

AD-A139 843

HOW TO MAKE GRAMMATICAL CHOICES IN TEST GENERATION(U)

1/1

UNIVERSITY OF SOUTHERN CALIFORNIA MARINA DEL REY

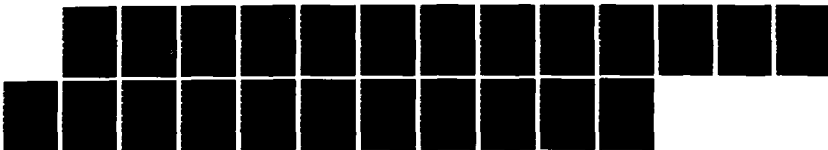
INFORMATION SCIENCES INST C M MATTHIESEN FEB 84

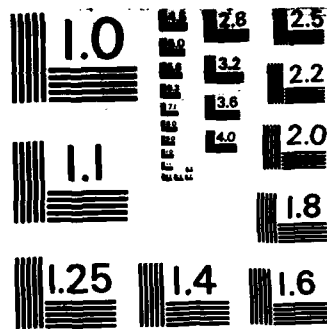
UNCLASSIFIED

ISI/RS-83-120 F49620-79-C-0181

F/G 6/4

NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

12

ISI Reprint Series
ISI/RS-83-120 ✓
February 1984

University
of Southern
California



Christian M. I. M. Matthiessen

How to Make Grammatical Choices in Text Generation

Reprinted from *The Tenth LACUS Forum 1983*

AD A139843

DTIC FILE COPY

DTIC
SELECTED
APR 5 1984
S A D

This document has been approved
for public release and sale, its
distribution is unlimited.

INFORMATION
SCIENCES
INSTITUTE



4676 Admiralty Way/Marina del Rey/California 90292-6695

213/822-1511

84 04 04 013

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER ISI/RS-83-120 ✓	2. GOVT ACCESSION NO. AD-A139843	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) How to Make Grammatical Choices in Text Generation	5. TYPE OF REPORT & PERIOD COVERED Research Report	
	6. PERFORMING ORG. REPORT NUMBER	
7. AUTHOR(s) Christian M. I. M. Matthiessen	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 90292-6695	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 February 1984	
	13. NUMBER OF PAGES 23	
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 Reprinted, with permission, from <i>The Tenth LACUS Forum</i> 1983, published by Hornbeam Press, Columbia, South Carolina.		
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) The problem of making purposefully controlled syntactic choices in text generation has not been given much attention in the past and there is no body of detailed and general solutions. This report briefly describes the problem and presents a solution to it by way of example. The tense expression of a clause is generated in the Nigel grammar, a large systemic grammar for text generation. The discussion shows how the problem of controlling tense is addressed in the Nigel framework. So-called systems are used to specify what the syntactic choices are and a semantic process, a so-called chooser, is defined for each system to exercise purposeful control of the choice alternatives specified by the system.		

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)

ISI Reprint Series

ISI/RS-83-120

February 1984

*University
of Southern
California*



Christian M. I. M. Matthiessen

How to Make Grammatical Choices in Text Generation

Reprinted from *The Tenth LACUS Forum 1983.*

**INFORMATION
SCIENCES
INSTITUTE**



213/822-1511
4676 Admiralty Way/Marina del Rey/California 90292-6695

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 and papers 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 90292-6695
USA

CONTENTS

1. INTRODUCTION	1
1.1 The process of grammatical choice: decomposition into systems	1
1.2 The process of semantic choice: choosers	2
1.3 Tense choosers	2
1.4 Novel contributions	2
1.5 Organization of the discussion	3
2. GRAMMATICAL CHOICE IN TEXT GENERATION	3
2.1 Four processes of text generation	3
2.2 Sentence generation	4
3. TENSE, THE AREA OF ILLUSTRATION	5
3.1 Grammar: tense resources	5
3.2 Choosing tense in conformity with time relations	7
4. INTERACTION BETWEEN CHOOSERS AND ENVIRONMENT: INQUIRIES	8
5. PRIMARY TENSE CHOOSER	10
5.1 Organization of chooser	10
5.2 The chooser in action	11
5.3 Additional chooser structure	11
6. SECONDARY TENSE CHOOSER	12
7. SECONDARY TENSE TYPE CHOOSER	14
8. TERTIARY TENSE CHOOSER	15
9. CONCLUSION	15
REFERENCES	16

Accession For	
NEIS CRA&I	<input checked="" type="checkbox"/>
ERIC TAB	<input type="checkbox"/>
Unpublished	<input type="checkbox"/>
Classification	<input type="checkbox"/>
Date	
Entered in	
Availability	
Dist	

A1



1. INTRODUCTION

In a functional account of language, semantics and grammar should be part of a model of communication: It is relative to this model that they are functional, i.e., that they can be motivated and understood in functional terms. Communication is a process; the central communicative construct is text. Text is a process; it is "a sociosemantic process" (see [Halliday 1978 78], p. 139). Communication is inherently interactive and there are two aspects to it: the process of producing text, text generation, and the process of constructing an interpretation, text comprehension. Ultimately, we want to account for both, but as a research strategy we can factor this task and construct two models, one for production and one for comprehension. These two models can then be used in designing a general model of communication. Obviously, this model will not just be the sum of the two separate models; it will entail revision, rejection and synthesis. The purpose of this paper is to present a contribution to a model of text generation.¹ If we assume that text is created to achieve goals, i.e., in a sense, any given text is a realization of a number of goals, one of the problems that can be identified for a model of text generation is the problem of controlling the process of grammatical choice purposefully. We can pose the following question:

How are grammatical choices made purposefully as part of the process of text generation in the creation of a text? Finding a solution to this problem is a challenge and a significant research task. Having a solution is critical to the success of any significant future text generation systems. At present there is no general and widely known solution to the problem of how to make grammatical choices. This paper presents a notation for stating how to make appropriate choices in the grammatical component (or stratum; grammar for short) of a text generation system. The type of grammar used in our research on text generation is systemic grammar.²

A notation for making grammatical choices is a semantic notation of grammatical choices, since it is a notation for stating what the choices mean. Any account written in the notation for a particular area of grammar is then a semantics for that area of grammar. One outcome of the enterprise is thus a factoring of the notion of semantics into (i) a semantics of grammatical choosing and (ii) other kinds.

