of the fact that de-passivization must, in any case, be done before the appropriate paradigm heuristic applies; thus it can easily be carried out before embedding procedures apply as well. Paradigm analysis then begins with the lowest (i.e. most embedded) sentence in the tree and moves up with the result that an entire embedded sentence (via an intermediate NP node) is functioning as theme or causal actant in a higher sentence.

The procedures outlined above for dealing with certain types of complex sentences--however partial and sketchy-- do indicate that it is possible to extend the verb paradigm recognition technique to complex sentences, i.e., its use is not limited to simple sentences.

Concluding Statement

Although many details have been omitted from this paper, we believe that such an approach is as linguistically and psychologically valid as other theoretical or pragmatic recognition procedures developed to date.[18] Innovations in linguistics and psycholinguistics such as Bever and Langendoen's use of "perceptual strategies," which explain as well as account for certain

---

[18]In addition to the automated IBM Recognition Grammar of Culicover, et al. (1969) and Winograd's work at MIT on Project MAC (1971), both of which we mentioned in the introduction, there is, of course, also the Halle-Stevens "Analysis-by-synthesis" model (1964), which must be considered as a candidate for a recognition-grammar model theoretically--even though Halle and Stevens were concerned with speech recognition per se. Although we can see that the "Analysis-by-synthesis" model is useful in explaining hallucinatory reconstruction of speech and related phenomena, we feel that neither their fully active model (nor a fully passive model for that matter) will prove to reconstruct accurately the complexities of normal human speech perception or to explain human language recognition and comprehension.

facts of historical syntax (1972), tend to confirm our working assumption, which is that there are perceptual cues in the surface structure of English that are critical in grammatical recognition. We hypothesize that these perceptual cues, when combined with lexical information and information regarding the patterned movements and deletions that take place in syntactic configurations (i.e., paradigms, transformations and incorporations) make up much of the "knowledge" that the speaker of a language gradually acquires and then uses in recognizing and understanding the sentences of his native language (or any language he learns and knows, for that matter). We are convinced that this knowledge overlaps with, yet is somehow different from, the knowledge and skills that are required if one wishes to produce grammatical sentences in a language.

Kelley (1968), for example, has postulated that comprehension is basic in language acquisition and that rules of production are not essential for comprehension but may develop alongside of the necessary comprehension rules to satisfy other goals and purposes. If Kelley's model is correct, it indicates that human recognition grammars--an essential component of comprehension--are not mere inverses of human production grammars. This, in turn, suggests that the most efficient computer-based recognition procedures will be based, not exclusively on generative production-type grammars, but also on principles and

rules similar to those that human beings would seem to employ expressly in language recognition and comprehension. On the other hand, the model also suggests that the most realistic recognition procedure will overlap to a considerable extent with a production-type grammar--utilizing, in different ways, a great deal of the same information. In the development of our recognition model, such psycholinguistic considerations proved to be most useful.

REFERENCES

Bever, T. and D. T. Langendoen, "The Interaction of Speech Perception and Grammatical Structure in the Evolution of Language," paper presented at the UCLA Conference on Historical Linguistics in the Perspective of Transformational Theory, February,1969. This paper is published in Linguistic Change and Generative Theory, R. Stockwell and R. Macaulay, eds., Bloomington: University of Indiana Press, 1972.

Chafe, Wallace, L., Meaning and the Structure of Language, (Chicago: University of Chicago Press), 1970.

Chomsky, Noam, Aspects of the Theory of Syntax, (Cambridge, Mass MIT Press), 1965.

Chomsky, Noam, "Some Empirical Issues in the Theory of Transformational Grammar." Reproduced by the Indiana University Linguistics Club, 1970.

Culicover, P., et al., "An Automated Recognition Grammar for English," Technical Report FSC 69-5007, IBM, Boston Programming Center, Cambridge, Mass. 02139, July,1969.

Fillmore, Charles J.,"Lexical Entries for Verbs," Foundations of Language, Vol. 4 (pp. 373-393), 1968a.

Fillmore, Charles J., "The Case for Case," in Universals in Linguistic Theory, Bach and Harms,eds. (New York: Holt, Rinehart, and Winston), 1968b.

Gruber, Jeffrey, "Functions of the Lexicon in Formal Descriptive Grammars," System Development Corporation, Santa Monica, Calif., TM-3770, December, 1967.

Gruber, Jeffrey, "Studies in Lexical Relations," MIT Ph.D. dissertation, September 1965. (Reproduced by the Indiana University Linguistics Club, January,1970.)

Hall, (Partee), Barbara, "Subject and Object in Modern English," MIT Ph.D. dissertation, June,1965.

Halle, M. and K. N. Stevens, "Speech Recognition: A Model and a Program for Research," in The Structure of Language, Fodor and Katz, eds. (Englewood Cliffs, N. J.: Prentice-Hall), 1964.

Halliday, M.A.K., "Categories of the Theory of Grammar," WORD, 17, 1961.

Halliday, M.A.K., "Some Notes on 'Deep' Grammar," Journal of Linguistics 2, 1966.

Halliday, M.A.K., "Notes on Transitivity and Theme in English," Journal of Linguistics, 3. 1967.

Kelley, Kalon, "The Role of Semantics in Early Syntactic Acquisition," unpublished paper dated July, 1968.

Lakoff, George, "The Nature of Syntactic Irregularity," The Computational Laboratory, Harvard University, Cambridge, Mass., Report No. NSF-16, December,1965.

Lees. R. B., Grammar of English Nominalizations, IJAL, Vol. 26, No. 3. Part II, July,1960.

UCLA English Syntax Project, (principal investigators: R. P. Stockwell, P. M. Schachter, and B. H. Partee), Integration of Transformational Theories of English Syntax, 2 Vols. prepared under contract No. AF 19 (628-6007, for the U. S. Air Force, L. G. Hanscom Field, Bedford, Mass., October, 1968, Report ESD-TR-68-419). (Later published in a revised and somewhate abridged form as The Major Syantactic Structures of English by Stockwell, Schachter, and Partee, New York: Holt, Rinehart and Winston, 1973.)

Winograd, Terry, Procedures as a Representation for Data in a Computer Program for Understanding Natural Language, Project MAC, MAC TR-84, Mass. Institute of Technology, 1971.

END