INTEGRATED DEVELOPMENT OF ENGLISH-SPANISH MACHINE TRANSLATION: FROM PILOT TO FULL OPERATIONAL CAPABILITY

TECHNICAL REPORT

Grant DPE-5543-G-SS-3048-00
from the U.S. Agency for International Development

Marjorie León and Lee A. Schwartz

Coordinated by Muriel Vasconcellos, Co-principal Investigator

October 1986

PAN AMERICAN HEALTH ORGANIZATION
Pan American Sanitary Bureau - Regional Office of the
WORLD HEALTH ORGANIZATION
525 Twenty-third Street, N.W.
Washington, D.C. 20037

TABLE OF CONTENTS

FIGURES

INTEGRATED DEVELOPMENT OF ENGLISH-SPANISH MACHINE TRANSLATION: FROM PILOT TO FULL OPERATIONAL CAPABILITY

Marjorie León[1] and Lee A. Schwartz[2]

1. INTRODUCTION

The Pan American Health Organization (PAHO), with headquarters in Washington, D.C., is a specialized agency which belongs to both the Inter-American and the United Nations systems. In the UN system, PAHO serves as Regional Office of the World Health Organization.

The four official languages of the Organization are Spanish, English, Portuguese, and French, but most documents are produced only in the working languages, Spanish and English. For this reason, development efforts in machine translation (MT) have concentrated on this pair of languages. The Organization's first MT system, SPANAM™,[3] which translates from Spanish into English, began to be used for production in 1980.

In late 1982 attention was turned to ENGSPAN™, the program for translation from English into Spanish. In August 1983 the U.S. Agency for International Development (AID) awarded the Organization a two-year grant for the accelerated development of ENGSPAN.[4] During the grant period, particular focus was placed on the introduction of new syntactic and semantic codes in the English dictionary, the building of a parser using an augmented transition network (ATN), and the development of techniques for synthesis of the Spanish

---

[1]Senior computational linguist, Machine Translation Program, Pan American Health Organization; with Muriel Vasconcellos, co-principal investigator under Grant DPE-5542-G-SS-3048-00 (see note 4 below).

[2]Computational linguist, Machine Translation Program, Pan American Health Organization.

[3]SPANAM™ and ENGSPAN™ are trademarks of the Pan American Health Organization.

[4]Grant DPE-5542-G-SS-3048-00, awarded to the Pan American Health Organization under letter dated 3 August 1983. Grant period extended through November 15, 1985. The funds supplied by the Program in Science and Technology Cooperation of AID covered approximately 40% of the development costs incurred during the grant period. The remaining portion continued to be funded by PAHO. The work described in the present eport represents the combined investment of the two institutions.

output. As of the end of the grant period, ENGSPAN had translated 612,973 words (2,452 pages) including both production texts and research corpora. The ENGSPAN dictionaries had 45,614 source entries and 47,545 target entries.

1.1 Hardware and Software Configuration

The translation programs were developed on PAHO's mainframe computer, currently an IBM 4381 running under DOS/VSE/SP. A version of the program is also installed and running on AID's IBM 3081 (OS/MVS).

The translation programs and supporting software are written in PL/1. ENGSPAN runs under VSE with a size parameter of 844K. The dictionaries are accessed using VSAM and require about 8 Mb each of disk storage. Production is in batch mode only. The speed of translation ranges from 300 to 600 wpm in clock time, depending on the level of utilization of the PAHO mainframe. In CPU time, the speed is approximately 1,400 wpm.

Word-processing workstations (Wang OIS/140) serve as remote job-entry terminals for text input and output. The source texts come from the regular flow of documentation in the Organization. There are no restrictions as to field of discourse or type of syntax, and the translation staff does not pre-edit a text before it is processed by the computer. On the other hand, all machine-translated output is postedited by a professional translator before it is delivered to the end user.[5]

1.2 Approach and Scope

ENGSPAN is based on the transfer approach to machine translation. It analyzes the English source string, applies transfer rules based on the contrastive analysis of English and Spanish, and then synthesizes the Spanish target string. The system is designed to provide comprehensive lexical and syntactic coverage, rather than an in-depth semantic analysis or a representation of knowledge.

Figure 1 shows the basic components of the system. The permanent data structures include the dictionaries and the grammar. After text input and segmentation, information is gathered about each lexical item through morphological analysis and from records retrieved from the source dictionary. The parser uses this lexical information and the system's ATN grammar to produce a structural analysis of the input string. When necessary, safety-net procedures are used to supplement the results of the parser. The structure of the target output string is determined by several different types of syntactic

---

[5]See Vasconcellos and León (1985) for further discussion of the history of the project and the Organization's use of its MT systems.

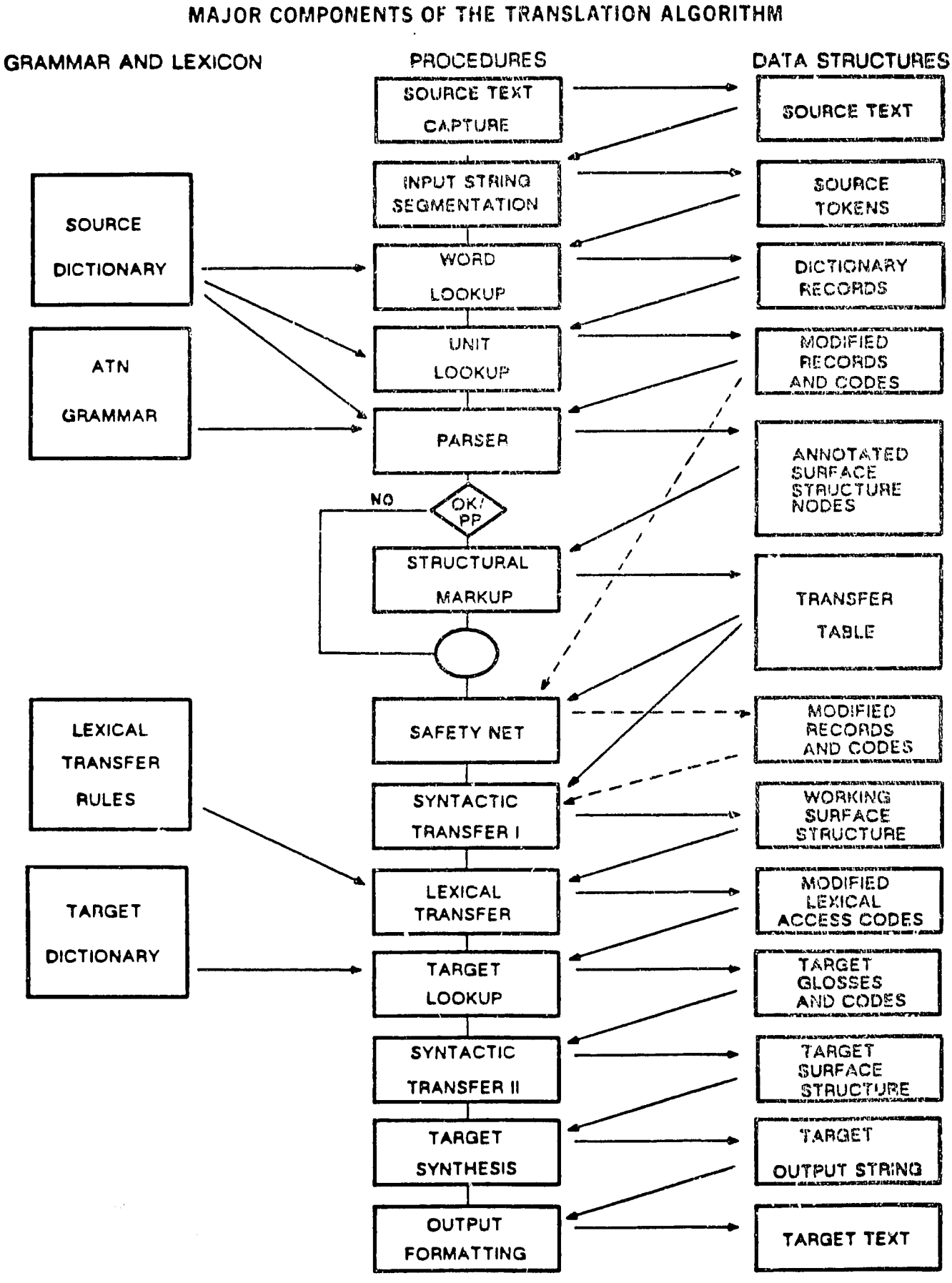


Figure 1. Major components of the translation algorithm.

transfer rules; lexical transfer rules are used to select appropriate glosses from the target dictionary. Finally, the surface form of each Spanish word is synthesized and the output text is formatted.

ENGSPAN has proven itself as a general-purpose MT system that can produce a usable translation of random text in a broad range of subject areas. The "robustness" of the system is achieved through combination of the following capabilities and features:

- Morphological analysis of inflected and derived words
- Gap analysis for not-found words
- Broad parsing coverage
- Safety-net analysis procedures
- Semantic generality of main target glosses
- Generality of transfer rules
- Dynamic synthesis rules

ENGSPAN can also become a domain expert--a source of authoritative terminology. This capability, of course, depends on the content of the dictionaries. Through the creative use of the various types of possible dictionary entries, the translator, terminologist, or dictionary coder can customize the system and fine-tune the output to meet highly specific needs.

1.3 Focus and Structure of the Report

This report provides an overview of the major components and capabilities of ENGSPAN, concentrating on those portions of the system which were developed during the grant period. Particular attention is given to the system's most unique feature, an ATN parser written in PL/1.

Chapter 2 gives a summary of the project's status at the beginning of the AID grant period. Chapter 3 outlines the organization of the dictionaries and the different types of dictionary entries. Chapter 4 describes the capabilities of the morphological analysis procedure. Chapter 5 is devoted to a detailed description of the parser and a step-by-step explanation of the parsing of a demonstration sentence. Chapter 6 discusses the backup analysis strategy known as the "safety net." Chapter 7 describes the functions of the transfer component and provides examples of the lexical transfer capabilities of the system. Chapter 8 describes the synthesis component and contains examples of Spanish output which depend on the system's syntactic transfer rules and synthesis capabilities. Chapter 9 identifies some of the areas for which future enhancements are planned.

2. STATUS AT THE BEGINNING OF THE GRANT PERIOD

At the time PAHO received the grant from AID, a pilot version of ENGSPAN was already being tested.[6] The SPANAM dictionaries had been reversed automatically and reviewed by the project staff. A corpus of over 50,000 words had been selected and the not-found words had been added to the dictionaries. The source dictionary contained 40,642 entries and the target dictionary had 41,979 glosses.

ENGSPAN and SPANAM use the same support programs to update and display the dictionaries. LEXDATE interprets mnemonically coded instructions and performs additions, changes, and deletions of dictionary entries. LEXVIEW produces a mnemonic side-by-side display of the entries for individual lexical items or groups of items which may be selected by range or by Boolean combinations of dictionary codes. These two programs were modified during the grant period to permit the coding of new types of entries and to incorporate descriptors for new dictionary codes.

A third support program, TEXTREP, converts the text received from the word processor into the format required by the translation program. The functions performed by TEXTREP include deblanking; identification of titles, sentence boundaries, and paragraph boundaries; reconcatenation of words broken by a facultative hyphen; separation of leading and trailing punctuation marks; marking of capitalization; and conversion of diacritics and special characters. This program remained basically unchanged during the grant period.

The pilot version of ENGSPAN included several modules which were originally written for SPANAM. ENGINIT sets the runtime parameters and initializes the high frequency dictionaries. MTSIOW contains both input and output procedures: it fetches the input strings to be translated and divides them into tokens, and it prepares the output text in one of three optional formats.

Another module which was originally incorporated "as is" from SPANAM is FINDUNIT, the procedure for retrieval of substitution units from the source dictionary. This procedure underwent substantial enhancement during the grant period.

The eight modules listed below, written specifically for ENGSPAN, were part of the pilot version of the program:

LEMMA	Morphological analysis of English verbs and nouns

---

[6] See León (1984) for a description of the initial stages in the development of ENGSPAN.

LOOKUP	Retrieval of lexical items from the source dictionary
VSTRING	Analysis of simple verb strings
POSAMBIG	Resclution of certain types of homographs
TLOOKUP	Retrieval of glosses from the target dictionary
NSYN	Synthesis of inflected forms of determiners, numeratives, adjectives, and nouns
VERBSYN	Synthesis of inflected forms of verbs
TARGPAT	Rearrangement of simple noun phrases and triggering of gender and number agreement for common sequences of modifiers.

With the exception of LEMMA and VERBSYN, all the modules listed above were modified extensively during the grant period. TARGPAT was replaced completely. Nevertheless, the pilot version of ENGSPAN provided the basic structure necessary for development and testing of the parser and contributed directly to implementation of the safety-net strategy. The output produced by early versions of the program was useful in the design of the parser and led to the development of new types of dictionary entries and new strategies for synthesis.

3. THE DICTIONARIES

The subject-area coverage of the system is determined by the content of the dictionaries. ENGSPAN's English source dictionary contains approximately 20,000 items of general vocabulary and 25,000 terms from the fields of medicine and public health. Any of these lexical items may have target glosses pertaining to different subject areas, reflecting different types of discourse, or representing the preference of a particular user.

3.1 Structure of the Dictionaries

The basic record has a fixed length of 161 bytes. The record is divided into 211 fields, the majority of which contain syntactic codes; others contain morphological and semantic codes. Several fields are also used to store information about the source and reliability of the entry and its relationship to other entries. Every lexical item is coded for at least one part of speech. Each part-of-speech entry takes a set of additional codes. When a lexical item in the source dictionary can function as more than one part of speech, the coding pertaining to its different functions is entered in the same record. This produces one conflated entry in which all the information about the lexical item is combined. When a word has many possible uses, the dictionary coder has the option of including only those that are most likely to occur, based on past experience with the type of texts to be translated. The inclusion of many infrequent or obscure possibilities can make it more difficult for the program to arrive at the correct analysis.

The source and the target dictionary entries are stored in separate files. The source entries are linked to their respective target glosses through a lexical number (LEX) which is assigned to each new pair as it is added to the dictionary. The first six digits of the 12-digit LEX constitute a unique identification number for a set of entries related to the same lexical item. The set of entries in the target dictionary can be viewed as a three-dimensional matrix (Figure 2). The first dimension corresponds to the part of speech. For each part of speech there may be many different context-sensitive glosses. These constitute the second dimension. Finally, in the third dimension, each alternative gloss can have specialized translations belonging to one or more microglossary. The last six digits of LEX serve as pointers to various types of sub-entries. Each slot in the matrix represents a potential dictionary entry, but storage is allocated only when the record is actually added to the file. The maximum number of target glosses that could be specified for a single source entry, if every slot in the matrix were used, is one million. In practice, of course, the program is not equipped with enough rules to distinguish among so many possibilities. At the present stage of implementation, the system utilizes approximately 4,096 different slots (16 positions on each vector), depending on the relationships among the entries. Some sets of entries have as many as 15 different glosses. For example, the dictionary contains the following entries related to the English source word meet:

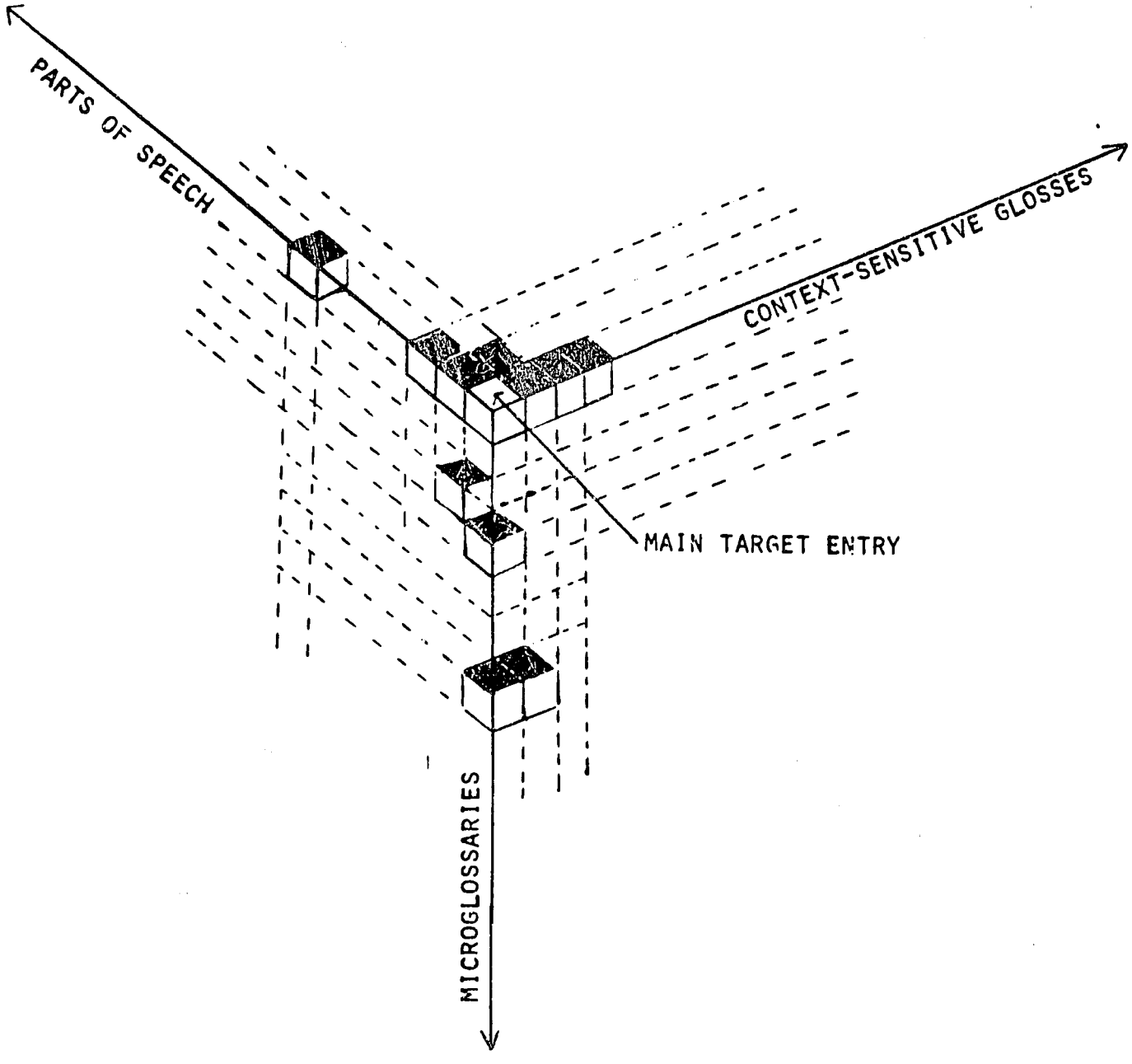


Figure 2. Array of possible entries in the target dictionary corresponding to one lexical item in the source dictionary.

meet (verb)	reun[ir]
meeting (noun)	reuni6n
met (verb)	sesi6n
	solucion[ar]
	satisfac[er]
	encontr[ar]
	encuentro
	confluencia
	atend[er]
	alcanz[ar]
	cont[ar]
	llen[ar]
	cumpl[ir]
	conoc[er]
	respond[er]

3.2 Coding of the Dictionary Entries

The basic set of dictionary codes was defined during the development of the pilot version of ENGSPAN. It reflects elements of both practical experience and linguistic theory. One of the references used for the coding of English words was Sager (1981). Many of the codes described below are used in both dictionaries; others apply to only the source or only the target.