1.1 The process of grammatical choice: decomposition into systems

The problem of purposeful grammatical choice entails another problem: How do we conceive of, organize, and state grammatical choice? (This is, of course, a question about how to handle paradigmatic organization.) Fortunately, this problem has already been addressed by systemic and stratificational linguists: Grammatical optionality can be handled by providing an account where each minimal grammatical choice point (system) is identified; this is the systemic approach,³ which will be

¹Text generation is a research task in computational linguistics, see particularly [Mann 82a], and computational tools can be used in the research, as they have in work by Davey and in our work; cf. [Davey 79] and [Mann 83]. They are helpful for testing and because they demand full explicitness, but are not necessary for research on text generation.

²This paper draws on work being carried out by W. Mann, M.A.K. Halliday, to both of whom I am greatly indebted for many insights, and myself. In particular, this paper rests on Halliday's representation of the grammar of English tense. I am very grateful to Bill Mann for comments on various versions of this paper. Responsibility for all errors rests entirely with me.

³The terms *system* and *systemic* are used in a special sense in the systemic approach to language. Systemic Linguistics. Following Firth, who developed linguistics in Britain from the 1930s, systemic linguists use a system to represent a choice between two or more options. Each choice point in the grammar is represented by a system. The concern with choices and their representation by systems is one of the characteristics of Systemic Linguistics. Others include the view of language as a communicative tool.

described below. In order to handle grammatical choice (optionality), we have to find a way of stating how to make an appropriate choice at each choice point. We will assume (without justifying the assumption here) that if we have an account of how to make a choice at each choice point, we also have an account of how to handle grammatical choice in general (cf. [Mann 82b]). So the task can be reduced to the task of stating how to make the right choice at a single choice point. This means that systemic grammar yields a decomposition of the task of choosing into systems. In addition, systems are organized into a network (the representation of the paradigmatic organization of grammar) and this organization also partly organizes the process of choosing.

1.2 The process of semantic choice: choosers

An account of how to choose in a system will be called a *chooser* (or *choice expert*). A chooser is a process that is composed of one or more steps that lead to the determination of which grammatical option to choose at a choice point. The decomposition of the task of handling grammatical optionality which the use of a chooser for each minimal grammatical choice point leads to is quite powerful. Each pair of a system and its chooser represents an account of choice on two strata: grammatical choice and semantic choice. I will refer to the theory of choosers as the *chooser framework*.

1.3 Tense choosers

The general problem of handling grammatical choice will be mapped onto the specific and concrete problem of controlling the grammar of English tense. (For a general account of the "anatomy" of choice, see in particular [Mann 82b].) For a particular clause of English, ... *who had been a nurse* ..., I will show how choosers can be applied to the grammatical tense alternations that English grammar offers us. By making the right distinctions, the semantics of tense formulated in terms of this chooser framework gives us the appropriate tense selection.

I have picked the area of tense, because it is sufficiently complex (some would even say messy) to demonstrate the capabilities of the chooser framework. Making appropriate tense choices is not a trivial task. (For a more detailed account than is possible here, see [Matthiessen 83a].) So in addition to simply illustrating the chooser framework used to account for a particular area of grammar, the tense example will serve to show how choosers together with a particular type of grammar, systemic grammar, serve to factor and make manageable tense, which has long been seen as one of the most intricate areas of English. Also, tense is an area where Halliday's analysis shows very well how systemic grammar decomposes the choosing problem in an interesting way; see e.g. [Halliday 76] and [Halliday 82].⁴

1.4 Novel contributions

As I have mentioned, the grammatical framework used in our work is that of systemic grammar. Consequently, the account of how to make grammatical choices is a proposal for systemic linguistics. In particular, it can be thought of a partial theory of systemic semantics, since "how to make grammatical choice" can be seen as semantic choice. However, to say that the account is intended as a systemic proposal is not to restrict it in scope or applicability. Rather, it is hoped that it can be of interest in any tradition where stratal organization and the representation of both paradigmatic and syntagmatic organization of grammar play important parts. Thus, there is a clear link to stratificational linguistics, for instance.

⁴Syntagmatic formulas in this area, such as *Aux --> Tense (M) (have - en) (be - ing)*, do not give us a handle on the choosing problem. Halliday's systemic account does.

The chooser framework constitutes a completely new development by Mann and Matthiessen in the context of the design of a particular text generation system (Penman) for a text generation grammar of the systemic kind (see [Halliday 69] etc.; cf. [Winograd 72], [Winograd 82], and [Davey 79]). The grammar and its chooser mechanism (together called Nigel) are not, however, dependent upon Penman; they constitute a general sentence generating component for text generation systems.

1.5 Organization of the discussion

To start with, I will make more precise what we understand by text generation (section 2). Next, I will introduce tense as the illustrative grammatical region to be used as an example of the application of the chooser framework (section 3). The rest of the paper (sections 4 through 8) deals with a presentation of the chooser framework with the help of the tense example.

2. GRAMMATICAL CHOICE IN TEXT GENERATION

There has been fairly little work done on text generation.⁵ One of the consequences is that the process of text generation is not very well understood as a research task. There is no established list of problems that have to be solved. I will not deal with this task of factoring text generation into research problems. Instead, I will merely assert that the process of making grammatical choices is a research task that emerges from such a factoring. I will give some credibility to this assertion by giving an indication of how grammatical choosing is part of text generation.

2.1 Four processes of text generation

Following an expository design for text generation systems by William Mann (see [Mann 83]), we can factor the generation process into four sub-processes called **Acquisition**, **Planning**, **Sentence generation**, and **Improvement**, each one of which is supported by resources. The first three work together to generate text. Improvement monitors the processes and criticizes the result and proposes remedies to improve it. Here we are only concerned with the first three.

Assume that our text generator is faced with the need for a text. There is a need to write a short biographical section about a particular person, Mrs. Jane McCluster. Based on this need, Acquisition searches for relevant information that might be included in the text or influence its organization. The knowledge support for this search is the **knowledge base** of the generator. The relevant information includes the facts that Jane is married, that she was a nurse before her marriage, that she grew up in town, that she married a farmer, that Jane had nursing experience, and so on.

Next, Planning responds to the information searched out by Acquisition and to the goals for the text derived from the initial text need. Planning uses a rhetoric for text organization. Planning's response is a plan for the text to be generated. The plan is global in scope, but it goes down to a level of detail where its units can be expressed by independent clauses. Here, the biographical section is the introductory part of a plan for a short story in which Jane plays an important role.⁶ There is a local plan for what can eventually be worded as *Jane McCluster, who had been a nurse before she*

⁵This is true of both linguistics and computational linguistics within Artificial Intelligence. The emphasis has been on discourse analysis and comprehension.

⁶For the example, I used Doris Lessing's "Little Tembe" from her African Stories. In general, short stories are far beyond explicit theories of text generation in various types of complexity, theories that can lead to computational models.

married, started a farm within a month of arriving. The aspect of it that is relevant for our discussion is the plan of the temporal relations represented by *who had been a nurse*: Apart from the time of telling the story, or NOW as we will call it, we need to keep track of two times, viz. the time at which Jane started the farm (which falls within the general time of the story), STORY-TIME, and a time before the general time of the story, the time at which Jane was a nurse, NURSING-TIME. An important goal that the clause *who had been a nurse* is intended to achieve is to tell us that Jane had had nursing experience before she started her nursing activities at the farm that the story goes on to tell us about. (Consequently, the tense combination will have to be chosen to signal that one step further back in time is taken, i.e., further back from STORY-TIME.) Figure 1 presents the time relations to be expressed diagrammatically.



Figure 1: Time relations to be expressed

Based on these plans for clauses created by Planning, Sentence generation produces a clause for each planned clause. The resource for Sentence generation is a grammatical component, i.e., a grammar.

2.2 Sentence generation

What does it take for Sentence generation to produce clauses with the help of its grammar? Here we can identify the problem addressed in the paper. A substantial part of the question we just asked is our initial question: How are grammatical choices made as part of the process of text generation? Briefly, the answer is that for each system of choice alternatives there is a corresponding process of choosing. This is the process of choosing mentioned in the introduction that we call a **chooser**. The grammar consists of hundreds of systems. Consequently, Sentence generation makes use of hundreds of choosers, one for each choice point.

Each chooser appeals to the environment of the sentence generation component in order to make an appropriate choice. The environment is the knowledge base and the text plan. The temporal relations diagrammed in Figure 1 are all in the environment, beyond choosers and grammar (or above, to use a vertical stratal metaphor). However, as we will see, choosers access this temporal information by presenting inquiries to the environment. It is on the basis of responses to these inquiries that choosers can direct appropriate choices of tense features to grammar. This aspect of sentence generation is summarized in Figure 2.

The figure can be thought of as representing a stratum of grammar, with grammatical choice, and a stratum of semantic choice, the process of controlling grammatical choice. Above this latter stratum, we find plans in the environment. As I have already indicated, the environment has its own organization. However, the important point for the present picture is that all that is above the level of choosers.

As is indicated in the figure, the paradigmatic organization of the grammar, represented by systematized features, is related through realization to the syntagmatic organization: Tense features are realized by fragments of a functional tense structure. For example, the selection of the feature future ultimately leads to the presence of *will* in the structure being built.

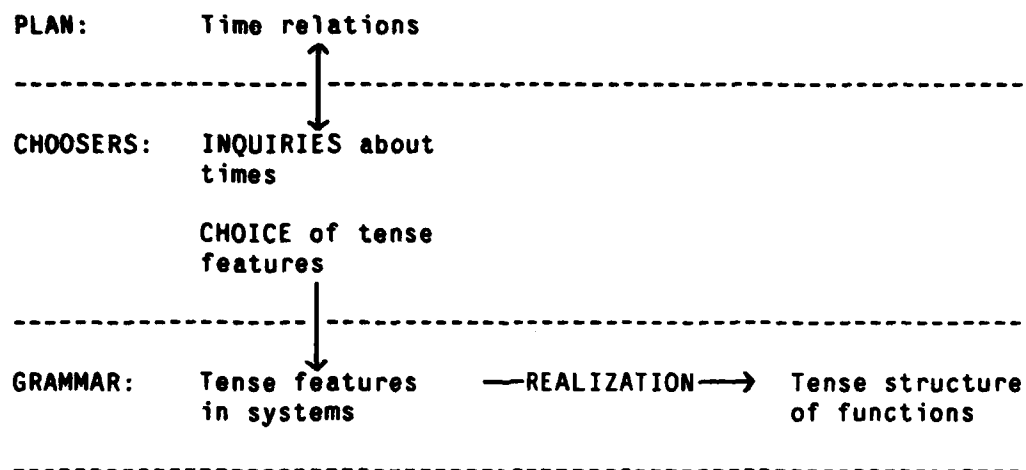


Figure 2: Partial organization of Nigel

Before describing choosers in more detail, I will introduce the area of grammar to be used for illustration, tense. This will also give me an opportunity of presenting some aspects of the systemic grammatical notation that we will need to understand how choosers work.

3. TENSE, THE AREA OF ILLUSTRATION

3.1 Grammar: tense resources

The part of the grammar of English tense that will be used in the illustration of the chooser framework is given in Figure 3. Each tense choice is represented by a system and the choice options by grammatical features like *past* and *future*. As the figure shows, the individual tense systems form a network of tense choices. For instance, once we have made a selection in the system SECONDARY TENSE, we proceed to SECONDARY TENSE TYPE, making a new choice. The loop back in the figure for the system SECONDARY TENSE means that it is possible to have a series of secondary tenses: After primary tense comes secondary, then tertiary, quaternary, and quinary. (For certain reasons, it is difficult to go beyond quinary tense, but not impossible; cf. [Halliday 76].) In systemic notation, these tenses can be called α (primary), β (secondary), γ (tertiary), δ (quaternary), and ϵ (quinary). The systems of secondary tense and secondary tense type can be thought of as schemas, standing for the specific instances of looping through them (as the loop back suggests). The first instance is secondary tense (or second order tense), the second is tertiary (third order) tense, and so on.

The network represents the choice organization of the grammar. Choices are re-expressed by realization processes in structural terms. The realizations of the tense terms for primary tense and secondary tense are given in the table in Figure 4. I will not go into the design of realization statements here, nor into the functional nature of the resultant structure; there is general discussion in e.g. [Matthiessen 83b].