3.2.1 Verbs

Verbs are coded for tense, person, conjugation class. irregular conjugation features, and the use of reflexive pronouns. Each verb is also coded for transitivity, phrasal and clausal complementation, and bound prepositions and adjuncts. Subject and object preferences can be specified as ±Human, ±Animate, and ±Concrete.

3.2.2 Nouns

Nouns are coded for number, gender, and irregular plural formation. Features which can be coded for nouns include Count, Bulk, Concrete, Human, Animate, Feminine, Proper, Collective, Device, Location, Time, Quantity, Scale, Color, Nationality, Material, Apposition (noun, infinitive, or clause), Verb nominalization, Body part, Condition, and Treatment. Nouns can also be specified for prepositional government.

3.2.3 Adjectives

Adjectives are coded for number and gender as well as for many of the noun features mentioned above. In addition, they can be coded Inflectable, Optionally inflectable, Comparable, General, Temporary, Resultative, and Tough-movement.

3.2.4 Adverbs

Like adjectives, adverbs are coded Inflectable, Optionally inflectable, and Comparable. Adverbs are also coded Time, Place, Manner, Interruptive, and Connector.

3.2.5 Function words

Other parts of speech belong to closed classes of function words. These include prepositions, conjunctions, determiners, numeratives, modifiers, adjuncts, prefixes, and pronouns. In ENGSPAN, these categories are further subdivided and coded in order to make distinctions which are required for parsing and for target synthesis.

3.3 Types of Entries

The source dictionary contains five different types of entries, all of which are stored in the same file. These are single words, substitution units (SU), analysis units (AU), delayed substitution units (DSU), and transfer units (TU). SUs and AUs are activated when the specified lexical items occur consecutively in the sentence. DSUs and TUs can cover either a continuous or a discontinuous string. The functions of each of these types of dictionary entries are discussed below.

3.3.1 Single-word entries

The basic dictionary entry contains a single lexical item. An item in the source dictionary may be up to 30 characters in length. The maximum length of an item in the target dictionary is 99 characters.

Words may be entered in the dictionary either with or without inflectional endings. Most nouns are entered only in the singular and adjectives only in the masculine singular. Verbs are entered as stems. Full-form entries are used for auxiliary verbs, words with highly irregular morphology, and homographs.

Several source entries--for example irregular forms of the same verb or alternate spellings of a word--may be linked to the same target gloss by assigning them the same LEX. Conversely, it is also by means of the LEX that a source word can be associated with more than one gloss in the target.

3.3.2 Substitution units

The basic SU contains from two to five words. When an SU is retrieved by the phrase lookup procedure, the SU record replaces the individual records corresponding to the words included in the unit. The LEX number in the SU record is linked to a single record in the target dictionary which contains the translation of the entire phrase. This type of unit is used for obtain-

ing the correct translation of names of organizations, titles of publications, slogans, etc., and it is also an efficient way of handling some fixed idioms, phrasal prepositions, and certain technical terminology. Examples of phrases which are entered as SUs are United Nations Children's Fund, Committee of the Whole, foot-and-mouth disease, sweet potato, live birth, in lieu of, close at hand, and by leaps and bounds.

When a longer unit is required, such as for names of organizations and titles of publications, a maximum of 25 words can be grouped together by nesting one SU inside another.

3.3.3 Analysis units

The AU, which also contains from two to five words, has several functions. It can be used to alert the analysis routines to the possible presence of a common collocation and provide information on its length and function. It can also be used to resolve the part-of-speech ambiguity of any of its members or to respecify the part of speech of the respective members. Finally, an AU can specify an alternative translation for one or more of its members.

When an AU is retrieved by the phrase lookup procedure, the record for each source word is retained in the representation of the sentence, but additional information is copied from the AU into the appropriate individual records. When the target lookup is performed, the gloss for each word is retrieved separately.

The use of AUs helps in the analysis and synthesis of conjoined modifiers and makes it possible for the parser to override the phrasal analysis when necessary. For example, by entering overseas post as an AU, the parts of speech can be specified as adjective and noun, thus eliminating the ambiguities of overseas as an adverb and post as a verb or prefix. In the AU for summary record, the parts of speech are specified, as well as an alternative gloss 'acta' for record. The AU for pension benefit indicates that the words are likely to function as a single noun phrase, while at the same time it also provides for an alternative gloss for one of its members ('prestación' for benefit). It does not, however, eliminate other possible analyses. The parser is still able to correctly analyze sequences such as with this type of pension benefits are higher and an increase in pension benefits everyone.

3.3.4 Delayed substitution units

The DSU is used to handle lexical items such as phrasal verbs (e.g. pay back, slow down, and wipe out) which can occur as noncontiguous words in the input. The unit is retrieved from the dictionary when the parser is analyzing the sentence. The syntactic and semantic codes in the DSU are compared with those in the individual records and the unit is accepted if it satisfies the conditions being tested by the parser. If the unit is accepted, it replaces the individual records and eventually triggers a different target gloss.

3.3.5 Transfer units

The TU is a rule which triggers a context-sensitive gloss for one or more words, based on the functions of those words as determined by the analysis component of the program. The rules are stored in the source dictionary and are retrieved after the analysis has been completed. If the conditions specified in the transfer rule are met, the desired target gloss is selected during the target lookup. For example, if the object of raise is coded as +Anim, the verb is translated as 'criar' instead of 'subir'.

TUs are discussed in detail in Section 7.2.

3.4 Microglossaries

A microglossary is a subset of dictionary entries which can be selected for a particular subject area, discourse register, or a specific user. Microglossary entries may be used in both the source and the target dictionaries. For PAHO's purposes, glosses pertaining to the subject area of international public health are usually the main target entry. Microglossaries are used to obtain special translations of terms in the fields of law, finance, sanitary engineering, agriculture, computer science, and biomedical research. The system design allows for up to 99 microglossaries with any number of entries in each one. (So far, only 16 microglossaries have been implemented.) The microglossaries to be consulted during the translation of a particular text are specified at run time. If more than one microglossary needs to be activated for the same translation, they are specified and consulted in order of priority.

4. MORPHOLOGICAL ANALYSIS

The dictionaries are supplemented by a resourceful morphological analysis component which can reduce a highly inflected word to a canonical form that has been entered in the lexicon. If the canonical form is not found, a gap analysis procedure checks for prefixes, suffixes, and orthographic clues and provides the syntactic analysis component with information about the probable functions of the word, thus making it possible to analyze a sentence which contains unknown words.

Capitalization and punctuation help to identify proper names, taxonomic names, place names, and alphanumeric ordinal and cardinal numbers. Percentages, temperatures, amounts of money, and dates receive special handling. The LEMMA procedure removes possessive endings, plural noun endings, comparative and superlative suffixes, the adverbial suffix, and verb inflections. Another procedure checks a list of prefixes. In the final phase of the lookup process, the word is checked for possible variations in spelling between American and British English.

Contractions and hyphenated or slashed compounds are separated, and each part is looked up individually, eliminating the need to have additional dictionary entries for every combination. If necessary, a compound can be entered as a phrase.

As a result of the lookup strategy used in ENGSPAN, more than 99 percent of the lexical items in a typical random text can be found in the source dictionary. It then becomes the job of the synthesis component to produce a meaning-preserving translation for a lemmatized form. The translator, of course, can always enter an inflected form or a compound in the dictionary in order to bypass the "creativity" of the program. For example, 'repetidamente' is an acceptable translation for repeatedly, and 'no dando vuelta' could fill in temporarily for unwinding, but unlikely needs its own full-form entry in the dictionary in order to avoid the system-generated translation, '*no similarmente'.

Examples of the synthesis of hyphenated compounds are given in Section 7.2.5 and 8.3.1.

5. THE PARSER

ENGSPAN uses an augmented transition network (ATN) parser, written in PL/I, to analyze its English input. In the initial stages of the development of the parser, a study was made of the ATNs discussed in Woods (1970), Bates (1978), and Winograd (1983). In view of the options presented therein, the parser was designed to perform a top-down, left-to-right sequential parse with both chronological and explicit backtracking, employing three basic data structures--a pushdown stack, a hold list, and a well-formed phrase list. The pushdown stack directs the parser through the network grammar; the hold list assists in the analysis of sentences containing extraposed material; and the well-formed phrase list allows for the saving and reuse of information generated during the analysis.

This chapter of the report is divided into six sections. Section 1 describes the ATN grammar; section 2, the output from the parser; section 3, the parsing process itself; section 4, additional facilities of the parser; section 5, the syntactic coverage of the parser; and section 6, future work to be done in the field of semantics. In section 2 the actual output structures produced by the parser during the analysis of one sentence are presented with explanation, and in section 3 an arc-by-arc account is given of how the parser goes about analyzing this same sentence. For the purpose of this presentation, a simplified version of the actual network grammar, or a "baby grammar," to use the terminology of Sager (1981), is used. This "baby grammar" contains most of the basic arcs needed to parse a declarative sentence, but it contains substantially fewer arcs than does the actual network because it does not provide for all possible distinctions within syntactic categories (see Appendix A).

The demonstration sentence is:

(1) It is clear that chronic malnutrition and hunger are linked to more general problems of poverty, inequitable distribution of land and other assets, underdevelopment, and adverse agroclimatic conditions.

This sentence was taken from a text translated by ENGSPAN. Three factors were considered in choosing it: its length, its complexity, and the use it makes of the different capabilities of the parser. In terms of length, it is a moderate-to-long sentence for English; in terms of complexity, it is fairly complex, consisting of more than one clause and containing several words which can function as different parts of speech in different contexts; and in terms of its use of the parser's capabilities, it employs those for backtracking, both explicit and chronological, as well as those for saving and reusing well-formed phrases.

5.1 The Grammar

ENGSPAN's ATN grammar resides outside the parsing algorithm and takes the form of a linked list of 11 networks, containing a total of 333 arcs, with which 141 conditions and 60 actions are associated (figures as of 30 October 1985). It differs most notably from the grammars presented in the sources cited above in its form, its separation of the clause and verb phrase networks, and its inclusion of conjunction arcs for the analysis of coordinate constructions.

The grammar is embodied in an augmented transition network which is not a single network but rather a system of subnetworks. The full version of the grammar contains subnetworks for the parsing of sentences, clauses, noun phrases, verb phrases, prepositional phrases, nonfinite verb phrases, hyphenated compounds, relative clauses, dependent clauses, adverbial phrases, and comparative clauses.

5.1.1 Arcs

Each subnetwork consists of a series of states connected by arcs. Associated with each arc is a label indicating the type of arc, and variables indicating the state at which the arc terminates, the condition that must be satisfied before the arc is traversed, and the action which must be performed after it is traversed. There are four types of arcs: category arcs, jump arcs, seek arcs, and send arcs. The category arc is the only arc the traversal of which results in the consumption of the current element of the input string. When the part of speech of the current word corresponds to the part of speech indicated by the category of the arc, the category arc can be traversed. The other arcs, i.e. the guiding arcs, direct the parser to another point in the network system, from which it is to continue processing. The jump arc directs the parser to another state within the same subnetwork; the seek arc signals the parser to make a recursive call to another subnetwork; and the send arc returns the parser from a recursive call.

5.1.2 Conditions

A condition can test any combination of the following types of information:

- The status of the switches set in the program;
- The lexical form of the current word, any word which has already been parsed, or a word which is within the look-ahead range (usually limited to one or two words);
- The features of the current word, any word which has already been parsed, or a word which is within the look-ahead range, as derived from the dictionary lookup and lemmatization routines;

- The characteristics and/or components of the phrase that is currently being parsed or of any phrase that has already been parsed;

- The status of the hold list.

Examples of portions of actual conditions which make the tests described above are presented below in pseudocode. "OK" means that the condition is satisfied; "NO," that it is not.

- Condition on the send arc from the clause network: If the semi-, list-, colon-, or title- switch is on and the current word is an end-of-string marker, then OK;

- Condition on a preposition arc in the nonfinite verb phrase network: If the current word is to or for, then OK;

- Condition on the head noun arc in the noun phrase network: If the current word is a singular count noun and a determiner has been parsed for the current phrase, then OK;

- Condition on the NP arc following the VP arc in the clause network: If a subject has not yet been parsed for the current clause, then OK;

- Condition on the send arc in the sentence network: If the hold list is empty, then OK.

5.1.3 Actions

An action can perform any combination of the following functions:

- Enter new information into a phrase information structure;

- Modify information in a phrase information structure;

- Examine the configurations of active phrase information structures and put the parser in a backtracking mode if they are not acceptable.

Examples of actions that perform these functions:

- Action on head noun arc in the noun phrase network: Assign the role H (for head word) to the word just parsed;

- Action on second object NP arc in the verb phrase network: If the last element in the phrase information structure has the role O, change it to I;

- Action on the send arc from a phrase which is not part of a coordinate construction: If the first word in the phrase was parsed as a correlative conjunction, backtrack to that word.

5.2 The Output

One of the most distinctive features of the parser is its output, which takes the form of three types of data structures. Two of these emphasize ambiguity resolution and one, structure. The most detailed form of output--that which represents the structure of the sentence--is a series of annotated phrase structures similar in content to those found in Winograd (1983).

The three types of output are: the side-by-side display, VERTPAT, and the phrase information structures.

5.2.1 Side-by-side display

The side-by-side display presents the input string on the left, all the possible parts of speech of each element in that string on the right, and a message indicating the success of the parse in the middle. The display generated for the demonstration sentence appears as Figure 3. This display was produced by a version of ENGSPAN which is used for parser development and which does not call the synthesis modules. The same type of output is produced when these modules are called, but in the place of the code display, the translation appears.

The input string needs no explanation; it is the sentence as formatted by the conversion program, TEXTREP. The message column is the parser's report card, noting the success of the parse. If the parser arrives at an analysis of the entire input string, "OK" appears in the message column. If it has not been able to analyze the string, either because its grammar is not equipped to do so, or because the miscoding of one or more words has blocked the analysis, the message is "NO." If it has run out of space (a limit is put on the number of information structures that the parser is allowed to create in the course of its work), the message is "TP" (for partial parse).

The code display gives the codes for all the possible parts of speech of each element in the input string, as derived from the dictionary-lookup and lemmatization routines. Just as the elements of the input string on the left side of the display are set off from one another by blanks, so are their corresponding sets of part-of-speech codes on the right. (For increased legibility, the number of the word to which each set of codes pertains is given either above or below it in Figure 3.) When there is more than one code for a given element, the first code identifies the part of speech that the parser has assigned to that element in the context in which it occurs in the input string. The code display shows all the part-of-speech ambiguities that the parser has faced in the course of its work and

Side-by-Side Display

Input String	Message Column	Code Display
It is clear that chronic malnutrition and hunger are linked to more general problems of poverty, inequitable distribution of land and other assets, underdevelopment, and adverse agroclimatic conditions.	OK	Word Numbers 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 r_ v_x a_v s_rt a_ n_ c_ n_ v_x a_v p_ m_k a_n n_ p_ n_ , a_ n_ p_ n_v c_ t_a n_ , n_ , c_ a_ a_ n_v. 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 Word Numbers

KEY

noun=n, verb=v, preposition=p, auxiliary=x, adjective=a, coordinate conjunction=c, subordinate conjunction=s, determiner=t, numerative=k, pronoun=r, modifier=m

Figure 3

how it has resolved them.

An examination of the code display will show that 10 words in the input string came into the parser from the lemmatization and dictionary-lookup routines with more than one part of speech. These words are given below with their corresponding parts of speech:

Word number	Lexical item	Parts of speech
2	is	verb, auxiliary
3	clear	adjective, verb
4	that	subordinate conjunction, pronoun, determiner
9	are	auxiliary, verb
10	linked	verb, adjective
12	more	modifier, numerative
13	general	adjective, noun
21	land	noun, verb
23	other	determiner, adjective
31	conditions	noun, verb

5.2.2 VERTPAT

VERTPAT is produced solely for the benefit of the synthesis modules; it is not used by the parser in the course of its analysis. This data structure takes the form of an array which contains the parts of speech which have been determined by the parser for each element of the input string that it has been able to analyze. The VERTPAT output for the demonstration sentence appears in mnemonic form below. The codes therein are the same as those appearing on the right side of the side-by-side display and the same as those appearing in the phrase information structures.