The tense expressions that the grammar in Figure 3 can generate are all exemplified as variants of

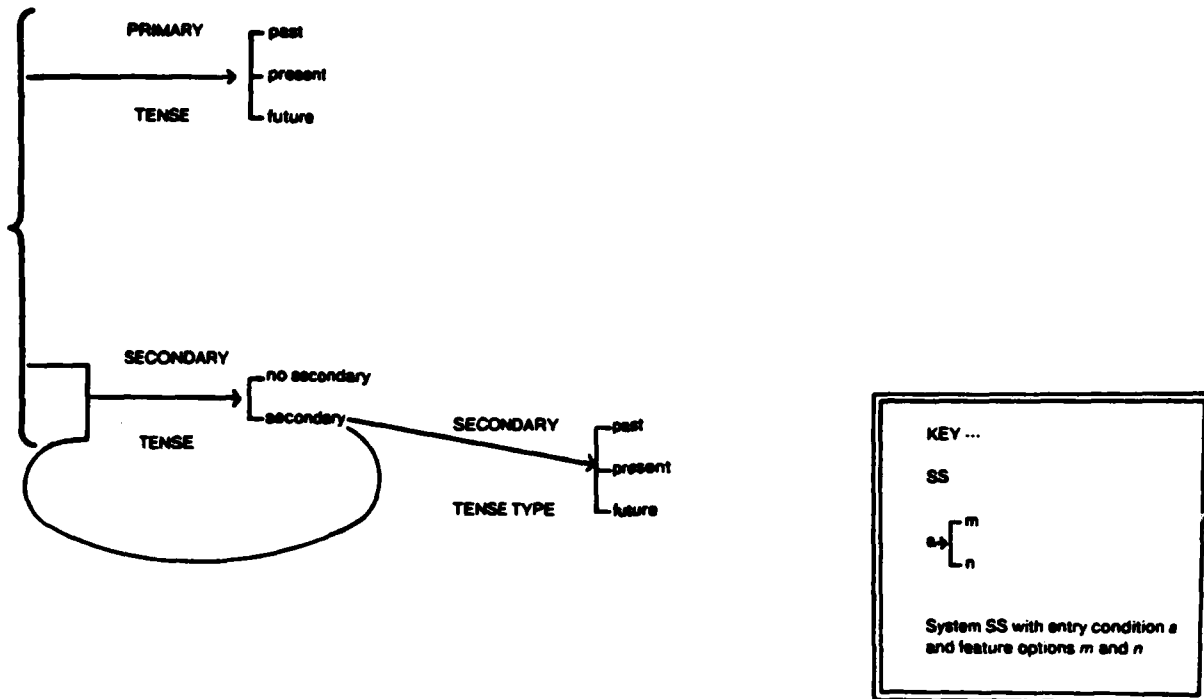


Figure 3: The grammar of tense

	past	present	future
primary	-ed	0	will
secondary	have ... en	be ... ing	be going to

Figure 4: Tense markers

who had been a nurse in Figure 5.⁷ The table determines two considerations for our tense chooser task, the selectivity of the task and the completion of the task.

1. **Selectivity.** The range of tense expressions shows what our tense choosers have to deal with: The choosers have to be clever enough to choose so that the outcome is *had been* rather than *has been*, *was*, or any of all the other competitors.
2. **Completion** We will have completed our task of finding the appropriate tense combination when we have reduced the table in Figure 5 to a single expression.

⁷The table reads as follows. Vertically, from the top, the choices of PRIMARY TENSE are represented. On the left hand side, first the features of SECONDARY TENSE, no secondary and secondary, are represented, then the three features of SECONDARY TENSE TYPE. As the bottom line indicates, the table is valid when there is no tertiary tense. The table would have to be expanded considerably for the selection of tertiary.

PRIMARY SECONDRARY		past	present	future
		no secondary	was	is
sec.	past	had been	has been	will have been
	pres	was being	is being	will be being
	fut	was going to be	is going to be	will be going to be
		no tertiary		
TERTIARY				

Figure 5: Tense combinations defined by tense grammar

3.2 Choosing tense in conformity with time relations

For each system, its chooser chooses a feature according to conceptual distinctions in the environment. As we have already seen in Figure 2, choosers thus mediate between environment and grammar as shown in Figure 6.

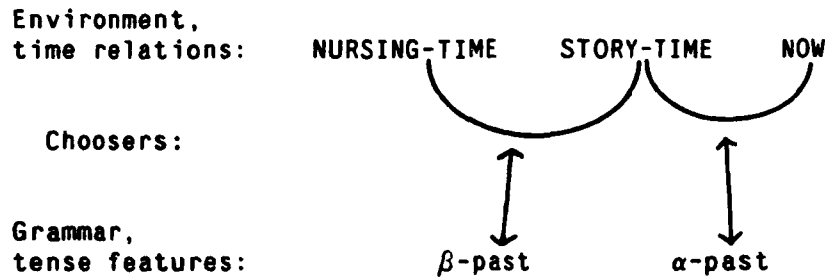


Figure 6: The boundary between environment and grammar

The precedence relation between NURSING-TIME and STORY-TIME will be expressed by the choice of secondary past and the precedence relation between STORY-TIME and NOW will be expressed by past from PRIMARY TENSE. To generate the tense expression *had been*, we will have to make the following choices. Every finite clause has a primary tense; the task is here to specify the type of primary tense we need. We choose feature past, since STORY-TIME precedes NOW. In the system SECONDARY TENSE, we choose feature secondary, because NURSING-TIME and STORY-TIME are discovered to be different times, i.e., because we want to remove NURSING-TIME one step further from NOW than the minimal tense expression *was* (as in *who was a nurse*) indicates. The next system, SECONDARY TENSE TYPE, lets us specify the direction of this step. We choose feature past

to express that NURSING-TIME precedes STORY-TIME.⁸ The grammar will re-express this feature for us in structural terms as the tense auxiliary *have*. With the selection of a secondary past tense, we have completed the tense chain. When we check to see if a higher order of tense is needed (looping back to SECONDARY TENSE) we use the same strategy as in the first traversal of SECONDARY TENSE and find that it is not needed.⁹

The reasoning we have just been through was informal. However, it is precisely this type of reasoning that the choosers for these systems must be capable of and this is the task we face.

4. INTERACTION BETWEEN CHOOSERS AND ENVIRONMENT: INQUIRIES

We can use dialogue as a metaphor for the interaction between choosers and the environment. The interaction between the two can thus be presented as an inter-stratal dialogue (cf. Figure 2 above). (We can present the interaction between a chooser and its system in the same way.) When the chooser process starts, the first tense chooser does not have any information from the environment. But in order to make a choice, it must obtain the relevant information. It must determine whether the conditions under which a grammatical feature like *past* is selected obtain, i.e., whether the choice conditions have been met. This is achieved by inquiries. An inquiry is one of the steps that constitute the chooser process. It is a request for information from the environment. Inquiries are the only mechanisms that choosers use to interact with the environment.

The inquiries from a chooser and the responses from the environment are the turns in the dialogue between these two components of a text generation system. As a set-up for the tense choosers, we need to establish what the time of speaking is since this is crucial to tense. It enters into temporal relations with other times and these relations constitute choice conditions. This happens in the following little conversation:¹⁰

Chooser: What concept represents the current time, the time at which the language is generated?

Environment: It's called NOW.

Chooser: I'll put the association of NOW with T_0 in the function association table.

⁸It is important that the grammar and choosers distinguish between the decision to have an additional tense (take a further temporal step) and the decision about the type of tense (the direction of the step). For instance, secondary tense can be used for a flash forward instead of a flashback, as in the example from the Lessing story: *As Bishop Angelo Roncalli, who would become Pope John XXIII, once said, he sought contacts "with people from both sides of the river" ...* (LA Times).

⁹When the conceptual temporal relations are more complex, the tense expression is also more complex. For example we can get *Jane will have been going to be a nurse in couple of weeks for ten years next Friday*, although need for such complex tense expressions is fairly rare; see [Halliday 76].

¹⁰The systems in Figure 3 represent the grammar of tense. However, before we reach one of these systems that encodes an option in tense, there are some inquiries which always have to be asked (no matter what tense combination is chosen) and that may be relevant not only to tense determination but also to the selection of a time adverbial. Consequently, these questions are asked early on in the generation of a clause; the information will be available to later choosers both for tense and for adverbials.

First, the chooser uses an inquiry that is intended to identify the time of speaking.¹¹ An identifying inquiry asks for a concept of a particular kind. The response is the concept that fits the description for the particular clause being generated. Identifying inquiries are thus open set inquiries.¹²

Next, the environment responds. The way it arrives at the response will of course vary with the particular design of the environment. What is important is only that the environment does give an appropriate response. Here the response is in the form of a symbol for a hub, NOW.

A hub is a conceptual locus in the environment. As a concept it participates in a number of relations and information about what category it belongs to, about its subtypes and instances and so forth can be derived. The conceptual locus is like the hub of a wheel (hence the term) in that it is at the centre of relations like these through which additional pieces of information can be accessed.

Thirdly, the chooser associates the response from the environment with a grammatical symbol; the hub NOW is associated with the variable T_0 .¹³ The conceptual information has thus been indexed in the grammar and is now available through T_0 . The table in Figure 7, the so-called function association table, shows the association of T_0 and NOW (at the top) and the other function-to-hub associations that will be established in the course of our example.

TIME- HUB	TENSE- FUNCTION
NOW	T_0
STORY-TIME	T_1
NURSING-TIME	T_2
NURSING-TIME	T_{PROC}

Figure 7: Function Association Table

The name of the present inquiry is *SpeakingTimeID*; we can represent the inquiry succinctly as (*SpeakingTimeID PROCESS*). The inquiry is applied to the grammatical function *PROCESS*, since information about the speaking time is available through the hub that is associated with this function.

¹¹ Here this inquiry is represented by an informal question in English. It has turned out to be highly useful, in fact crucial to the work, to have an informal question like this in addition to the formal version of the inquiry. In presenting the chooser framework, the informal English questions are also quite helpful and I will continue to use them here.

¹² There are two kinds of inquiries, identifying ones and branching ones, the second of which we will meet shortly.

¹³ An important aspect of most recent systemic grammars is that grammatical structure is stated in terms of grammatical functions and not in terms of grammatical categories (classes) like noun phrase and noun. Any single constituent may serve more than one grammatical function; each of these functions can have a hub-association. For mnemonic reasons, it is convenient to use T_0 , T_1 , etc. as names for the variables with which particular time concept values are associated. Instead of these, I could have used grammatical the grammatical functions that carry the realizations of the tense selections. For instance, the time of speaking for any clause would be associated with the function *FNRE*. We need a number of functions $TEMPO_0$, $TEMPO_1$, etc. in any case for the information about temporal auxiliaries and tense morphemes.

The procedure for establishing a hub for the time of the encoded by the clause, represented by T_{PROC} ¹⁴ is quite parallel; a chooser inquiry, an environment response, and an association of the hub symbol NURSING-TIME that constitutes the response with the grammatical function symbol T_{PROC} .

This chooser work concludes the set-up for the tense choosers.

5. PRIMARY TENSE CHOOSER

Before going through the primary tense selection for *who had been a nurse*, I will present the chooser.

5.1 Organization of chooser

The whole chooser is represented diagrammatically in Figure 8.

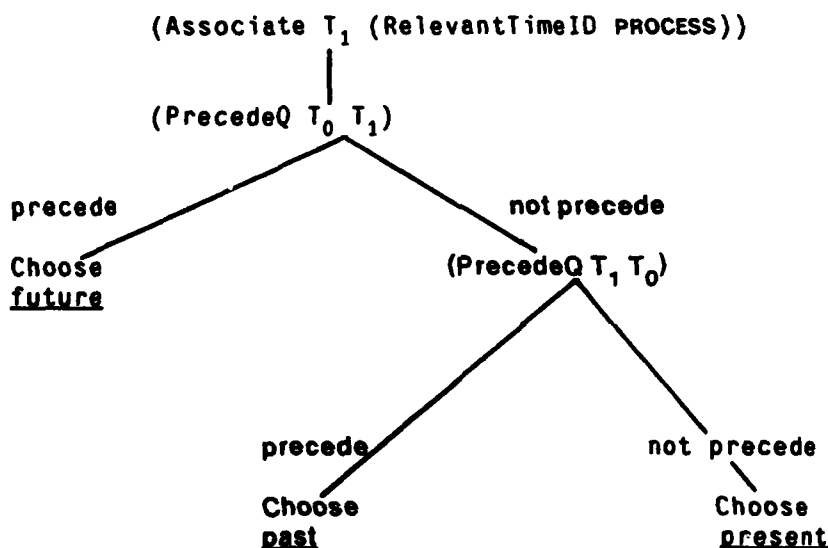


Figure 8: The chooser of primary tense

The first inquiry, RelevantTimeID, is quite parallel to the two we have met so far, SpeakingTimeID and EventTimeID, and is of the identifying kind and associates a hub, which comes from an open set of responses, with a grammatical function: (Associate (RelevantTimeID PROCESS) T_1) In our example, STORY-TIME is identified by the inquiry and then associated with T_1 .