VERTPAT ON COMPLETE PARSE:

R	V	A	S	A	N	C	N	V	A
P	M	A	N	P	N	,	A	N	P
N	C	T	N	,	N	,	C	A	A
N	.								

Key: A=adjective, C=coordinate conjunction, D=adverb, G=not (negative), M=modifier, N=noun, P=preposition, R=pronoun, S=subordinate conjunction, T=determiner, V=verb.

5.2.3 Phrase Information Structures

All the information in VERTPAT, and much more, is found in the phrase information structures. A phrase information (PHINFO) structure is created

each time the parser begins the analysis of a new phrase. Each PHINFO structure is printed out as a table. PHINFO structures 18 and 19, produced during the parse of the phrase land and other assets, are reproduced below and are used in the following discussion to demonstrate the components of the PHINFO structure.

```
NP  PHRASE 18  COMPLETE
CALLED BY PHRASE 17
CATEGORY     ROLE     LOCATION  MODIFIES
   N          H          21                LAND
   C                     22                AND
   NP                    19
HEAD= LAND
NUMBER= PL
CONHS= 21,24

NP  PHRASE 19  COMPLETE
CALLED BY PHRASE 18
CATEGORY     ROLE     LOCATION  MODIFIES
   T          T          23        24      OTHER
   N          H          24                ASSET
HEAD= ASSET
NUMBER= PL
CONHS= 24
```

The PHINFO structures for the demonstration sentence are listed in Appendix B.

The first line of the PHINFO table reports on the type of the phrase, its number, and its status. In structure 18 this line reads as follows:

Type of phrase	Number of phrase	Status messages
NP	PHRASE 18	COMPLETE

Each phrase type corresponds to a network type. The following types appear in the output for the demonstration sentence: ST (sentence), CL (clause), NP (noun phrase), VP (verb phrase), PP (prepositional phrase), and DC (dependent clause). The PHINFO structure being discussed here was produced for a noun phrase. The number of the phrase, 18, is the number that the parser assigned to the structure when it was created.

There are several flags associated with the PHINFO structure which indicate the status of the phrase represented by that structure. These flags include COMPLETE, REUSED, POPFLAG, and RNR.

Flag	Status
COMPLETE	The phrase has been successfully parsed.

Flag	Status
REUSED	The PHINFO structure in question was put on the well-formed phrase list at one point in the course of the parse and was picked up and reused at a later point (after a misparse of an earlier portion of the input string had been corrected).
POPFLAG	The phrase in question is inactive, i.e. it is on the well-formed phrase list.
RNR	A constituent has been raised out of the phrase (right node raising has taken place).

The second line of the PHINFO structure gives the number of the calling phrase. For structure 19 this line reads:

CALLED BY PHRASE 18

A phrase X is a calling phrase of a phrase Y, if X was the phrase that was being parsed when a seek arc was matched and, as a result, the parse of Y was begun. The parse of phrase 19 was the direct result of the matching of a seek arc in the NP network during the parse of phrase 18. The only PHINFO structure in which no information is given about the calling phrase is the first one, that for the sentence (ST). When the parser begins its analysis of an input string it immediately creates a ST PHINFO structure. All other phrase information structures are created only as a result of a seek arc having been matched.

The body of the PHINFO table is yet another table. This table contains information about the elements that make up the phrase--their parts of speech, their roles in the phrase, and their participation, if any, in modification relationships. The element table for phrase 19 is as follows:

CATEGORY	ROLE	LOCATION	MODIFIES	
T	T	23	24	OTHER
N	H	24		ASSET

The number of lines in the body of the element table corresponds to the number of elements in the phrase. The NP which is represented by the PHINFO structure given above is made up of two elements, in this case two lexical items. The first element is a determiner (T); its role in the phrase, not surprisingly, is that of determiner (T); it is word number 23 in the input string; it modifies word number 24; and its base lexical form is other. The second item in the phrase is word number 24; it is a noun which functions as the head (H) of the phrase; it does not modify any element of the phrase; and its base lexical form is asset.

There are no phrasal elements in this noun phrase, though there are in its calling phrase, phrase 18:

CATEGORY	ROLE	LOCATION	MODIFIES
N	H	21	LAND
C		22	AND
NP		19	

The phrasal element here is NP 19. No lexical items are associated with this phrasal element in the information structure. The associated lexical items can be found in phrase information structure 19.

Additional information about the phrase corresponding to a given PHINFO structure is found below the element table. The form that this information takes varies with the type of phrase. The additional information for NP structures 18 and 19 is given below.

NP PHRASE 18	NP PHRASE 19
HEAD= LAND	HEAD= ASSET
NUMBER=PL	NUMBER=PL
CONHS= 21,24	CONHS= 24

Noun phrases, verb phrases, prepositional phrases, relative clauses, and dependent clauses are assigned heads. The heads of prepositional phrases, dependent clauses, and relative clauses are outside of those phrases proper. The head of a prepositional phrase is the head of the object of its preposition, the head of a dependent clause is the head of the calling verb phrase, and the head of a relative clause is the head of the phrase which it modifies. In noun phrases and verb phrases, the HEAD slot contains the number of the word in the phrase which functions as the head. The headword of phrase 18 has the base lexical form land; the headword of phrase 19 has the base lexical form asset.

The NUMBER slot contains different types of information for different types of phrases. In noun phrases and prepositional phrases, the slot contains information as to the plurality of the phrase; in a verb phrase, it contains the person of the subject of the verb; and in a clause, it holds information on whether the clause is declarative or interrogative, dependent or relative. In phrase 18, NUMBER is PL, indicating the plurality of the entire coordinate construction; in phrase 19, NUMBER is also PL, indicating the plurality of that phrase.

In PHINFO structures for the conjuncts of a coordinate NP, the CONHS slot will contain the word number of the head of the phrase which the structure represents, followed by the numbers corresponding to the heads of all the subsequent conjuncts in the construction. Noun phrase 18, an initial conjunct, has two entries in the CONHS slot, 21 and 24, corresponding to the heads of the coordinate construction. Noun phrase 19, the rightmost conjunct,

has only one entry in the CONHS slot, 24, which is its own head.

There is one other piece of information which may appear in the PHINFO structure and which is present only in structures for verb phrases and interrogative clauses, the VERBFEATURES. These features carry information on tense, aspect, mood, do-support, and negation, as well as on whether the head of the phrase is a simple or a phrasal verb. Any syntactically permissible combination of the following features may be activated in a PHINFO structure: finite, passive, participle, past, progressive, subjunctive, imperative, future, perfect, negative, modal, DO, and adjunct.

5.2.4 Why a phrase-structure information representation?

A few words are in order about the intention behind the PHINFO structure representation of a sentence. It can be seen from structures 18 and 19 that the conjuncts of a coordinate construction are chained one onto the next; they do not appear as sister elements in one coordinate construction phrase information structure. Similarly, a prepositional phrase does not appear in an NP structure in which it is the sister of an NP, but rather in an NP structure in which it is the sister of the constituents of an NP (see, for example, phrase structure 12 in Appendix B). The PHINFO structure representation is not to be taken as a linguistic tree diagram in disguise; it is a diagramatic representation of the information extracted by the parser from the input string in the order in which it is extracted and in the units in which it is extracted. A linguistic tree structure could be constructed directly from it, but the result would be a highly, probably overly, structured tree, with many levels of embedding.

The networks could be redesigned so that the PHINFO structures could be translated directly into a flatter, more standard phrase structure tree diagram. If an NP arc originating at the first state of the NP network and terminating at a state at which a PP arc originated were added to the NP network, a prepositional phrase postmodifying a noun phrase could also be parsed as its sister rather than its daughter. Similarly, If an NP loop arc (an arc which originates and terminates at the same state) and a conjunction loop arc were added to the arcs leaving the NP network at state 1, the conjuncts of a coordinate NP could be parsed as sisters of one another rather than as daughters. The placement of an NP arc on the first state of the NP network, however, invites left recursion. Since a top-down parser such as the ATN does not handle left recursion well, this situation is to be avoided if possible.

The creation of a phrase structure tree is not the goal of the parser. The goal is to analyze the sentence, breaking it up into units of information which can be used by ENGSPAN to produce a Spanish translation. The phrase information structure is only an intermediary product of ENGSPAN and should be considered as such.

5.2.5 Output for parses of varying degrees of success

When a string has been completely parsed ("OK" appears in the message

column of the side-by-side display), all phrase information structures are closed off (i.e. the COMPLETE flag is on); the length of VERTPAT matches that of the input string; and all the codes in the side-by-side display contain an underbar. When a string is only partially parsed ("PP"), or when the parse is blocked ("NO"), not all the phrase information structures are closed off; VERTPAT is only as long as the longest path that was carved through the network system. The codes in the side-by-side display which correspond to those words to the right of the word that the parser was analyzing when it decided to give up are not separated by underbars.

5.3 The Parse

Once again, the sentence being analyzed is:

(1) It is clear that chronic malnutrition and hunger are linked to more general problems of poverty, inequitable distribution of land and other assets, underdevelopment, and adverse agroclimatic conditions.

The first task on the parser's agenda when it analyzes a sentence is to make a pass through the input string, mark all the hyphenated compounds and phrases between parentheses and double dashes, and find the last occurrence of a non-correlative coordinate conjunction. It then creates a PHINFO structure for a sentence, makes the first word of the input string the current word, and sets the pointer which will guide it through the network system to the address of the first arc from the first state of the sentence network.

Whenever the network pointer is set to a new state of a network, the parser determines the "compatibility" of the current word with the arcs that leave the current state. This determination is made in different ways for the different types of arcs. A word is compatible with a category arc if the label of the arc matches one of the possible parts of speech of the word; it is compatible with a seek arc of a given type if it can begin a phrase of that type; and it is always compatible with jump arcs and send arcs. If the current word has been determined to be compatible with an arc, and if there is either no condition associated with that arc or the condition associated with that arc is satisfied, the arc can be traversed.

When the parse of a sentence begins, the network pointer is set to the address of the first state of the sentence network. Two arcs originate from this state: a conjunction category arc and a clause seek arc. These arcs leave the state in a specified order, as do the arcs from every state of the network. It is in this order that they are checked for compatibility with the current word, and it is in this order that the conditions associated with them are tested.

When the decision is made that the current word meets the requirements of an arc, the parser places the address of that arc in an array and tests the next arc. Once all the arcs from the state have been tested, the array of addresses is placed on the pushdown stack and the network pointer is set to

the address which appears on the top of the stack.

At the current state, the conjunction category arc is ordered and tested first. It fails the compatiblity test. The next arc, the CL arc, is tested. It passes the compatibility test. Since there is no condition associated with the arc, all the requirements for traversing it are met. The address of the arc is put on the stack and the network pointer is subsequently set to that address.

The matching of a seek arc triggers a chain of events. A PHINFO structure is created for the type of phrase corresponding to the type of arc matched. This is the second structure created for the sentence (the first having been an ST structure). It is a CL structure for the first clause in the sentence. After creating the structure, the parser assigns to its CALLER slot the number of the phrase structure which was current at the time of the recursive call to the clause network. It then sets a pointer variable within the structure to the address of the seek arc that was matched. When the parser completes the analysis of the phrase it is now starting, it will consult this variable to know to which phrase and to which arc on which network it should return.

Phrase 2 is now the current phrase. The current word is still the first word of the sentence, it. The parser resets its network pointer to the address of the first state of the network which corresponds in type to the phrase being sought, namely a clause. Once again the parser goes through the steps of checking the arcs from the current state and putting on the stack the addresses of those arcs that can be taken by the current word. There are four arcs leaving the current state: an NP seek arc, a VI seek arc, an AP seek arc, and a conjunction category arc. The VI arc leads to the parse of an imperative clause; the AP arc, to the parse of an adverbial phrase; and the conjunction arc, to the parse of a correlative conjunction as the first element of a clause. The NP arc is the only arc with which the current word is compatible.

Once again a seek arc is matched. Another PHINFO structure is created, this one for an NP. The CALLER slot is set to 2, and the network pointer is set to the first state of the NP network. This state can be exited by a conjunction arc, a determiner arc, a pronoun arc, a gerund arc, and a send arc. The current word is compatible with the pronoun arc and the send arc. The condition on the send arc, namely that there be an NP on the hold list, is not satisfied. The condition on the pronoun arc, namely that the case of the current word not be objective if the NP is being parsed as a subject, is satisfied.

For the first time, the current word has matched a category arc. The part of speech corresponding to the category label, PRON (R), is entered both in VERTPAT and in the element table of the current PHINFO structure, and the number corresponding to the position of the word in the input string, namely 1, is entered in the location slot for that element of the phrase. The parser checks to see if there is an action associated with the arc just taken. There

is. It assigns the current word the role of head (H) and sets NUMBER to singular (SG).

After the action associated with an arc is performed, the network pointer is set to the address of the next state. Associated with each arc is a variable which indicates at which state in the network that arc terminates. The number assigned to this variable will indicate the number of the next state in the network to which the parser should proceed. The pronoun arc terminates at state 5 in the noun phrase network. Hence the current state is now state 5.

The word it is finally consumed, and the current word is set to the next word in the input string, is. At the current state, seek arcs are available to begin the parse of phrases which can postmodify an NP. The current word cannot begin any such phrase, so the parser jumps to state 6. At state 6, comma, conjunction, and send arcs are available. The current word is not compatible with either of the category arcs. It is compatible, however, with the send arc, as are all words. The condition associated with this arc requires that the noun phrase have a head and that it be syntactically well formed. This condition is satisfied and the send arc is taken.

This is the first time that a send arc is being traversed. The parser has several things to do. First it turns on the COMPLETE flag in the PHINFO structure of the phrase it has just completed. Next, it looks to see which seek arc it was that was matched in order to produce this phrase. It sets the network pointer to the address of this arc and continues the parse from the NP seek arc on the CL network.

The parse of the first element in the clause, an NP, has just been completed. NP is entered in the element table. The location of the NP is set to 3, the number of the PHINFO structure which corresponds to the phrase. The parser's final task with respect to this NP, now that it has parsed it and put it in its calling phrase, is to perform the action that is associated with the seek arc which triggered the analysis of the phrase. It assigns the role of subject (S) to the phrase.

The parser has now finished with the subject. The current state is set to the second state of the CL network (the state at which the NP arc terminates). There are two paths that the parser can follow with the current word, a verb phrase path and an adverbial phrase path. The current word, is, cannot begin an adverbial phrase, so there actually is only one way to go. The VP seek arc is traversed.

The current PHINFO structure is now phrase 4, the current word is is, and the current state is the first state of the verb phrase network. At this state there are arcs for a correlative conjunction, adverb, modifier, auxiliary, and main verb. The current word is compatible with the latter two and satisfies the condition on subject-verb agreement that is associated with both of them.

This is the first point at which there has actually been more than one path available to the parser. The addresses of both the auxiliary and the main verb arcs are placed on the stack, and the parser sets its network pointer to the top address, that of the auxiliary arc. The action associated with this arc causes the FINITE bit to be turned on in PHINFO. The current word is now set to the next word in the sentence, clear.

Having traversed the auxiliary arc, the parser arrives at state 2 of the verb phrase network. There it finds another auxiliary arc and a main verb arc. Although the current word can function as a verb, the parser can not take the main verb arc because the condition requiring the correct sequence of tenses is not satisfied. At this point there is no path that can be taken in a forward direction through the network. The parser must backtrack chronologically. Chronological backtracking involves the resetting of the current word to the last word for which there was a choice of paths to be taken, as well as the resetting of the network pointer to that last choice, i.e. the address on the top of the stack. Here the current word is reset to is, and the network pointer is reset to the address of the main verb arc from state 1 of the verb phrase network. All the information that was entered into any output structure about the word is is erased.

The verb arc takes the parser to the third state of the verb phrase network. The action associated with the verb arc assigns is as the headword of the current phrase. Is is now entered in the element table as a verb rather than as an auxiliary.

Once again, the current word is clear. From state 3 there is a path to follow. The arcs from this state are for the coda of the verb. There is an NP arc for verbs that are transitive, bitransitive, and/or take a complement; a DC arc for verbs that take a clause as object; an SN arc for verbs that take for-to and -ing complementation, an adjunct arc for phrasal verbs, and a category arc for the word not. There are also send, conjunction, and jump arcs. The send and jump arcs are for intransitive verbs. The jump arc takes the parser to the state at which it can parse an optional adverbial phrase; the send arc allows the parser to leave the verb phrase. The only arc with which the current word is compatible is the NP arc. Associated with this arc is a condition that tests whether the headword of the current phrase is coded to take a complement or an object. Since is is coded to take a complement, the condition is satisfied. The arc is traversed and the parser moves to the NP network.

With clear, the parser jumps through the network until it gets to state 3, where it can take an adjective arc. On the way to that state it stops to look carefully at one arc, the V2 arc from state 1. This arc would cause the parser to advance to the second state of the VP network. The current word can function as a verb in the right environment, so the condition associated with the arc is tested. It is not satisfied, however, because it requires that the current word be a present participial verb form.

There are two adjective arcs from state 3. The first is for an adjec-

tive which functions as a premodifier of the head of an NP, and the second is for an adjective which itself functions as the head of an NP. Since the current phrase is being parsed as a complement of the verb is, the current word could function as either. Both arcs must be considered. The addresses of both are placed on the stack. The premodifying arc is tried first because it is ordered first. This arc takes the parser back to state 3 with the next word, that, as the current word. The jump arc is taken to state 4.

At state 4 there are several noun arcs and a jump arc. There is no category arc for any of the possible parts of speech of that, i.e. no pronoun arc, no subordinate conjunction arc, and no determiner arc. The jump arc is not available to the current word because the condition is that the current word be either a comma or a conjunction. The only option for the parser is to backtrack chronologically and traverse the head adjective arc with clear as the current word.