The second inquiry is of a new kind; it is a branching inquiry, PrecedeQ. A branching inquiry defines a closed set of possible responses, usually two. These responses constitute choice conditions. For instance, the response T_0 precedes T_1 , is the condition under which the feature future is chosen. As can be seen from the chooser of primary tense, more than one response (condition) may be needed to determine the selection of a feature. In the chooser, two conditions have to be met

¹⁴The time of the process of the clause, T_{PROC} , is called *event time* by e.g. [Reichenbach 47]. Just as in systemic grammar in general, the term *process* is intended to be neutral between events (dynamic) and states (stative).

in order for present to be chosen: T_0 does not precede T_1 & T_1 does not precede T_0 . This is an example of a conjunction of choice conditions. As we will see below, choice conditions may also be disjunctive: There is diversification between choosers and systems.

Grammatical options are stated in a factored step by step way and the chooser semantics reflects this decomposition, since there is a chooser for each grammatically recognized option.

As the diagram shows, branching inquiries have closed set answers, represented by *precede* and *not precede* for PrecedeQ. They give choosers the structure of decision trees, with a branch for each response.

5.2 The chooser in action

Now, let's traverse the chooser of primary tense for our example *who had been a nurse*. First, the time that is related to the time of speaking, T_0 , is identified. As before, we can present the generation of our example as a dialogue between the environment and Nigel's choosers, using the informal English rendering of the inquiries:

Chooser: What concept represents the time to which NURSING (PROCESS) is relevant?
(i.e., (RelevantTimeID PROCESS))

Environment: It's called STORY-TIME.

Chooser: I'll put the association of STORY-TIME with T_1 in the function association table. (i.e., (Associate STORY-TIME T_1))

Next, we determine what relation obtains between the two times T_0 and T_1 .

Chooser: Does the moment or interval of time NOW (T_0) strictly precede the moment of interval STORY-TIME (T_1)?

Environment: No, it does not.

Chooser: Does the moment or interval of time STORY-TIME (T_1) strictly precede the moment or interval NOW (T_0)?

Environment: Yes, it precedes it.

Chooser: Then I choose feature past.

The choice of the feature past eliminates primary future and present in the table in Figure 5.

5.3 Additional chooser structure

Before we proceed to SECONDARY TENSE, let me indicate some further structure of the chooser of primary tense not yet presented. The feature past is chosen because there is a precedence relation: STORY-TIME precedes T_0 and we get *had* rather than *has* (which is present) or *will have* (which is future). However, some other branching inquiries come before the one about precedence. This is so because there may be other reasons for choosing the feature past; see the chooser fragment in Figure 9 which adds CounterfactualityQ above the first questions about precedence in the primary tense chooser.

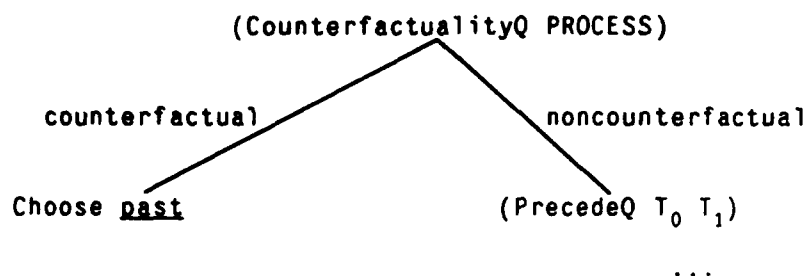


Figure 9: Addition to PRIMARY TENSE chooser

past may also be selected to convey counterfactuality, as in *I wish Henry spent less time looking for a wife and if Henry and Anne were here, I would know*. So before inquiring about precedence relations between T_0 and T_1 , the chooser presents an inquiry about the counterfactuality of the proposition to be expressed. If it is counterfactual, past is selected. Otherwise, the chooser presents the normal inquiries about precedence.¹⁵

6. SECONDARY TENSE CHOOSER

The second tense system we reach is SECONDARY TENSE; see Figure 3 above.¹⁶ This system controls the option of having at least one tense auxiliary (the secondary option, with *have*, *be* or *be going to*) or none (the no secondary option).

The chooser of SECONDARY TENSE is diagrammed in Figure 10. The path we will take appears in boldface.

The first inquiry is SameAsQ, a branching inquiry. Two times, T_1 and T_{PROC} are compared for identity. If they are the same, represented by "same" in the diagram, there is no time ordering to express in the clause in addition to the one that is always expressed in tensed clauses in English (PRIMARY TENSE; see above). Hence, the action taken by the chooser if the response is "same" is to choose the feature no secondary, giving us for example *who was a nurse*.

However, if the two times are not the same ("different" in the chooser diagram), there is an ordering of T_1 with respect to T_{PROC} to be expressed, and the feature secondary is chosen. The realization (expression) of this feature in the grammar is the addition to the grammatical structure being built of a tense auxiliary to express the ordering of the two times. The auxiliary is *have*, *be*, or *be going to*. The

¹⁵This discussion brings out a general property of choosers. Choosers are not restricted to having only one occurrence of the choice of a grammatical feature like past. Rather, a feature can be chosen under as many different circumstances as seem appropriate. In other words, the chooser approach does not commit us to the position that there is only one meaning for each linguistic form or even for each systemic option.

¹⁶Here I will not go into exactly how and when a system is reached. Suffice it to say that each system has an input or entry condition consisting of a single grammatical feature, a disjunction of features or a conjunction, as can be seen in the diagram of the tense grammar above. In general, any one of the systems whose entry condition has been satisfied may be entered next in the generation process. The control of system entry is a grammatical matter; choosers do not have entry conditions, nor do they control the entry of their systems.

present system does not introduce any selectivity among these three options. That is left to the next system, SECONDARY TENSE TYPE. As already noted, this factoring of the statement of options is a general feature of Nigel.

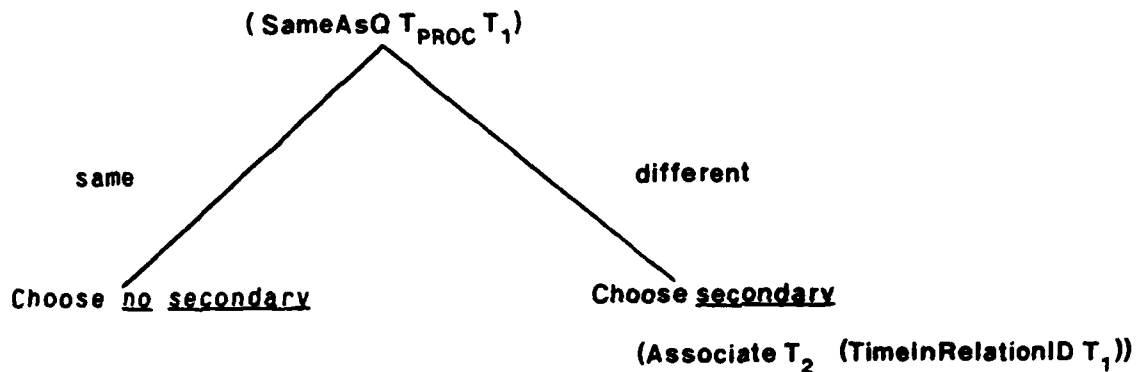


Figure 10: SECONDARY TENSE chooser

If the two times are different, the SECONDARY TENSE chooser chooses secondary, but in addition it asks a further non-branching question. This identifying inquiry, TimeInRelationID, identifies the time in relation to which T_1 is to be ordered.¹⁷ The hub symbol that constitutes the response is assigned to the grammatical function T_2 .

Chooser: Is STORY-TIME (T_1) identical to NURSING-TIME (T_{PROC})? (i.e., (SameAsQ T_1 , T_{PROC}))

Environment: No, they are different.

Chooser: Then I choose secondary.

Then, the chooser determines that STORY-TIME (T_1) is to be related to NURSING-TIME and this time concept is associated with T_2 .

To get a feeling for what progress we have made, look at the table in Figure 5 above: The choice of secondary has decreased the number of tense competitors; the column labeled no secondary has been excluded.

Once it has been established that there is a "secondary tense", the next step is to determine the type of tense to be expressed, i.e., to determine whether we have a future tense, a present tense, or a past tense. This is the task of SECONDARY TENSE TYPE (immediately to the right of SECONDARY TENSE in the diagram in Figure 3) and its chooser. Using our spatial metaphor, we can say that we have decided to take a step in time; now we have to decide on the direction of the step.

¹⁷ This time may be T_{PROC} and is indeed this time in our example. However, in more temporally complex examples, there may be an additional reference time that is different from T_1 , as in *Jane had been going to be a nurse*. For this reason, the SECONDARY TENSE chooser always asks for a symbol for this time and associates it with T_2 . As we will see, the TERTIARY TENSE chooser will check whether this is the same as T_{PROC} or not.

7. SECONDARY TENSE TYPE CHOOSER

The chooser of the SECONDARY TENSE TYPE system is represented by the diagram in Figure 11. The first chooser action is an inquiry of the branching type. Notice that the two parameters of the inquiry -- T_1 and T_2 -- have already been assigned hubs by earlier systems. There is thus a flow of hubs from one system to another.¹⁸

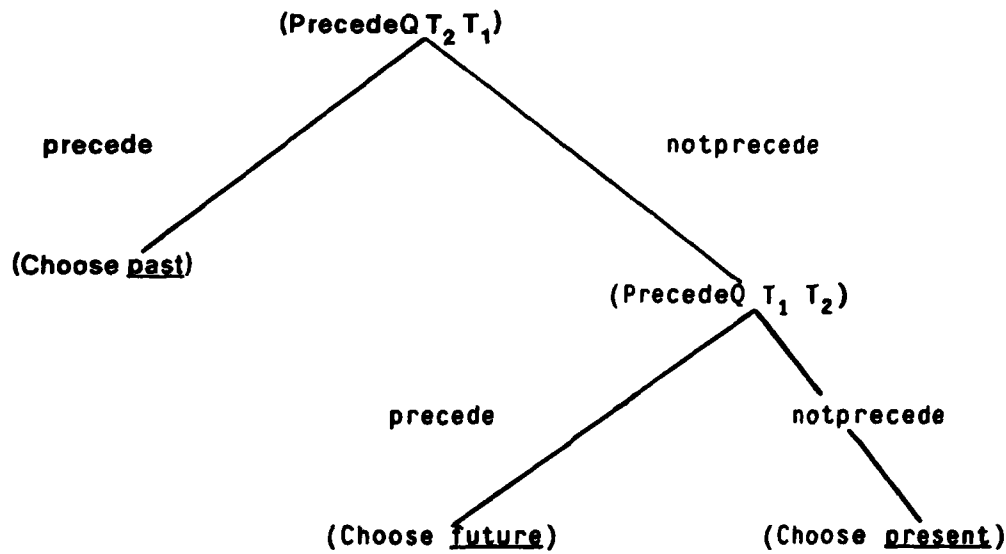


Figure 11: SECONDARY TENSE TYPE chooser

When the answer to (PrecedeQ $T_2 T_1$) is "precede", past (realized by *have*) is chosen. When the answer is "notprecede", an additional branching question is asked. The result for the structure of the chooser is that it has the form of a decision tree with two branching nodes.

The second inquiry is also PrecedeQ, but the order of the parameters has been changed: (PrecedeQ $T_1 T_2$). When the answer is "precede", the feature future (realized by *be going to*) is chosen. This is really the only reasonable outcome, since the inquiry is embedded under the answer "notprecede".¹⁹ With the present design, both choices, past and future, are the result of positive answers. If T_1 neither precedes nor follows T_2 , present is chosen.²⁰

¹⁸ If the two functions T_1 and T_2 had not been given hubs, the inquiry (PrecedeQ $T_1 T_2$) would have been meaningless at this point. In general, it is overall dependency in the network of the grammar of one system upon the features of other systems that ensures that the choosers of different systems create hubs in such a way that they feed other chooser questions and maintain an uninterrupted flow of hubs.

¹⁹ Precedence is here strict precedence.

²⁰ In other words, the so-called progressive is chosen for a tense reason (a relation between two times), rather than for an aspectual reason. Actually, it is probably necessary to add a further inquiry at this point in the chooser. There are fairly good reasons for asking whether T_2 includes T_1 , and choosing present if the answer is positive. The inclusion interpretation of the so-called progressive has a long tradition -- Sweet, Jespersen, etc. -- and has recently been argued and studied by e.g. [Schachter 81] and [Merrill 82]. I will not pursue the issue here.