Clear is made the head of the noun phrase; the current word is set again to that, and the current state becomes state 10. A PP seek arc and a jump arc originate at this state. The jump arc is taken to state 6, where a send, a conjunction, and a comma arc are available. The send arc is taken after the parser makes sure that the noun phrase is syntactically well formed.

The NP just parsed is entered into the element table of the VP PHINFO structure with the role of compl ment (C). The network pointer is set to state 4 of the VP network. Arcs are available for the parsing of an adjunct, a second NP, a dependent clause, and a reduced complement clause (for-to complementation). That is compatible with both the NP and the DC arcs. The condition on the NP arc, namely that the head of the verb phrase be bitransitive, is not satisfied; the arc can not be taken. The condition associated with the DC arc, i.e. that either the main verb or the adjectival complement of the verb be coded for a clause, is satisfied, and the arc is traversed.

The DC network is entered at its first state. The subordinate conjunction arc is taken to state 2, where the CL arc is matched. The current word is chronic. The NP arc from state 1 of the clause network is traversed, and on that network the premodifying adjective arc at state 3 is matched. The head adjective arc is not considered to be a candidate arc, as it was when clear was being parsed; the environment is not right for an adjectival NP.

The parser traverses the premodifying adjective arc with chronic as the current word. Now malnutrition becomes the current word, and the parser jumps to state 4. There are four noun arcs originating at this state: two premodifying noun arcs, a head noun arc, and a possessive noun arc. Only the head noun arc and the second premodifying arc can be taken. There are two premodifying noun arcs because there are two strategies for deciding whether a noun should be taken as a premodifier or a head. The first strategy, to try the noun as a premodifier before trying it as a head, is operative when the noun does not precede a conjunction; the second strategy is operative otherwise. The current word takes the head arc and leaves the address of the second premodifying noun arc on the stack. Should it turn out that the current word is

actually a premodifying noun, the premodifying arc can be picked up when the parser backtracks.

Malnutrition is entered in the PHINFO structure. The action associated with the current arc marks this noun as the head noun. It also sets NUMBER to singular (SG) and determines the modification relationships within the phrase. Chronic modifies malnutrition, so the word number corresponding to the latter is entered in the MODIFIES slot of the former.

The current word is set to and; the current state, to state 5. The jump arc with the condition that the current word be a conjunction is taken. This arc terminates at state 6, where a send arc, a comma arc, and a conjunction arc originate. The condition on the send arc is satisfied, as is the condition on the conjunction arc. Both addresses are put on the stack. The conjunction arc is ordered first and is therefore taken first. At state 7 there is one arc, an NP arc. The current word, hunger, is compatible with this arc, which is therefore taken.

With hunger, the parser takes several jumps through the NP network until it matches the head noun arc at state 4. With are, the new current word, the parser looks at the arcs from state 5 to see if there are any NP coda arcs with which it is compatible. There are none. This state has an RC arc for a relative clause, a DC arc for a clause in apposition (e.g. the fact that he left...), an SN arc for for-to complementation (e.g. the desire to be free...), and a PP arc for prepositional phrase modification. It also has a jump arc, which is the only arc that can be matched.

The current state is now state 6. The send arc from this state is traversed and the phrase is closed off. The NP just parsed is now entered in the element table of the PHINFO structure of the calling NP, its sister conjunct. The action associated with the conjunct NP arc from state 7 is performed. This action calls for checking the configuration of the coordinate construction to make sure that it is well formed. If it is not well formed, the parser will backtrack to a specific point in the parse (explicit backtracking) and try again. The construction is determined to be well formed, and the parser moves on to state 8, where it takes a send arc out of the NP network.

The NP just parsed, a coordinate NP, is added to the element table of the PHINFO structure for the clause (which is part of the dependent clause) and is marked as the subject. The string parsed up to this point is: "It is clear that chronic malnutrition and hunger." With the current word, are, the parser has the same options available to it at this point in the network system as it did when the second word of the input string, is, was the current word. The VP arc is taken, and the addresses of both the auxiliary and the main verb arc from state 1 of the verb phrase network are placed on the stack. The auxiliary arc is traversed first. At state 2, with the current word linked, the condition associated with the main verb arc is not satisfied. Linked is coded as both a verb and a resultative adjective. The parser employs the strategy of trying to analyze a word which might be either a passive verb or a resultative adjective as an adjective, as long as it is not

followed immediately by a prepositional phrase which could mark the agent of the verb.

Are is reparsed as the head of the verb phrase, and linked is parsed on the NP network just as clear was. At state 10 of the NP network the PP arc is taken, since linked is coded in the dictionary as commonly preceding the preposition to, which is the next word.

On the PP network, to causes the parser to take the preposition arc from state 1, and more causes it to select the NP arc from state 2. Once again the parser is on the noun phrase network. With the current word as more, both the modifier and the numerative arcs are matched. The modifier arc is traversed, and the address of the numerative arc is saved on the stack. General is parsed as an adjective at state 3. General is also a noun, but since the parser has no noun arc that it can take from the current state, the noun possibility is not entertained. Problems, the next word, causes the parser to take a jump arc to state 4, where it traverses the head noun arc.

The NP more general problems has been parsed with more as a modifier of the adjective general rather than as a numerative modifying the head problems. The phrase is truly ambiguous. The parser only gives one reading of a truly ambiguous phrase, i.e. the reading that it arrives at first by taking the arcs in the order in which they leave the states. If the phrase were not truly ambiguous and the adjective was either not comparable or it formed the comparative only by affixation, backtracking would have taken place and more would have been parsed as a numerative.

We return to the parser at state 5, the NP coda state. With the current word, of, the parser takes the PP arc and saves the address of the jump arc on the top of the stack. On the PP network the parser takes the preposition arc from state 1. With poverty then as the current word, it takes the NP arc from state 2. The path that the parser paves through the network with poverty is identical to that which it paved with malnutrition as the current word. After poverty has been parsed as the head noun, a comma becomes the current word. The parser takes the jump arc from state 5 and puts the address of the RC arc on the stack. Two RC arcs leave the current state. The first is taken if the current word is a relative pronoun; the second one is taken in all other situations (for unmarked relative clauses and appositives).

At state 6 the parser traverses the comma arc and puts the address of the send arc on the stack. The NP network is re-entered with inequitable as the current word when the parser takes the NP conjunct arc from state 7. When analyzing the string inequitable distribution of land and, the parser goes through almost the same steps that it went through to parse the string general problems of poverty,. The only difference in the paths for these strings is that the former includes a conjunction arc, and the latter a comma arc. This is the only difference, in spite of the fact that the former string contains an ambiguity that the latter does not. The word land comes into the parser as both a noun and a verb. Because it follows the preposition of, however, and it is not a participial form, there is actually no part-of-speech ambiguity.

With the remaining strings in the sentence, namely other assets, underdevelopment, and adverse agroclimatic conditions, the parser paves paths in the noun phrase network similar to the those for the other noun phrases in the sentence. The verb-noun ambiguity of the lexical item conditions gives it no problem because the word is first seen at state 4 of the NP network and from that state there is no verb arc that it could take with conditions.

The parser is faced with one major problem when it finishes analyzing adverse agroclimatic conditions. As it closes off one NP conjunct and adds it to the element table of the previous conjunct, it finds that it has parsed land and other assets, underdevelopment, and adverse agroclimatic conditions as one coordinate NP. After the parser enters the NP inequitable distribution + PP in the PHINFO structure of the NP containing the elements poverty and a comma, it performs the action associated with the conjunct NP arc. This action checks the well-formedness of the configuration of the coordinate construction. The coordinate construction is made up of two complete NP conjuncts that are separated only by a comma. This is not a well-formed construction. The parser must backtrack to the last point in the parse at which it attached a conjunct, i.e. the last point at which it had the choice of taking a send arc to complete a phrase or taking a comma or conjunction arc to parse a right conjunct.

Starting from the top of the stack, the parser and discards all the addresses it finds until it gets to the address of an arc that was matched for a comma or a conjunction. The first such address is that of the send arc from state 6 of the NP network. This arc was matched for the last comma in the sentence.

The parser gets ready to return to that point in the parse at which it had just finished parsing underdevelopment as the head of an NP. The first thing it does is save all the phrase information structures it had completed after the point in the parse to which it is now returning. The structure for the NP adverse agroclimatic conditions is put on the well-formed phrase list, as are copies of every complete phrase which contains as a constituent, immediate or otherwise--either the word that the parser is now trying to reparse, or any word following it. This means that a copy of each of the PHINFO structures created in the course of the parse of the string inequitable distribution of land and other assets, underdevelopment, and adverse agroclimatic conditions is placed on the list. A phrase on the well-formed phrase list is inactive, but it can be reactivated if and when the parser returns to the word which begins that phrase and it is again seeking a phrase of that type.

The parser encounters problems again when it tries to close off the NP underdevelopment and enter it into the NP containing the elements other, assets, and comma. Just as before, the action on the conjunct NP arc rejects the coordinate construction in which the last conjunct is separated from the conjunct to its left by only a comma. Once again, the parser goes through the stack to find the address of the last arc which was matched for a comma or conjunction. It finds the address of the send arc from the NP network, matched for the comma parsed after the word assets. The phrase other assets

is closed off and entered in the PHINFO structure for the object of the preposition of. Now the parser has the correct conjoined structure as the object of the preposition. It puts this NP in the PHINFO structure for the PP and it finds itself on the prepositional phrase network with the current word, which is a comma (the second to last in the sentence). The comma arc from state 3 is traversed. The analysis is blocked. With the current word, once again underdevelopment, the parser cannot take the PP conjunct arc, the only arc originating at state 4. Once again, it has to backtrack--this time chronologically.

When the parse is blocked because there is no path to take through the network, backtracking is chronological. When the parse is blocked by an action, which tests the configuration of a phrase and finds it unsatisfactory, the parser can backtrack explicitly, i.e. to a specified point in the parse history. This is because the tests made in the action can diagnose why there was a misparse and at what point it occurred. Since the present scenario is that there is nowhere to go, the parser must go back to the last address that it placed on the stack.

The problem facing the parser is an attachment problem. Rather than trying to close off the prepositional phrase once its object had been parsed, the parser tried to parse a conjunct for that phrase. So now it goes back and closes off the prepositional phrase of land and other assets and enters it in the PHINFO structure of the NP already containing the elements inequitable and distribution. The parser is faced with the decision of whether to close off a phrase or try to parse a conjunct for it. It follows its strategy (determined by the ordering of comma, conjunction, and send arcs) of trying to parse a conjunct first. Still with a comma as the current word, it takes the comma arc to state 9 and from there it takes the jump arc to state 7. Now it is looking for an NP that begins with the current word, underdevelopment. There are two phrases that it can take off the well-formed phrase list, those which correspond to the conjuncts of the phrase underdevelopment, and adverse agroclimatic conditions. These phrases are reactivated, and everything falls into place. Each conjunct of the object of the preposition of (following problems) is closed off and entered in the calling conjunct; the initial conjunct is closed off and placed in the calling prepositional phrase; the prepositional phrase is closed off and placed in the calling noun phrase; that noun phrase is closed off and placed in the calling prepositional phrase; and so on up the line until the sentence is closed off with all the elements of the input string consumed.

The parser is now finished with the string. It tells the message board to give the string an OK and it sends VERTPAT and the phrase information structures back to ENGSPAN, which works on the transfer of the English structure to Spanish structure and the synthesis of the Spanish output.

5.4 Other Facilities of the Parser

Not all the facilities of the parser were employed in the analysis of

the foregoing sentence. The demonstration sentence contained no hyphenated compounds, no parenthetical material, no words that were not in the dictionary, and no relative clauses. In the following paragraphs the manner in which the parser handles these phenomena will be discussed.

5.4.1 Hyphenated compounds

When the parser receives a sentence to analyze, it checks for the presence of hyphenated compounds. When it finds one, it marks each element in the compound. A special seek arc, the HY arc, accomodates the compound. This arc is only compatible with a word that is marked as the first word of a hyphenated compound. Words marked as participating in a hyphenated compound are only compatible with HY arcs and arcs on the HY network. HY arcs are found in the noun phrase network for the parsing of descriptors and heads. Hyphenated verbal compounds must be entered in the dictionary.

5.4.2 Parenthetical material

Parenthetical material is defined as material within parentheses or between double dashes. In the intial pass that the parser makes through the input string, the left parentheses and double dashes are marked as the beginning of parenthetical material. A variable is set in the copies of the dictionary records of these markers to record the length of the parenthetical phrase. When the parser reaches the beginning of the parenthetical phrase or clause, it enters the main part of speech of each of its constituent elements in VERTPAT. The analysis of the phrase is left for the safety-net routines.

5.4.3 Words not in the dictionary

A word which is not in the dictionary comes out of the dictionary-lookup and lemmatization routines with at least one part of speech. Initially, it is assigned the parts of speech verb, noun, and adjective. Information from potential prefixes and/or suffixes can cause additional parts of speech to be assigned. For example, if the not-found word ends in ly, it is also assigned the part of speech adverb. Information about affixes is used in the same way for words that are in the dictionary as for words that are not. The only difference in the way in which the parser treats found and not-found words is that it automatically considers the latter to satisfy any semantic test.

5.4.4 Relative clauses

For the parsing of relative clauses, and Wh-questions too, the parser makes use of a hold list which keeps up to three extraposed phrases. These phrases can be picked up by special send arcs from the first states of the NP and PP networks. To demonstrate how relative clauses are parsed, a quick trip will be taken through the parse of a modified version of the demonstration sentence, given below. The output for this sentence is found in Appendix C.

(1') It is clear that chronic malnutrition and hunger are linked to more general problems that have already been discussed.

We have already followed the parse through the word problems. Whereas in the demonstration sentence this noun was followed by a prepositional phrase, parsed by taking the PP arc from state 5 of the NP network, in the modified sentence this noun is followed by a relative clause, parsed by taking the RC arc from that state. We will start the parse of the sentence above at state 1 of the RC network.

The current word is that. The parser takes the subordinate conjunction category arc and arrives at state 2. The action associated with this arc is performed. The number of the calling phrase (the NP more general problems) is placed on the hold list. The current word is set to have and the CL arc from state 2 is traversed.

In the clause network, the NP arc from state 1 is matched. It is matched not for the current word but for the phrase on the hold list. A phrase on the hold list is compatible with jump arcs and send arcs, as well as with seek arcs that are of the same type as the phrase, provided that certain constraints are not violated. When a seek arc is matched for a phrase on the hold list, as happens here, a hold flag is turned on.

When the parser enters the NP network, it creates, as usual, a phrase information structure. Because the hold flag is on, however, it changes the type of the phrase to NH (noun phrase on hold) and copies the contents of the phrase information structure for the phrase on the hold list into the new phrase information structure. After doing this, it takes the phrase off of the hold list and puts it in a holding variable so that it will be available for future use, in the event that the relative clause about to be parsed turns out to be part of a coordinate construction.

The send arc from state 1 of the NP network tests for the hold flag; if it is on, the arc can be taken. This is the case; therefore the arc is taken. The action associated with the arc turns off the flag, and the parse of the clause continues. The string have already been discussed is parsed as a verb phrase, and the clause is completed. The action associated with the CL arc from state 2 of the RC network is designed to make sure that the phrase placed on the hold list at state 1 has been removed.

The send arc from state 3 of the RC network is taken; the phrase is closed off and entered in the calling phrase. For all other phrases, the parser follows suit. When the ST phrase is closed off, the parse is successfully completed.

5.4.5 Overall coverage

Interim progress reports have followed the development of the parser and the expansion of its syntactic capabilities. During the final reporting period, 1 February 1985 - 31 October 1985, the following constructions were added to the parser's repertoire: gerunds; interrupting, parenthetical, and appositive phrases and clauses; non-restrictive relative clauses; reduced relative clauses beginning with adjectives; relative clauses not marked by

relative pronouns and beginning with pronouns or with human or collective nouns; and dependent clauses with subjunctive verbs. The side-by-side display of sentences demonstrating these and other capabilities appears as Figure 4. In order to demonstrate the overall syntactic capability of the parser, a side-by-side display (with unedited translation) for 20 pages of running text selected from the ENGSPAN production flow is given in Appendix D. Eighty-six percent of the input strings were parsed. Of these, the average length of a successfully parsed string was 14 words and the longest string parsed was 42 words.

The simulator presented the following tasks. (1)	OK OK	t_ n_ v_ t_ a_vnp n_v: (k_) v_ t_ n_ p_ n_v n_; (
controlling a car in wind gusts; (2) following a lead car	OK	k_) v_nap t_ a_ n_ p_ a_n n_v; (k_) v_ t_ n_ p_c n_
at constant speeds; (3) passing a car between obstacles in	OK	p_ t_ a_ n_.
the adjacent lane.		
The passing task tested the driver's ability to complete	OK	t_ v_ n_v v_a t_ n_ n_ p_ v_a t_ a_ n_v p_ t_ a_ n_ p_ k_
a successful pass of a lead car for three different passing		a_ v_ n_.
distances.		
The second decisionmaking task, the emergency stop, was	OK	t_ k_vnd n_ n_v , t_ n_ n_v , v_x g_ t_ a_ n_v.
not an optional task.		
The degree of behavioral impairment was found to be	OK	t_ n_ p_ a_ n_ x_v v_ p_ v_x d_ a_ p_ t_ n_ n_v n_v.
roughly proportional to the treatment dose level.		
Laboratory studies have shown marijuana to impair	OK	n_ n_v x_v v_ n_ p_ v_ a_ c_ a_ n_v a_ p_ n_va.
perceptual and perceptualmotor functions important to		
driving.		
While none of the car control measures showed any	OK	s_ r_k p_ t_ n_ n_ n_ v_ t_r n_ p_mo n_ , t_ n_ p_ t_ a_
decrement under marijuana, the response to the peripheral		n_v n_ n_v v_ t_ a_v n_v , s_rt v_x d_ a_v p_ t_ n_v n_v.
signal detection task showed an increasing delay, which was		
linearly related to the drug dose.		
It will answer many questions, and hopefully, inspire	OK	t_ x_n v_n k_ n_v , c_ d_ , v_ a_ r_.
additional ones.		
The task of hiring and assigning staff to perform the	OK	t_ n_v p_ v_n c_ v_ n_v p_ v_ t_ n_v v_x k_nr s_rt x_ x_v
work is one which must be completed prior to training.		v_a p_ n_v.
The first step in organizing special training is to	OK	t_ k_d n_v p_ v_ a_ n_v v_x p_ v_ t_ n_ c_ n_ a_ p_ t_ a_
describe the skills or knowledge necessary for a specific		n_v.
position.		
Deciding on the necessary equipment is a surprisingly	OK	v_ t_ a_ n_ v_x t_ d_ a_n n_.
complex problem.		
Continuing education might take place in the form of	OK	a_ n_ x_ v_ p_ t_ n_v p_ v_ a_ n_ c_ n_ p_ n_.
introducing new concepts or methods of application.		
They should be made to feel that what they do is	OK	r_ x_ x_v v_a p_ v_ s_rt r_tc r_ v_x v_x a_ c_ v_x t_ n_v
important and has an effect on immunization activities and		p_ n_ n_ c_ t_ n_ r_ v_. t_r n_ , s_ x_ v_x a_ n_v p_ v_
the community they serve. Each participant, who should	OK	n_ p_ k_nr p_ t_r n_ p_ n_ c_ n_v , x_n x_v v_a p_ v_ t_
have actual experience in using microcomputers in one of		n_ c_ t_ n_ v_ t_rc n_v.
these areas of education or training, will be asked to		
prepare a demonstration or a paper describing that		
experience.		

Figure 4. Typical constructions handled by the parser.

We request that each potential participant send biographical information and a 100-word abstract of the demonstration or paper, so that we can select those who will make the greatest contribution to a useful exchange of information at the symposium.	OK	r_ v_n s_rl t_r a_n n_ v_ a_ n_ c_ t_ k_ 76 n_v n_va p_ t_ n_ c_ n_ , s_ r_ x_n v_a r_ s_ x_n v_ t_ a_ n_ p_ t_ a_ n_v p_ n_ p_ t_ n_.
The Consultative Group on International Agricultural Research (CGIAR), established in 1971, is an informal association of governments, international and regional organizations, and private foundations dedicated to supporting a system of agricultural research centers and programs around the world.	OK	t_ a_ n_v p_ a_ a_ n_ (n_) , v_a p_ n_ , v_x t_ a_ n_ p_ n_ , a_ c_ a_ n_ , c_ a_ n_ v_a p_ v_a t_ n_ p_ a_ n_ n_ c_ n_v p_m t_ n_a.
It is clear that chronic malnutrition and hunger are linked to more general problems of poverty, inequitable distribution of land and other assets, underdevelopment, and adverse agroclimatic conditions.	OK	r_ v_x a_v s_rl a_ n_ c_ n_ v_x a_v p_ m_k a_n n_ p_ n_ , a_ n_ p_ n_v c_ t_a n_ , n_ , c_ a_ a_ n_v.
Such disparities are a measure both of the developing countries' predicament and of the gains that might result from research.	OK	t_ n_ v_x t_ n_v c_rl p_ t_ a_ n_ n_ c_ p_ t_ n_v s_rl x_ v_n p_ n_v.
To be useful, therefore, CIAT's research must be equally diverse, versatile, and responsive to both social and biological realities.	OK	p_ v_x a_ , d_ , n_ n_v x_ v_x d_ a_ , a_ , c_ a_ p_ c_rl a_ c_ a_n n_.
The search for inherent resistance to pests and diseases has now been extended from the existing genotypes in the germplasm collection to new, advanced lines bred at CIAT.	OK	t_ n_v p_ a_ n_ p_ n_ c_ n_ x_v d_ x_v v_a p_ t_ a_ n_ p_ t_ n_ p_ a_ , a_v n_v v_ p_ n_.
International networks of cooperating programs evaluate CIAT's improved materials for effective resistance and yield in a variety of actual farming environments, leading to the selection of superior lines for release by many countries in the region.	OK	a_ n_ p_ a_ n_ v_ n_ a_v n_ p_ a_ n_ c_ n_v p_ t_ n_ p_ a_ a_vn n_ , v_a p_ t_ n_ p_ a_n n_v p_ n_v p_ k_ n_ p_ t_ n_.
Such yield increases will require genetically-improved varieties with resistance to pests and disease and improved production practices.	OK	t_ n_v n_v x_n v_ d_ 76 v_a n_ p_ n_ p_ n_ c_ n_ c_ a_v n_ n_v.

6. THE SAFETY NET

ENGSPAN is geared to free syntax and is designed to be challenged daily by random text. Even so, it cannot handle every input string that a writer in English is capable of producing. In its present state of development, the parser is able to complete the analysis of approximately 60 percent of the strings in a text which is submitted for translation without prior review by the project staff. After correction of spelling and punctuation errors and some basic dictionary work, the success rate increases to about 70 percent. If a text is subjected to detailed analysis and thorough dictionary updating, the percentage of successful parses can exceed 80 percent.

The ATN grammar is designed primarily for parsing complete sentences. It can also handle titles and other types of incomplete sentences such as those found in lists and in segments of the input bounded by colons and semicolons. In addition, a runtime parameter allows the parser to accept certain types of sentence fragments even if they are punctuated as complete sentences. Two other program parameters affect the parser's ability to complete the analysis. On the one hand, the size of the parser's work space is determined by a runtime parameter which specifies the maximum number of phrase information structures that can be created. On the other hand, the length of the input string is limited to 70 words. Longer strings are divided arbitrarily, which may result in two unanalyzable fragments.

There are many other possible reasons for an unsuccessful parse. The input may contain a misspelled function word (e.g. form instead of from) or punctuation errors, or it may not follow standard rules of English syntax. A key lexical item or idiomatic expression may be missing from the dictionary or may be coded incorrectly. An ill-advised SU may prevent a necessary word-by-word analysis. The sentence may contain a large number of homographs which cause the parser to try unsuccessful paths, or the sentence may be so long that the parser uses up its work space before completing the analysis. Complicated sequences of conjoined, embedded, or interruptive phrases can also cause problems. Finally, of course, the particular syntactic construction may not have been accounted for in ENGSPAN's grammar.

In order to ensure that ENGSPAN produces a Spanish translation for every sentence, a safety net strategy has been built into the program. The safety net procedures allow the transfer component to operate with or without a parse. Whenever possible, the safety net makes use of information returned by the parser. When a partial parse is produced, the active phrase information structures are used, provided that they have been assigned a headword. If the parser has tried all possible paths through network and rejected all of them, the longest path (which is returned in VERTPAT) is used.

Part-of-speech disambiguation for a word that does not belong to an active surface structure node is accomplished in one of two ways. If the word was parsed as part of the longest path, it is assigned the part of speech specified on the category arc originally traversed by the parser. Otherwise,

the part of speech is determined by a backup homograph resolution procedure. This procedure looks at a maximum of two tokens to the left and right, compares the syntactic environment with the coding of the word in question, and selects the most probable part of speech.

In addition to homograph resolution, the safety net includes procedures for phrase-level analysis and the identification of subject-verb-object relationships in simple clauses. Often these capabilities make it possible to produce an acceptable translation; sometimes the result is identical to a translation based on a complete parse. The safety net procedures will have problems, however, if the parser is pursuing an incorrect analysis when it exhausts its work space or if the unanalyzed portion of the sentence contains several contiguous homographs.

7. THE TRANSFER COMPONENT

7.1 Surface Structure Markup

When the parse has been successful, the next step is to inspect the phrase information structures produced by the parser and use the information that they contain to fill in a structural transfer table. This table provides a common source of data for the lexical and syntactic transfer procedures, no matter how the analysis was produced. The safety net procedures do not create phrase information structures, but they can contribute information to this table.

Structural markup also involves deriving additional information needed for translation. For example, a noun phrase that has been parsed as the subject of a verb in the passive voice must be marked as the underlying direct or indirect object of the verb. When a subject has been raised out of an embedded clause, the two roles of that noun phrase are indicated. For complements and reduced relative clauses, the modification relationships that were established by the parser are associated with specific lexical items. Infinitive phrases that function as verbal complements may be marked for restructuring as dependent clauses. Gerund phrases that function as the object of a preposition may be marked for synthesis as nominalizations.

7.2 Lexical Transfer

The selection of alternative target glosses based on sentence context is governed by the AUs and TUs (see Section 3.3). Both these types of units offer the possibility of changing the LEX that is used for retrieval. AUs are applied before the sentence is parsed and have been described in Section 3.3.3 above. TUs are applied just prior to target dictionary lookup. Most types of TUs depend on the information produced by the parser, but the safety net strategy allows some units to take effect even when the sentence cannot be parsed.

Lexical transfer itself is accomplished by retrieving from the Spanish target dictionary the record corresponding to the LEX of the English entry. This record may contain coding that causes the target lookup procedure to retrieve an alternative entry. Codes in the main target entry indicate the existence of additional records for microglossary entries, irregular plural nouns, irregular verb nominalizations, and adjectival translations for premodifying nouns and present participles.

The TU is a lexical transfer rule which is stored in the source dictionary. The rule contains one or more conditions to be tested and an action to be performed. If a source item has been coded as the trigger for a TU, the units pertaining to that item are retrieved and tested in sequential order. If the conditions specified in the unit are met, its actions are performed. If the conditions are not met, the next unit is retrieved and tested. This

process continues until a unit is found for which the conditions are met or until all units for the word have been checked.

TUs can be written for both single words and substitution units. A "NOT" operator can also be used in TUs. A unit that contains a "NOT" operator will be accepted only when the last condition tested by the unit is not met. The following sentences contain examples of TUs. The translations are shown in Figure 5.

7.2.1 Triggers for special translations of verbs

In each case below the affected verb is indicated with a solid underscroe and the trigger with a broken line.

The surface or underlying object is animate:

(2) The children were raised in Mexico.

The surface or underlying object is a specific word:

(3) They raised the price of beef.

(4) They raised the American flag.

(5) The participants raised many interesting questions.

(6) The funds raised by the volunteers were an important contribution.

The subject is human:

(7) The patient gave birth to twins.

The subject is not animate:

(8) The concept gave birth to a new development effort.

The subject is not human but is animate:

(9) The dog gave birth to six puppies.

The verb has a specific type of coda:

(10) He asked to continue working.

(11) The professor asked that the students return the books.

(12) The professor asked when the students would return with the books.

(2) The children were raised in Mexico.	OK TU	(2) Los niños se criaron en México.
(3) They raised the price of beef.	OK TU	(3) Aumentaron el precio de carne de res.
(4) They raised the American flag.	OK TU	(4) Izaron la bandera americana.
(5) The participants raised many interesting questions.	OK TU	(5) Los participantes plantearon muchas preguntas interesantes.
(6) The funds raised by the volunteers are an important contribution.	OK TU	(6) Los fondos recaudados por los voluntarios son una contribución importante.
(7) The patient gave birth to twins.	OK TU	(7) La paciente dió a luz gemelos.
(8) The concept gave birth to a new development effort.	OK TU	(8) El concepto produjo un esfuerzo nuevo de desarrollo.
(9) The dog gave birth to six puppies.	OK TU	(9) La perra parió seis cachorros.
(10) He asked to continue working.	OK TU	(10) Pidió continuar trabajando.
(11) The professor asked that the students return the books.	OK TU TU	(11) El profesor pidió que los estudiantes devolvieran los libros.
(12) The professor asked when the students would return with the books.	OK TU	(12) El profesor preguntó cuando los estudiantes regresarían con los libros.
(13) He asked for a raise.	OK	(13) Pidió un aumento.
(14) We have made up for the loss of time.	OK TU	(14) Hemos compensado la pérdida de tiempo.
(15) The child came down with chickenpox.	OK TU	(15) El niño contrajo varicela.
(16) The patient became hysterical.	OK TU	(16) El paciente se puso histérico.
(17) The woman became ill during the night.	OK TU	(17) La mujer se enfermó durante la noche.
(18) The worker held the employer responsible for the accident.	OK TU	(18) El obrero responsabilizó al empleador del accidente.
(19) In this respect, the government wished to reiterate its position.	OK TU	(19) En este sentido, el gobierno deseaba reiterar su posición.
(20) The plaintiff had no legal grounds for his actions.	OK TU	(20) El demandante no tenía fundamentos jurídicos para sus acciones.
(21) Our efforts paved the way for future research.	OK TU	(21) Nuestros esfuerzos prepararon las condiciones para investigaciones futuras.
(22) We need a rough estimate of the number of users.	OK TU	(22) Necesitamos una estimación aproximada del número de usuarios.
(23) The pesticide did not affect the female reproductive organs.	OK TU	(23) El plaguicida no afectó los órganos reproductivos femeninos.
(24) The female rats received a higher dose.	OK TU	(24) Las ratas hembras recibieron una dosis mayor.

Figure 5. Transfer units

(25) The methods of crop protection that have been used by small farms for centuries must be improved by the introduction of modern agricultural concepts.	OK TU TU	(25) Los métodos de protección de cultivos que han sido empleados por explotaciones agropecuarias pequeñas por siglos deben ser mejorados mediante la introducción de conceptos modernos agrícolas.
(26) The Organization published a study on the effects of the drugs.	OK	(26) La Organización publicó un estudio sobre los efectos de los medicamentos.
(27) The agency is responsible for the project.	OK	(27) La agencia es responsable del proyecto.
(28) There are two bauxite-producing companies in the country.	OK TU	(28) Hay dos empresas productoras de bauxita en el país.
(29) The results of the study showed a five-fold increase.	OK TU	(29) Los resultados del estudio mostraron un aumento de cinco veces.
(30) This house has burglar-proof windows.	OK TU	(30) Esta casa tiene ventanas a prueba de ladrones.
(31) Accident-prone employees increase the cost of insurance.	OK TU	(31) Empleados proclives a accidentes aumentan el costo de seguro.

Fig. 5, page 2

The verb is followed by bound preposition:

(13) He asked for a raise.

(14) We have made up for the loss of time.

The object of the bound preposition is a condition:

(15) The child came down with chickenpox.

The complement is an adjective:

(16) The patient became hysterical.

The complement is a specific word:

(17) The woman became ill during the night.

(18) The worker held the employer responsible for the accident.

7.2.2 Triggers for special translations of nouns

The noun is the object of a specific preposition:

(19) In this respect, the government wished to reiterate its position.

The noun is modified by a specific word:

(20) The plaintiff had no legal grounds for his actions.

The noun is the object of a specific verb:

(21) Our efforts paved the way for future research.

7.2.3 Triggers for special translations of adjectives

The head of the noun phrase is not concrete:

(22) We need a rough estimate of the number of users.

The head is not animate:

(23) The pesticide did not affect the female reproductive organs.

The head is not human:

(24) The female rats received a higher dose.

7.2.4 Triggers for special translations of prepositions

The object is a time word or the object is a deverbal noun:

(25) The methods of crop protection that have been used by small farms for centuries must be improved by the introduction of modern agricultural concepts.

The preposition is governed by a noun:

(26) The Organization published a study on the effects of the drugs.

The preposition is governed by an adjective:

(27) The agency is responsible for the project.

7.2.5 Special translations for parts of hyphenated compounds

It is possible to elicit a special translation for a word when it occurs as part of a hyphenated compound:

(28) There are two bauxite-producing companies in the country.

(29) The results of the study showed a five-fold increase.

(30) This house has burglar-proof windows.

(31) Accident-prone employees increase the cost of insurance.

7.3 Syntactic Transfer

While the TUs are the mechanism for selecting the most appropriate gloss from the target dictionary, the syntactic transfer rules determine the structure of the Spanish sentence. Using the slot-and-filler approach, the algorithm applies as many rules as necessary in order to generate a grammatically correct and stylistically acceptable sentence in the target language. Whenever possible, the linear order of the phrasal units is preserved in order to maintain the overall cohesion of the text. In many instances, correct Spanish syntax is achieved through the manipulation of function words. Topicalization is used to handle certain instances of preposed noun phrases. One situation that requires rearrangement of entire phrases is that of a clause in which the

finite verb is the final element.

The transfer component includes three different types of syntactic transfer rules. The first type is triggered by a lexical item in the English source text or the presence of a particular surface structure in English. The second type is a general syntactic transformation that converts a surface structure of English to a surface structure of Spanish; the rules are independent of the lexical items involved. These two types of syntactic transfer rules are applied prior to the target lookup. The third type of rule depends on the characteristics of the lexical item in the Spanish target text and is applied after target lookup.

Examples of rules triggered by specific lexical items in the source are the transformation of there + be into the Spanish 'haber' construction and introduction of a first person subject in the sentence Let us begin. General rules include the ordering of the elements of the verb string, the insertion and ordering of clitic pronouns, the suppression of subject pronouns, and the nominalization of gerundive phrases that function as the object of a preposition. Examples of rules that require access to information about the target lexical items are subject-verb agreement, the ordering of the elements of the noun phrase, the triggering of the subjunctive mood and imperfect tense, and the transformation that reverses the syntactic roles of the subject and object with verbs such as 'gustar' and 'faltar'.

Syntactic transfer works best when the input string has been successfully parsed. Some rules, such as those that deal with subject raising and verb preposing, can only be applied if the necessary information is available from the parse. For other rules, however, the transfer algorithm contains sufficient rules to make up for the lack of a complete parse. For example, verb strings with a limited amount of interruptive material can usually be restructured correctly. Most noun phrases, including many with multiple premodifiers and conjoined structures, can also be identified and rearranged into Spanish word order.

8. THE SYNTHESIS COMPONENT

When the transfer component has completed its work, each target lexical item is associated with a group of features which specify the inflectional endings to be synthesized. Additional codes indicate whether the item is to receive a derivational prefix or suffix and whether the item is to be preceded by function words such as the definite article, personal 'a', or the preposition 'de'.