In our example we want to express a time relation where the time of the "nursing event" precedes the story time, which corresponds to the hub for T_2 . So the following dialogue ensues:

Chooser: Does NURSING-TIME (T_2) precede STORY-TIME (T_1)?

Environment: Yes, it does.

Chooser: Then I choose past.

With the choice of past, we can again eliminate a section of the table in Figure 5. We have excluded the future column.

8. TERTIARY TENSE CHOOSER

The next system we reach is TERTIARY TENSE. Its chooser is quite parallel to that of SECONDARY TENSE.²¹ The purpose is to find out whether there is a need to express an additional time relation or not. Through the choice of secondary in the SECONDARY TENSE system, we have already expressed the pastness of the time of Jane's being a nurse with respect to the general time of the story. Consequently, there is no need to establish any further time relations:

Chooser: Is NURSING-TIME (T_2) identical to NURSING-TIME (T_{PROC})?

Environment: Yes, it is.

Chooser: Then I choose no tertiary.

The choice of no tertiary eliminates all of the boxes that could have been included in the table in Figure 5.²²

The choice of the feature no tertiary concludes the tense reasoning for our example. An independent set of processes work on the features chosen and re-express them as linguistic structure, yielding *had been*.

9. CONCLUSION

The table in Figure 12 shows how the problem of selecting an appropriate tense combination to express conceptual time relations has been decomposed by systemic grammar.

This decomposition into minimal choice points, represented by systems, has made it possible for us to formulate simple tense choosers that selected the right tense combination for our example step by step and together achieved the right result.

²¹This is what is to be expected, given Halliday's analysis of tense as linearly recursive; cf. section 3.1 above.

²²This feature is used to represent the second time the feature no secondary in the SECONDARY TENSE schema is instantiated.

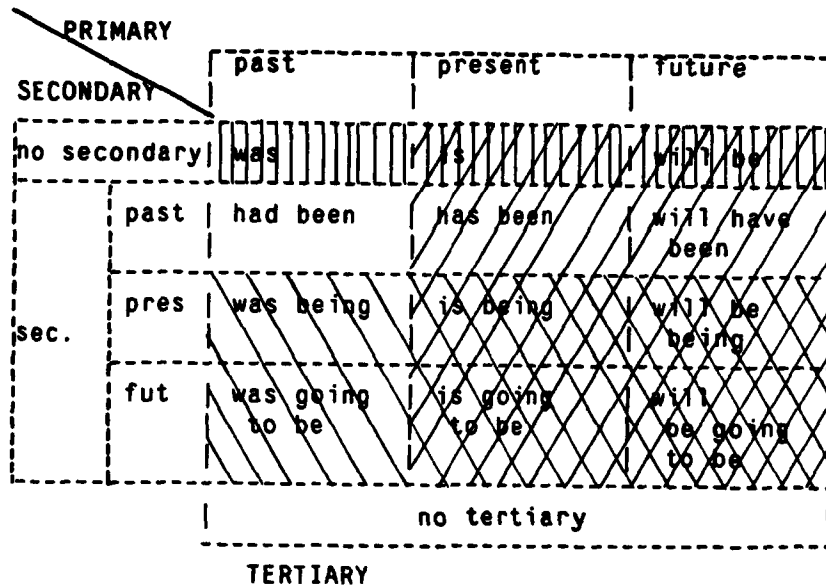


Figure 12: The steps in choosing tense for *had been*

The area of tense has served to introduce the chooser framework and point to some of its important properties like the decomposition of the decision process and the flow of hubs from one independent chooser to another.

The general question about how to make grammatical choices was mapped onto a more specific version of it about tense. However, choosers are not restricted to areas similar to tense. Experience clearly shows that the chooser approach is quite general across different areas of grammar. This leads us to believe that the chooser framework goes quite some way towards solving the problem that our initial question "How are grammatical choices made as part text generation?" raises. An important task is now to increase our understanding of English semantics by extending detailed coverage to diverse areas of English grammar.

REFERENCES

- [Davey 79] Davey, A., *Discourse Production*, Edinburgh University Press, Edinburgh, 1979.
- [Halliday 69] Halliday, M.A.K., "Options and functions in the English clause," *Brno Studies in English* 8, 1969, 81-88.
- [Halliday 76] Halliday, M.A.K., "The English verbal group," in Kress, G. (ed.), *Halliday: System and Function in Language*; Oxford University Press, 1976. Originally circulated in 1966
- [Halliday 82] Halliday, M.A.K., *A Short Introduction to Functional Grammar*, To be published by Batsford, 1982.
- [Halliday 1978 78] M.A.K. Halliday, *Language as social semiotic*, Edward Arnold, 1978.
- [Mann 82a] Mann, W. C., et al., *Text Generation: The State of the Art and the Literature*, USC/Information Sciences Institute, RR-81-101, April-June 1982 . Appeared as "Text Generation" in April-June 1982 AJCL
- [Mann 82b] Mann, William C., *The Anatomy of a Systemic Choice*, USC/Information Sciences Institute, Technical Report RR-82-104, October 1982. To appear in *Discourse Processes*.
- [Mann 83] William C. Mann, "An Overview of the Penman Text Generation System," in *Proceedings of the National Conference on Artificial Intelligence*, pp. 261-265 , AAAI, August 1983 . Also appeared as Information Sciences Institute report ISI/RR-83-114 .
- [Matthiessen 83a] Matthiessen, C., "Choosing Primary Tense in English," *Studies in Language* 7, 1983.
- [Matthiessen 83b] Matthiessen, C., "The systemic framework in text generation: Nigel," in W. Greaves & J. Benson (eds.), *Systemic Perspectives on Discourse*, Ablex, 1983.
- [Merril 82] Merrill, P., *Telicity and the Progressive in English*, 1982. UCLA MA thesis
- [Reichenbach 47] Reichenbach, H., *Elements of Symbolic Logic*, Macmillan, 1947.
- [Schachter 81] Schachter, P., *Explaining Auxiliary Order*, Indiana University Linguistics Club, 1981.
- [Winograd 72] T. Winograd, *Understanding Natural Language*, Academic Press, 1972.
- [Winograd 82] Winograd, T., *Language as a Cognitive Process. Syntax*, Addison-Wesley Publishing Company, Inc., 1982.

END

FILMED

5-84

DTIC