Verb synthesis is performed first, since the resulting form may serve as the input to other synthesis procedures. The next task is the generation of the surface forms of the reflexive, direct object, and indirect object pronouns. Then all elements of the noun phrase are inflected for gender and number. This step is followed by prefixation and the insertion of function words. The final steps in synthesis are the introduction of phonologically determined variants and the adjustment of capitalization and diacritics.

8.1 Synthesis of Spanish Verbs

The algorithm for the synthesis of Spanish verb forms is based on principles of generative morphology and phonology. The program synthesizes regular verbs and most of the irregular verbs in all tenses and moods except the future subjunctive and in all persons except the second person plural. The verb is entered in the target dictionary in its stem form. Binary codes are used to specify the conjugation class and 11 exception features which govern the synthesis of irregular forms. Only one dictionary entry is needed for each verb. A small number of highly irregular stems and full forms (74 in all) are listed in a table. The majority of "stem-changing" verbs require no special synthesis coding. The procedure consists of a series of morphological spellout rules; raising, lowering, diphthongization, and deletion rules based on phonological processes; stress assignment rules; and orthographic rules to handle predictable spelling changes.

Spanish verbs are also coded to permit synthesis of the nominalization corresponding to 'the action or process of'. The verb synthesis procedure is invoked to produce nominalized forms which correspond to the first or third person present indicative or present subjunctive, or to the past participle or infinitival form of the verb. Other forms are produced by a combination of stem alterations and derivational affixes. Irregular forms require a separate target dictionary entry.

8.2 Synthesis of Noun Phrases

The synthesis of noun phrases involves the reordering of long premodifying strings, adjectival heads, conjoined or hyphenated heads, phrasal possessives, many types of hyphenated modifiers, expressions of quantification, and embedded parenthetical material. It also includes the handling of dates, proper names, and place names.

The synthesis procedure scans the sentence from right to left until it encounters the head of a noun phrase. The constituents are then examined from both the left and right boundaries of the phrase and are placed on one of five different stacks depending on their role in the phrase and on the context-free and context-sensitive information contained in each dictionary record.

The same algorithm is used for all sentences, regardless of the status of the parse. Built-in constraints limit the scope of the phrase when no boundary information has been supplied by the parser. In addition to reordering the constituents, the procedure provides for the concordance of gender and number, the insertion of 'de' and the definite article, and for the insertion of personal 'a'.

8.3 Examples of Spanish Synthesis

Figures 6 through 10 give unedited ENGSPAN translations of sentences that illustrate some of the capabilities of the rules for syntactic transfer and procedures for target synthesis. Figure 6 shows the translation of some of the sentences used in preceding portions of the report. In the remaining figures the examples are grouped by the type or scope of the syntactic structure involved, and the specific capability being shown is indicated below. None of the examples depends on the presence of a multiple-word entry in the source dictionary.

8.3.1 Noun phrases

Figure 7 shows how ENGSPAN has synthesized the following examples of noun phrases:

A long string of premodifiers:

(32) relevant geographic, climatic, demographic, political, economic and social characteristics

Other types of conjoined modifiers:

(33) safe, effective use of pesticides

(34) highly qualified and very dedicated staff

(35) 15 subregional and national production centers

(36) morbidity and mortality statistics

A modifier of conjoined heads:

(37) the national plans and programs

Section 3.3.3	OK	Sección 3.3.3
Greater control of pension benefits is needed.	OK	Se necesita mayor control de prestaciones de jubilación.
With this type of pension benefits are higher.	OK	Con este tipo de jubilación beneficios son mayores.
An increase in pension benefits everyone.	OK	Un aumento de jubilación beneficia a todos.
Section 5	OK	Sección 5
(1) It is clear that chronic malnutrition and hunger are linked to more general problems of poverty, inequitable distribution of land and other assets, underdevelopment, and adverse agroclimatic conditions.	OK	(1) Es claro que +la malnutrición+ crónica y hambre están vinculadas a problemas más generales de pobreza, distribución desigual de tierra y otros recursos, subdesarrollo, y condiciones agroclimáticas adversas.
(1') It is clear that chronic malnutrition and hunger are linked to more general problems that have already been discussed.	OK	(1') Es claro que +la malnutrición+ crónica y hambre están vinculadas a problemas más generales que ya se han tratado.

Figure 6. Examples from text

(32) relevant geographic, climatic, demographic, political, economic and social characteristics	OK	(32) características relevantes geográficas, climáticas, demográficas, políticas, económicas y sociales
(33) safe, effective use of pesticides	OK	(33) uso seguro, efectivo de plaguicidas
(34) highly qualified and very dedicated staff	OK	(34) personal altamente calificado y muy dedicado
(35) 15 subregional and national production centers	OK	(35) 15 centros de producción subregionales y nacionales
(36) morbidity and mortality statistics	OK	(36) estadísticas de morbilidad y de mortalidad
(37) the national plans and programs	OK	(37) los planes y programas nacionales
(38) health policy and strategies	OK	(38) política y estrategias de salud
(39) the university students	OK	(39) los estudiantes universitarios
(40) insect control	OK	(40) control de insectos
(41) the conditioning factor	OK	(41) el factor de condicionamiento
(42) the dividing line	OK	(42) la línea divisoria
(43) approximately 5 months' imports	NO	(43) las importaciones de aproximadamente 5 meses
(44) a mixture of 60 ml anhydrous diethyl ether and 1 ml methanol	OK	(44) una mezcla de 60 ml de éter dietílico anhidro y 1 ml de metanol
(45) one of the most important areas	OK	(45) una de las áreas más importantes
(46) three thousand physicians and three million people	OK	(46) tres mil médicos y tres millones de personas
(47) the industrial workers' mental ability	OK	(47) la capacidad mental de los obreros industriales
(48) the elderly, the handicapped or the working population	OK	(48) los ancianos, los minusválidos o la población activa
(49) the poorest	OK	(49) los más pobres
(50) some additional ones	OK	(50) algunos adicionales
(51) oral rehydration therapy (ORT) manuals	OK	(51) manuales de terapia de rehidratación oral (TRO)
(52) 3-*nitraniline-4-sulfonic acid	SD NO	(52) ácido 3-nitraniline-4-sulfónico
(53) central/regional mini-courses	OK	(53) mini-cursos centrales/regionales
(54) iron containing foods	OK	(54) alimentos que contienen hierro
(55) a three generation reproduction study	OK	(55) un estudio de reproducción de tres generaciones
(56) other nutrition-related diseases	OK	(56) otras enfermedades relacionadas con nutrición
(57) small-headed open-pollinated Chinese cabbage	OK	(57) repollo chino de cabeza pequeña de polinización abierta

Figure 7. Noun phrase

(38) health policy and strategies

A noun premodifier synthesized as an adjective:

(39) the university students

A noun premodifier synthesized in the plural:

(40) insect control

A gerundive premodifier synthesized as a noun:

(41) the conditioning factor

A gerundive premodifer synthesized as an adjective:

(42) the dividing line

Phrasal quantifiers:

(43) approximately 5 months' imports

(44) a mixture of 60 ml anhydrous diethyl ether and 1 ml methanol

(45) one of the most important areas

(46) three thousand physicians and three million people

A phrasal possessive:

(47) The industrial workers' mental ability

An adjectival noun-phrase head:

(48) the elderly, the handicapped or the working population

(49) the poorest

A pronominal noun-phrase head:

(50) some additional ones

Embedded parenthetical material (synthesized in correct grammatical position):

(51) oral rehydration therapy (ORT) manuals

A long hyphenated premodifier:

(52) 3-nitraniline-4-sulfonic acid

A hyphenated head with slashed premodifiers:

(53) central/regional mini-courses

A hyphenated modifier consisting of noun plus present participle:

(54) iron-containing foods

A hyphenated modifier consisting of quantifier plus noun:

(55) a three-generation reproduction study

A hyphenated modifier consisting of noun plus past participle:

(56) other nuttrition-related diseases

(57) small-headed open-pollinated Chinese cabbage [sic!]

8.3.2 Verb phrases

Figure 8 shows how ENGSPAN has dealt with the typical problems of Spanish verb phrase synthesis evoked by the sentences below.

Generation of SE with the passive:

(58) The agency [Ø] has defined a set of protocols, lines, modems, etc. which are recommended depending on the type of access or support necessary.

Rearrangement of adverbs:

(59) These projects have always been adequately managed.

Incomplete verb phrase:

(60) The delegates were not present, but the others were.

(61) The attorney was ready, but the judge wasn't.

Emphatic DO:

(62) On the other hand, the government does need support from all sectors.

(58) The agency has defined a set of protocols, lines, modems, etc. which are recommended depending on the type of access or support necessary. OK	(58) La agencia ha definido un conjunto de protocolos, líneas, modems, etc. que se recomiendan dependiendo del tipo de acceso o apoyo necesario.
(59) These projects have always been adequately managed. OK	(59) Estos proyectos siempre se han administrado adecuadamente.
(60) The delegates were not present, but the others were. OK	(60) Los delegados no estaban presentes, pero los otros sí.
(61) The attorney was ready, but the judge wasn't. OK	(61) El abogado estaba preparado, pero el juez no.
(62) On the other hand, the government does need more support from all sectors. OK	(62) Por otro lado, el gobierno sí necesita más apoyo de todos los sectores.
(63) Documenting need is inherently difficult. OK	(63) La documentación de necesidad es intrínsecamente difícil.
(64) The patients are eager to participate in the study, but the questionnaires are too complex. OK	(64) Los pacientes están ansiosos por participar en el estudio, pero los cuestionarios son demasiado complejos.
(65) The delegate stated that his country felt that the resolution should be approved unanimously. OK TU	(65) El delegado declaró que su país creía que la resolución se debe aprobar por unanimidad.
(66) The administrator may not have time to handle special requests. OK	(66) El administrador quizás no tenga tiempo para manejar solicitudes especiales.
(67) The nurse warmed up the milk for the baby. OK	(67) La enfermera calentó la leche para el bebé.
(68) The nurse gave them hot tea to warm them up. OK	(68) La enfermera les dió té caliente para calentarlos.

Figure 8. Verb Phrase

Nominalization of participial phrase:

(63) Documenting need is inherently difficult.

Use of ser vs. estar

(64) The patients are eager to participate in the study, but the questionnaires are too complex.

Use of preterit vs. imperfect (trigger indicated by broken underline):

(65) The delegate stated that his country felt that the resolution should be approved unanimously.

Subjunctive triggered by modal:

(66) The administrator may not have time to handle special requests.

Phrasal verbs:

(67) The nurse warmed up the milk for the baby.

(68) The nurse gave them hot tea to warm them up.

8.3.3 Clause-level solutions

Figure 9 shows how ENGSPAN has solved the following structural problems at the level of the clause. (Referents and triggers indicated with broken underline.)

Subject pronoun deletion:

(69) They should realize that they have a responsibility to the community they serve.

Synthesis of pronouns:

(70) The administrator receives the mail and then distributes it within the office.

(71) We sent you the report yesterday.

(72) The defendant [Ø] pleaded guilty after he consulted with his lawyer.

(73) The secretary had the report, but couldn't show it to us.

(69) They should understand that they have a responsibility to the community they serve.	OK	(69) Deben comprender que tienen una responsabilidad a la comunidad que ellos sirven.
(70) The administrator receives the mail and then distributes it within the office.	OK	(70) El administrador recibe el correo y entonces lo distribuye dentro de la oficina.
(71) We sent you the report yesterday.	OK	(71) Le enviamos el informe ayer.
(72) The defendant pleaded guilty after he consulted with his lawyer.	OK	(72) El demandado se declaró culpable después que consultó con su abogado.
(73) The secretary had the report, but couldn't show it to us.	OK	(73) El secretario tenía el informe, pero no podía mostrárnoslo.
(74) These activities will benefit the children of Africa.	OK	(74) Estas actividades beneficiarán a los niños de Africa.
(75) The delegate presented the proposal to the chairman of the committee.	OK	(75) El delegado le presentó la propuesta al presidente del comité.
(76) We lack the necessary resources.	OK	(76) Nos faltan los recursos necesarios.
(77) A number of groups in Argentina would like to see such arrangements made.	OK	(77) A varios grupos en Argentina les gustaría ver dichos arreglos hechos.
(78) Beans are a good source of protein.	OK	(78) El frijol es una fuente buena de proteína.
(79) The recommendations of the consultant are very timely.	OK	(79) Las recomendaciones del consultor son muy oportunas.
(80) The potential to attract and retain an excellent public health care does exist, but higher salaries are required.	OK	(80) Si existe el potencial para atraer y retener una atención excelente de salud pública, pero se requieren salarios mayores.
(81) The problems we are to consider are difficult and will not be resolved easily.	OK	(81) Los problemas que nosotros consideraremos son difíciles y no se resolverán fácilmente.
(82) The patient was given his medicine before every meal.	OK	(82) Al paciente se le dió su medicina antes de cada comida.
(83) Population groups living in rural areas have been singled out for special attention.	OK	(83) Grupos de población que viven en áreas rurales se han individualizado para atención especial.
(84) There is no limit to the number of documents that can be sent at the same time.	OK	(84) No hay límite al número de documentos que se pueden enviar al mismo tiempo.
(85) Has there been an increase this year?	NO	(85) ¿Ha habido un aumento este año?
(86) When did the problem of the Falkland Islands begin?	OK	(86) ¿Cuándo empezó el problema de las Islas Falkland?
(87) Were the results reported to the authorities?	OK	(87) ¿Los resultados se informaron a las autoridades?
(88) Have the students taken the exam?	OK TU	(88) ¿Han rendido los estudiantes el examen?

Figure 9. The clause level

Generation of the personal a:

(74) These activities will benefit [Ø] the children of Africa.

Generation of the dative of interest:

(75) The delegate [Ø] presented the proposal to the chairman of the committee.

Reversal of subject and object:

(76) We lack the necessary resources.

(77) A number of groups in Argentina would like to see such arrangements made.

Subject-verb agreement (subject is singular in Spanish):

(78) Beans are a good source of protein.

Subject-complement agreement:

(79) The recommendations of the consultant are very timely.

Fronting of the verb phrase:

(80) The potential to attract and retain an excellent public health care staff does exist, but higher salaries are required.

Synthesis of BE + infinitive as future tense:

(81) The problems we are to consider are difficult and will not be resolved easily.

Use of impersonal SE for bitransitive passive constructions:

(82) The patient [Ø] was given his medicine before every meal.

Synthesis of relative clause for participial phrase:

(83) Population groups living in rural areas have been singled out for special attention.

Synthesis of negative verb phrase:

(84) There is no limit to the number of documents that can be sent at the same time.

Synthesis of interrogative sentences:

(85) Has there been an increase this year?

(86) When did the problem of the Falkland Islands begin?

(87) Were the results reported to the authorities?

(88) Have the students taken the exam?

8.3.4 Sentence-level solutions

Figure 10 shows how ENGSPAN has solved the structural problems raised by the examples below at the level of the sentence. (Referents and triggers indicated by broken underline.)

Agreement of relative pronoun and complement with main subject:

(89) The cases for which these rules were intended have not been included.

Agreement in a reduced relative clause and insertion of relative marker:

(90) The system administrators must use the set of programmed routines designed for the system [Ø] they are using.

Subjunctive in a dependent clause:

(91) Unless industrial countries take constructive steps to deal with these problems, their own recovery will be undermined.

(92) We hope that the package arrives on Monday.

Subject-raising:

(93) The government wants the people to buy the apartments.

(94) The committee asked the chairman to postpone the meeting.

(95) The cause of the problem is believed to be a leak in the underground pipeline.

(96) Official intermediaries other than commercial banks were allowed to lend up to 20 percent of their portfolio at freely determined interest rates.

(89) The cases for which these rules were intended have not been included.	OK	(89) No se han incluido los casos para los cuales estas reglas estaban concebidas.
(90) The system administrators must use the set of programmed routines designed for the system they are using.	OK	(90) Los administradores de sistema deben emplear el conjunto de rutinas programadas diseñadas para el sistema que ellos están empleando.
(91) Unless industrial countries take constructive steps to deal with these problems, their own recovery will be undermined.	OK TU TU	(91) A menos que países industriales tomen medidas constructivas para tratar estos problemas, se minará su propia recuperación.
(92) We hope that the package arrives on Monday.	OK	(92) Esperamos que el paquete llegue el lunes.
(93) The government wants the people to buy the apartments.	OK	(93) El gobierno desea que las personas compren los departamentos.
(94) The committee asked the chairman to postpone the meeting.	OK TU	(94) El comité le pidió al presidente que aplazara la reunión.
(95) The cause of the problem is believed to be a leak in the underground pipeline.	OK	(95) La causa del problema se cree que es una fuga en las cañerías subterráneas.
(96) Official intermediaries other than commercial banks were allowed to lend up to 20 percent of their portfolio at freely determined interest rates.	OK	(96) A intermediarios oficiales con excepción de bancos comerciales se les permitió prestar hasta 20 por ciento de su cartera a tasas de interés libremente decididas.

Figure 10. The sentence level

9. FUTURE DIRECTIONS

In the final months of the grant period, Dr. Naomi Sager examined the output of the parser and some unedited translations produced by ENGSPAN. Dr. William Cressey studied samples of raw Spanish output and the corresponding postedited versions. In her report, Dr. Sager stated that "the number and complexity of constructions handled by the grammar has increased to the point where relatively few types of input sentences are not parsed adequately for subsequent translation." Dr. Cressey found the lexical and morphological aspects of the Spanish output to be adequate, and also noted "many accomplishments in the area of syntactic structure." In his section devoted to recommendations for future development, Dr. Cressey pointed out that the project team is in the best position to identify "what aspects of the system, if developed further, would be most likely to yield the greatest overall improvement per effort expended."

PAHO's success in the development and implementation of machine translation has sparked the interest of other international agencies and governments and nonprofit institutions in the Americas. The availability of ENGSPAN on a microcomputer would provide an effective and low-cost means of making health information available in Spanish-speaking countries. At the end of the grant period a study was done on the feasibility of adapting ENGSPAN to the microcomputer environment. Following this study, a decision was taken to seek additional resources in order to convert the PL/1 source code to the 'C' programming language and develop a complete software package for the microcomputer.

ACKNOWLEDGMENTS

Special thanks are expressed to Susana Santangelo for her patience with ENGSPAN's firsts attempts at Spanish and her unselfish zeal in conveying her knowledge of Spanish and experience as a translator to the system's dictionaries and its developers.

The authors would like to recognize the contribution made by Dr. R. Ross Macdonald in the early stages of the development of ENGSPAN. Dr. Macdonald's untimely death in 1983 deprived us of his further participation in the project.

The advice and encouragement received from our consultants, Dr. Naomi Sager and Dr. William Cressey, are greatly appreciated.

We would also like to thank Lucretia Vanderwende and Patricia Schmid, who did the bulk of the source dictionary coding, and Plinio de Almeida, PAHO's system programmer, for his assistance in porting the system to AID's computer.

Finally, we owe a debt of gratitude to Dr. John Daly, our Project Officer, for his enthusiasm and confidence in the project, and to the U.S. Agency for International Development for its financial support.

REFERENCES

Bates, Madeleine 1978 The theory and practice of augmented transition network grammars. In: Natural Language Communication with Computers, ed. by Leonard Bole. New YorkN Springer. pp. 191-260.

León, Marjorie 1984 Development of English-Spanish machine translation. Paper presented at the International Conference on Machine Translation (Cranfield, England, 13-15 February 1984).

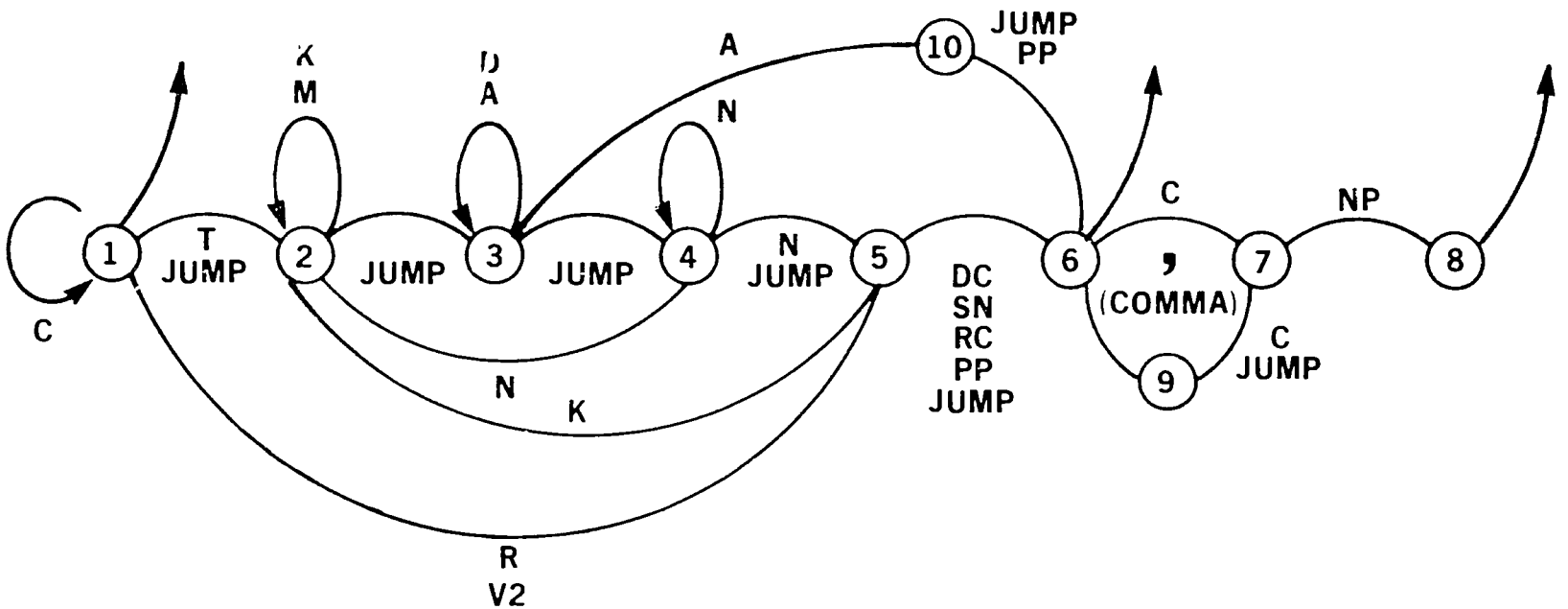
Sager, Naomi 1981 Natural Language Information Processing: A Computer Grammar of English and Its Applications. Addison-Wesley, Reading (Mass.).

Vasconcellos, Muriel, and Marjorie León 1985 SPANAM and ENGSPAN: Machine Transla tion at the Pan American Health Organization. Computational Linguistics 11(2-3):122-136.

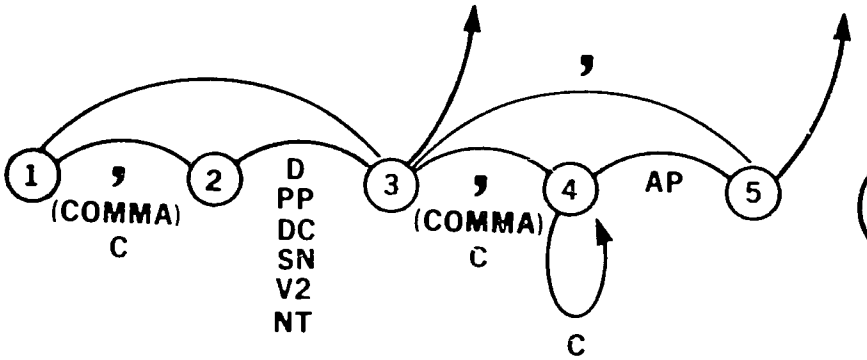
Winograd, Terry 1983 Language as a Cognitive Process. vol. 1. Syntax. Addison=Wesley, Reading (Mass.).

Woods, William A. 1970 Transition network grammars for natural language analysis. Communications of the Association for Computing Machinery 13(10):591-606.

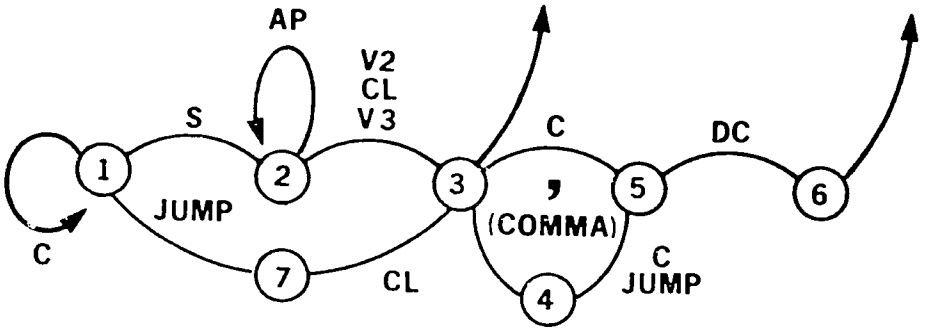
NOUN PHRASE NETWORK



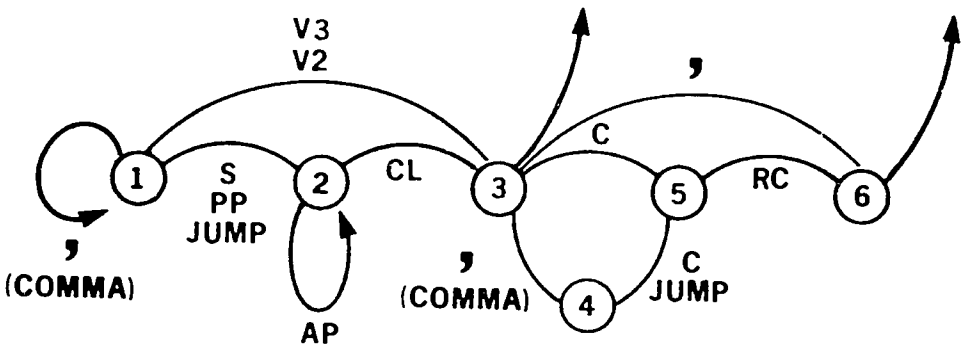
Key: A = adjective, AP = adverbial phrase, C = coordinate conjunction, CL = clause, D = adverb, DC = dependent clause, G = not (negative), K = numerative, M = modifier, N = noun, NP = noun phrase, NT = time noun phrase, P = preposition, PP = prepositional phrase, R = pronoun, RC = relative clause, S = subordinate conjunction, SN = sentence nominalization (for-to complementation), T = determiner, V = verb, VI = imperative verb phrase, VP = verb phrase, V2 = non-finite verb phrase, V3 = verb phrase with "be" deleted, X = auxiliary.



ADVERBIAL PHRASE NETWORK

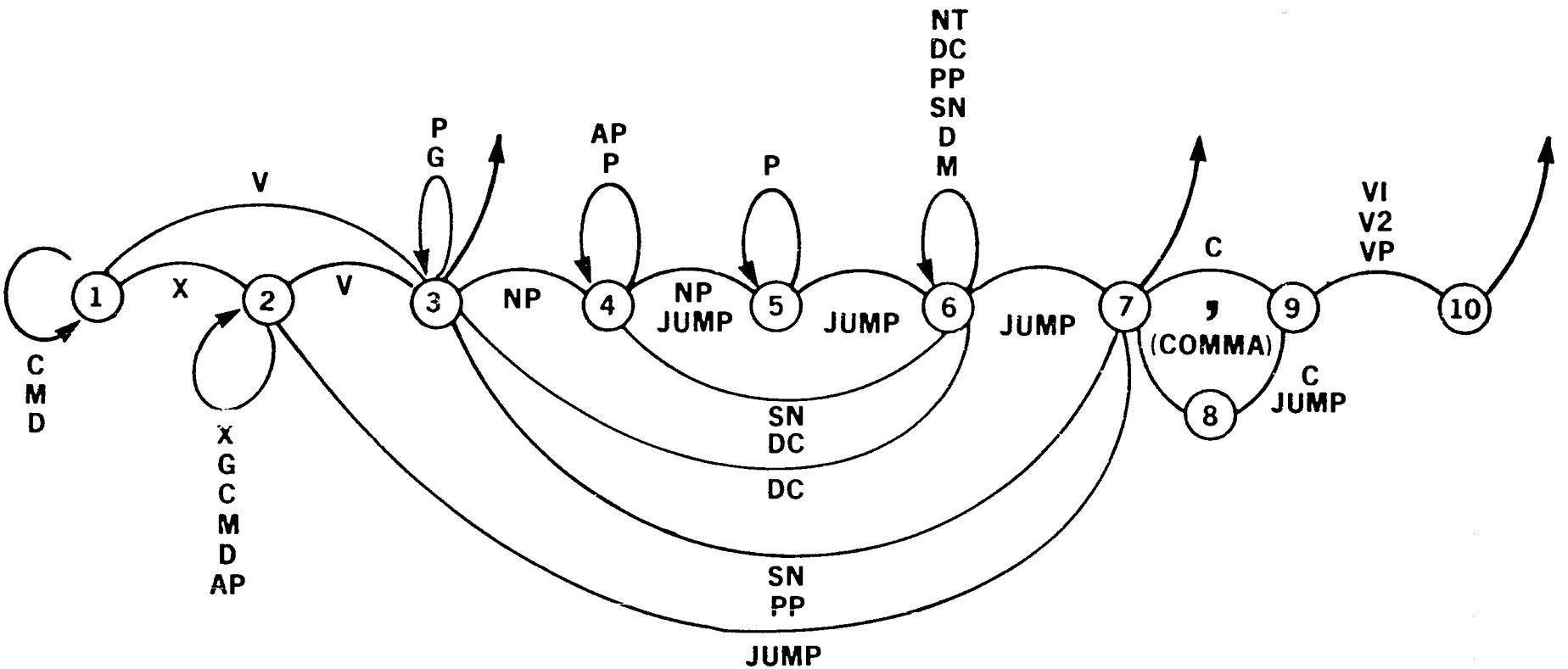


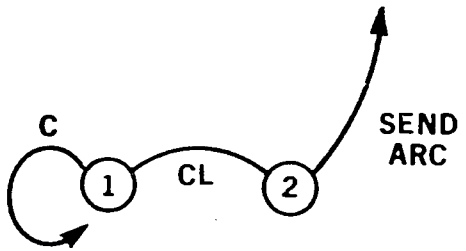
DEPENDENT CLAUSE NETWORK



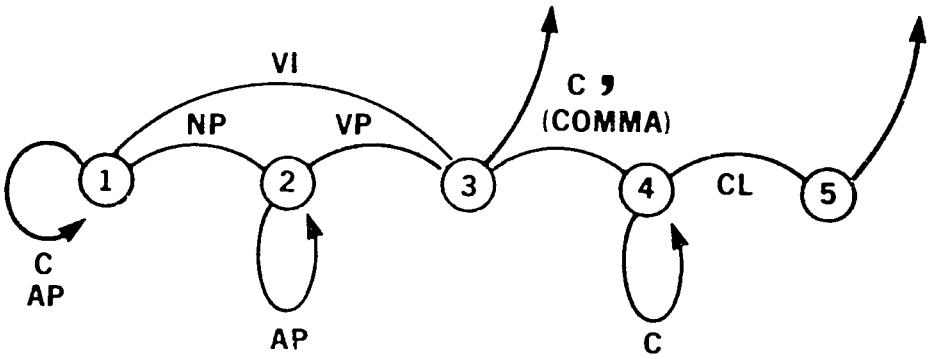
RELATIVE CLAUSE NETWORK

VERB PHRASE NETWORK

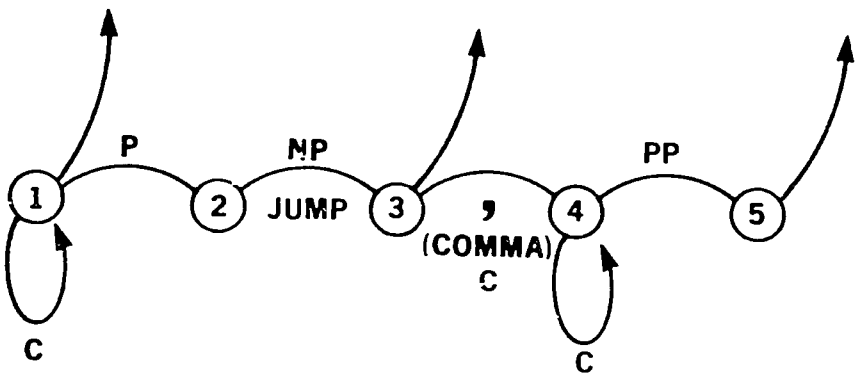




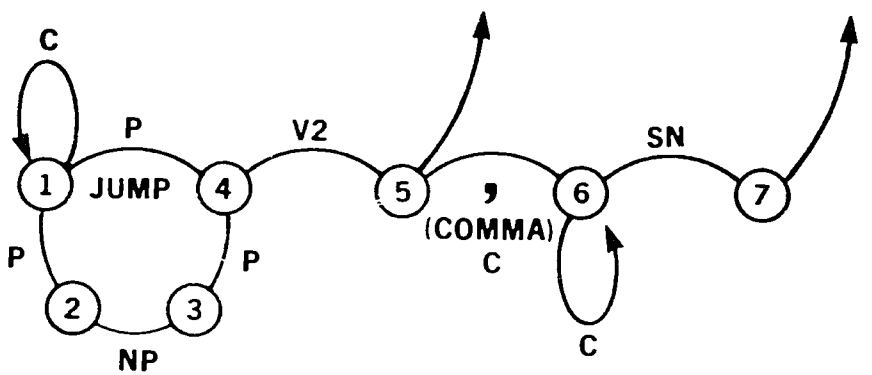
SENTENCE NETWORK



CLAUSE NETWORK



PREPOSITIONAL PHRASE NETWORK



COMPLEMENTATION NETWORK

APPENDIX B

PHRASE INFORMATION STRUCTURES FOR THE DEMONSTRATION SENTENCE

```
ST  PHRASE  1  COMPLETE
CATEGORY     ROLE     LOCATION  MODIFIES
   CL                    2


CL  PHRASE  2  COMPLETE
CALLED BY PHRASE  1
CATEGORY     ROLE     LOCATION  MODIFIES
   NP         S           3
   VP                     4


NP  PHRASE  3  COMPLETE
CALLED BY PHRASE  2
CATEGORY     ROLE     LOCATION  MODIFIES
   R          H           1                IT
HEAD= IT
NUMBER=SG


VP  PHRASE  4  COMPLETE
CALLED BY PHRASE  2
CATEGORY     ROLE     LOCATION  MODIFIES
   V          V           2                IS
   NP         C           5
   DC                     6
HEAD= IS
NUMBER=3S
VERBFEATURES: FINITE


NP  PHRASE  5  COMPLETE
CALLED BY PHRASE  4
CATEGORY     ROLE     LOCATION  MODIFIES
   A          H           3         1      CLEAR
HEAD= CLEAR


DC  PHRASE  6  COMPLETE
CALLED BY PHRASE  4
CATEGORY     ROLE     LOCATION  MODIFIES
   S                      4                THAT
   CL                     7
HEAD= IS
```

```
CL  PHRASE  7  COMPLETE
CALLED BY PHRASE  6
CATEGORY     ROLE     LOCATION  MODIFIES
   NP         S          8
   VP                   10
NUMBER= D


NP  PHRASE  8  COMPLETE
CALLED BY PHRASE  7
CATEGORY     ROLE     LOCATION  MODIFIES
   A          D          5          6      CHRONIC
   N          H          6                 MALNUTRITION
   C                     7                 AND
   NP                    9
HEAD=MALNUTRITION
NUMBER=PL
CONHS= 06,08


NP  PHRASE  9  COMPLETE
CALLED BY PHRASE  8
CATEGORY     ROLE     LOCATION  MODIFIES
   N          H          8                 HUNGER
HEAD= HUNGER
NUMBER=SG
CONHS= 08


VP  PHRASE 10  COMPLETE
CALLED BY PHRASE  7
CATEGORY     ROLE     LOCATION  MODIFIES
   V          V          9                 ARE
   NP         C         11
HEAD= ARE
NUMBER= 3P
VERBFEATURES: FINITE


NP  PHRASE 11  COMPLETE
CALLED BY PHRASE 10
CATEGORY     ROLE     LOCATION  MODIFIES
   A          H         10          6      LINKED
   PP                   12
HEAD= LINKED

PP  PHRASE 12  COMPLETE
CALLED BY PHRASE 11
CATEGORY     ROLE     LOCATION  MODIFIES
   P                    11                 TO
   NP         O         13
HEAD= PROBLEM
NUMBER= PL
```

```
NP  PHRASE 13  COMPLETE
CALLED BY PHRASE 12
CATEGORY     ROLE     LOCATION  MODIFIES
   M          D          12        13      MORE
   A          D          13        14      GENERAL
   N          H          14                PROBLEM
   PP         D          14
HEAD= PROBLEM
NUMBER= PL


PP  PHRASE 14  COMPLETE
CALLED BY PHRASE 13
CATEGORY     ROLE     LOCATION  MODIFIES
   P                     15                OF
   NP         O          15
HEAD= POVERTY
NUMBER= PL
CONNS= 16,19,26,31

NP  PHRASE 15  COMPLETE
CALLED BY PHRASE 14
CATEGORY     ROLE     LOCATION  MODIFIES
   N          H          16                POVERTY
   ,                     17
   NP                    16
HEAD= POVERTY
NUMBER= PL
CONNS= 16,19,26,31


NP  PHRASE 16  COMPLETE
CALLED BY PHRASE 15
CATEGORY     ROLE     LOCATION  MODIFIES
   A          D          18        19      INEQUITABLE
   N          H          19                DISTRIBUTION
   PP         D          17
   ,                     25
   NP                    26
HEAD= DISTRIBUTION
NUMBER= PL
CONNS= 19,26,31


PP  PHRASE 17  COMPLETE
CALLED BY PHRASE 16
CATEGORY     ROLE     LOCATION  MODIFIES
   P                     20                OF
   NP         O          18
HEAD= LAND
NUMBER= PL
```

```
NP  PHRASE 18  COMPLETE
CALLED BY PHRASE 17
CATEGORY     ROLE     LOCATION  MODIFIES
   N          H          21                LAND
   C                     22                AND
   NP                    19
HEAD= LAND
NUMBER= PL
CONHS= 21,24

NP  PHRASE 19  COMPLETE
CALLED BY PHRASE 18
CATEGORY     ROLE     LOCATION  MODIFIES
   T          T          23        24      OTHER
   N          H          24                ASSET
HEAD= ASSET
NUMBER= PL
CONHS= 24


NP  PHRASE 21  COMPLETE  REUSED
CALLED BY PHRASE 26
CATEGORY     ROLE     LOCATION  MODIFIES
   A          D          29        31      ADVERSE
   A          D          30        31      AGROCLIMATIC
   N          H          31                CONDITION
HEAD= CONDITION
NUMBER= PL
CONHS= 31

NP  PHRASE 26  COMPLETE  REUSED
CALLED BY PHRASE 16
CATEGORY     ROLE     LOCATION  MODIFIES
   N          H          26                UNDERDEVELOPMENT
   ,                     27                ,
   C                     28                AND
   NP                    21
HEAD= UNDERDEVELOPMENT
NUMBER= PL
CONHS= 26,31
```

APPENDIX C

PHRASE INFORMATION STRUCTURES FOR A RELATIVE CLAUSE

NP PHRASE 13 COMPLETE
CALLED BY PHRASE 12

CATEGORY	ROLE	LOCATION	MODIFIES	
M	D	12	13	MORE
A	D	13	14	GENERAL
N	H	14		PROBLEM
RC		14		

HEAD= PROBLEM
NUMBER= PL

RC PHRASE 14 COMPLETE
CALLED BY PHRASE 13

CATEGORY	ROLE	LOCATION	MODIFIES	
S	O	15		THAT
CL		15		

HEAD= PROBLEM

CL PHRASE 15 COMPLETE
CALLED BY PHRASE 14

CATEGORY	ROLE	LOCATION	MODIFIES
NH	S	16	
VP		17	

NUMBER= R

NH PHRASE 16 COMPLETE
CALLED BY PHRASE 15

CATEGORY	ROLE	LOCATION	MODIFIES	
M	D	12		MORE
A	D	13		GENERAL
N	H	14		PROBLEM

HEAD= PROBLEM
NUMBER= PL

VP PHRASE 17 COMPLETE
CALLED BY PHRASE 15

CATEGORY	ROLE	LOCATION	MODIFIES	
X	H	16		HAVE
D		17		ALREADY
X	B	18		BEEN
V	V	19		DISCUSS

HEAD= DISCUSS
NUMBER= 3P
VERBFEATURES: FINITE PASSIVE PERFECT

DESCRIPTION OF THE COUNTRY AND ITS HEALTH SITUATION	OK	DESCRIPCION DEL PAIS Y SU SITUACION DE SALUD
Country: SURINAME - Date: November 1984	OK OK OK OK	País: SURINAME - Fecha: Noviembre de 1984
Country Narrative Summary	OK	Resumen Narrativo de País
Suriname	OK	Suriname
0. Relevant Geographic, Climatic, Demographic, Political, Economic and Social Characteristics	OK	0. Características Relevantes Geográficas, Climáticas, Demográficas, Políticas, Económicas y Sociales
0.1 Geography and Climate - Suriname is situated on the north coast of the South American Continent and shares borders with Guyana, Brazil and French Guiana. The land area totals about 163,000 km^2 including approximately 90,000 km^2 of disputed area with Guyana and French Guiana. Most of the population lives along the coastal strip which is about 3% of the total land area. The sparsely populated interior consists mainly of dense tropical forest with small settlements along the rivers. The climate is tropical with an average temperature of *26¡C; the mean annual rainfall is 2,000 to 2,500 mm. There has been no major natural disaster in recent history.	OK OK TU OK OK OK TU SD OK OK OK	0.1 Geografía y Clima - Suriname está situado en la costa norte del Continente sudamericano y comparte fronteras con Guyana, Brasil y Guayana Francesa. La superficie terrestre totaliza cerca de 163,000 km^2 incluyendo aproximadamente 90,000 km^2 de área disputada con Guyana y Guayana Francesa. La mayoría de la población vive a lo largo de la franja costera que es cerca de 3% de la superficie terrestre total. El interior dispersamente poblado consta principalmente de bosque tropical denso con asentamientos pequeños a lo largo de los ríos. El clima es tropical con una temperatura media de 26¡C; la precipitación anual media es 2,000 a 2,500 mm. No ha habido desastre principal natural en historia reciente.
0.2 Demographic Factors - The population is about 360,000 and consists mainly of Hindustani, Creole, Javanese, Amerindians, and Chinese. About two thirds of the population is concentrated in the capital city of *Paramaribo.	OK OK TU SD OK	0.2 Factores Demográficos - La población es cerca de 360,000 y consta principalmente de indostani, criollo, javanés, amerindios, y chino. Cerca de dos tercios de la población está concentrada en la ciudad capital de Paramaribo.
0.3 Political System - Administratively the country is divided into nine districts including the capital district of *Paramaribo. These districts are administered by Commissioners with limited authority.	OK SD OK CK	0.3 Sistema Político - Administrativamente el país está dividido en nueve distritos incluyendo el distrito capital de Paramaribo. Estos distritos son administrados por Comisionados con autoridad limitada.
0.4 Economy - Suriname has an unevenly developed economy with a few capital-intensive modern enterprises dominated by foreign capital and producing most of the country's output and exports. In addition to the two bauxite-producing companies, there are also large agricultural enterprises engaged in rice research, cultivation, and production and in wood processing.	OK OK TU OK TU	0.4 Economía - Suriname tiene una economía desigualmente desarrollada con unas pocas empresas modernas capitales-intensivas dominadas por capital extranjero y generando la mayoría de producción y exportaciones del país. Además de las dos empresas productoras de bauxita, hay también empresas grandes agrícolas ocupadas en investigaciones de arroz, cultivo, y producción y en procesamiento de madera.

Despite the large-scale immigration in recent years and the rapid growth of the public sector, unemployment is still around 18% of the labor force. The overall balance of payments position has remained relatively strong in recent years. This is mainly the result of the imposition of the levy on bauxite production and a substantial increase in foreign aid. Official reserves at the end of 1982 were equivalent to approximately 5 months' imports.

0.5 Social and Family Characteristics - The literacy rate is about 83% for men and about 75% for women. Through literacy programs now being executed, there should be 100% literacy within a few years. Attendance at primary school is compulsory and remains high throughout secondary school. The University of Suriname includes Faculties of Medicine, Social Sciences, Engineering, and Law. A number of housing programs have been implemented by the Government to improve public housing. The road and telecommunication systems are extensive and of good quality. The national carrier, Suriname Airways, provides air links with neighboring countries, the United States of America, and Europe. The Government of Suriname has one television station and one radio station.

1. Development of the Health System

1.1 Health Policy and Strategy - After the Revolution of February 1980, there has been a reorientation of the developmental policies based on the principles of renovation in the political administrative order, the socioeconomic order, and the education system. Within this framework, health policies were defined with the Alma-Ata Declaration as a background. Health is to be available, accessible, affordable, and acceptable to the population and the policy focus is directed at the development of primary health care. The participation of the community in this process is essential.

A pesar de la inmigración en gran escala en años recientes y el crecimiento rápido del sector público, el desempleo es todavía alrededor de 18% de la fuerza laboral. La posición global de balanza de pagos ha permanecido relativamente fuerte en años recientes. Esto es principalmente el resultado de la imposición de la recaudación en producción de bauxita y un aumento sustancial en ayuda extranjera. Las reservas oficiales al final de 1982 fueron equivalentes a las importaciones de aproximadamente 5 meses.

0.5 Características Sociales y Familiares - La tasa de alfabetismo es cerca de 83% para los hombres y cerca de 75% para mujeres. A través de programas de alfabetismo ahora ejecutándose, debe haber el alfabetismo de 100% dentro de unos pocos años. La asistencia en escuela primaria es obligatoria y permanece alta en todo escuela secundaria. La Universidad de Suriname incluye a Facultades de Medicina, Ciencias Sociales, Ingeniería, y Ley. Varios programas de vivienda han sido instrumentados por el Gobierno para mejorar vivienda pública. El camino y sistemas de telecomunicación son exhaustivos y de calidad buena. El portador nacional, *Vías Aéreas de Suriname, proporciona nexos de aire con países vecinos, los Estados Unidos de América, y Europa. El Gobierno de Suriname tiene una estación de televisión y una estación de radio.

1. Desarrollo del Sistema de Salud

1.1 Política y Estrategia de Salud - Después de la Revolución de febrero de 1980, ha habido una reorientación de las políticas del desarrollo basadas en los principios de renovación en el orden administrativo político, el orden socioeconómico, y el sistema de educación. Dentro de este marco, las políticas de salud fueron definidas con la Declaración de Alma-Ata como unos antecedentes. La salud estará disponible, accesible, alcanzable, y aceptable a la población y el foco político está dirigido al desarrollo de atención primaria de salud. La participación de la comunidad en este proceso es esencial.

The Government is committed to the strategy of health for all through its national health policies, and there is a high degree of congruence between national health policies and the Regional Strategies for attaining health for all by the year 2000. Population groups living in rural areas have been singled out for special attention in the development of primary health care. In view of their vulnerability, mothers and children also receive special attention. No special programs have been implemented for the elderly, the handicapped or the working population. A number of medical and social provisions, however, have been in existence for a long time for these groups of the population.

1.2 The Health System - The Ministry of Health has a regulatory, legislative, coordinating, and initiating role. It is also directly responsible for the health care of approximately one fourth of the population through its health institutions. This comprises two major hospitals, a regional health service responsible for primary health care in the coastal area, and preventive medicine services. The private sector is made up of two hospitals and a large number of private medical practitioners. A privately organized but publicly funded foundation is responsible for primary health care in the interior of the country where 10% of the population lives. In 1981, a State Health Insurance Foundation started, which at this stage, covers curative medical services for state employees.[1] The next stage comprising medical services for the underprivileged population will start in 1984 and will cover another significant part of the population. In a later stage, small independent employers will be invited to join the scheme. Larger employers already have insurance schemes for their employees and families and often provide their own health services. It is estimated that the majority of the total population will be covered by the State Health Insurance eventually.

El Gobierno está comprometido a la estrategia de salud para todos a través de sus políticas nacionales de salud, y hay un grado alto de congruencia entre políticas nacionales de salud y las Estrategias Regionales para logro de salud para todos en el año 2000. Grupos de población que viven en áreas rurales se han individualizado para atención especial en el desarrollo de atención primaria de salud. En vista de su vulnerabilidad, madres y niños también reciben atención especial. Ningunos programas especiales se han instrumentado para los ancianos, los impedidos o la población activa. Varias provisiones médicas y sociales, sin embargo, han sido en existencia por un tiempo largo para estos grupos de la población.

1.2 El Sistema de Salud - El Ministerio de Salud tiene una función reguladora, legislativa, coordinadora, y de iniciación. Es también directamente responsable de la atención de la salud de aproximadamente un cuarto de la población a través de sus instituciones de salud. Esto comprende dos hospitales principales, un servicio de salud regional responsable de atención primaria de salud en el área costera, y servicios de medicina preventiva. El sector privado está constituido de dos hospitales y un gran número de médicos privados. Una fundación privadamente organizada pero públicamente financiada es responsable de atención primaria de salud en el interior del país donde vive un 10% de la población. En 1981, una Fundación de Seguros de Salud Estatal comenzó, que en esta etapa, cubre servicios médicos curativos para empleados estatales. La próxima etapa que comprende servicios médicos para la población menos privilegiada comenzará en 1984 y cubrirá otra parte significativa de la población. En una etapa posterior, a empleadores pequeños independientes se les invitará que acojan el esquema. Empleadores más grandes ya tienen esquemas de seguro para sus empleados y familias y a menudo proporcionan sus propios servicios de salud. Se estima que la mayoría de la población total estará cubierta por el Seguro de Salud Estatal eventualmente.

Currently, the health information system is very deficient. OK
Separate subsystems exist with little or no coordination OK
or central collection. Specific plans exist to improve the OK
situation in 1984-1985. A national definition of coverage OK
has not been made. Three levels of care exist in the OK
country. In general, accessibility and referral have to be OK
acceptable, especially in view of the very low population
density in some areas. Over the past decade most attention OK TU
went to the development of adequate secondary and tertiary
care. Presently, more attention is being given to the OK TU
development of primary health care services.

1.4 Community Involvement - A number of efforts have been OK OK
undertaken to increase community participation in various
parts of the country. In addition, the Ministry of Health OK TU
continues to rely strongly on a traditional health
education department.

1.5 Mobilization of Resources - The country is dependent OK OK
on external supplies and equipment. National drug OK
registration was started in 1981. In 1983, the OK
Pharmaceutical Services Department of the Ministry of
Health was organized. A list of essential drugs with some OK
200 generic names was prepared in 1983 and is being used by
the State Health Insurance and government institutions. OK
National research and health technology policies are not
yet defined. There is one major research institute in the OK
health field while smaller units of the Ministry such as
the Central Laboratory have some research functions. A OK
well organized and documented medical library is offered at
the Faculty of Medicine, and a new Documentation Center is
being planned for the Ministry of Health. The development OK
of human resources in the health sector continues to be a
priority.

Actualmente, el sistema de información de salud es muy deficiente. Subsistemas separados existen con poca o ninguna coordinación o colección central. Planes específicos existen para mejorar la situación en 1984-1985. No se ha hecho una definición nacional de cobertura. Tres niveles de atención existen en el país. En general, la accesibilidad y referencia tienen que ser aceptables, especialmente en vista de la densidad demográfica muy baja en algunas áreas. En el último decenio la mayoría de atención fue al desarrollo de atención adecuada secundaria y terciaria. Actualmente, más atención se está prestando al desarrollo de servicios de atención primaria de salud.

1.4 Participación Comunitaria - Varios esfuerzos se han realizado para aumentar participación comunitaria en diversas partes del país. Además, el Ministerio de Salud continúa dependiendo firmemente de un departamento tradicional de educación de salud.

1.5 Movilización de Recursos - El país depende de suministros y equipo externo. El registro nacional de medicamentos se comenzó en 1981. En 1983, se organizó el Departamento de Servicios Farmoquímicos del Ministerio de Salud. Una lista de medicamentos esenciales con algunos 200 nombres genéricos se preparó en 1983 y está siendo empleado por el Seguro de Salud Estatal e instituciones del gobierno. Las investigaciones y políticas de tecnología de salud nacionales no son todavía definidas. Hay un instituto principal de investigaciones en el campo de salud mientras unidades más pequeñas del Ministerio como el Laboratorio Central tienen algunas funciones de investigaciones. Una biblioteca médica bien organizada y documentada se ofrece en la Facultad de Medicina, y un Centro nuevo de Documentación se está planificando para el Ministerio de Salud. El desarrollo de recursos humanos en el sector salud continúa siendo una prioridad.

Although there are enough well-qualified physicians PP
available, there continues to be a need for qualified mid-
level personnel, as well as for specific high-level
personnel, such as health administrators, sanitary
engineers, statisticians, and entomologists. The School of OK
Dental Auxillaries is continuing to produce graduates who
are starting to work in rural areas in cooperation with the
regional health services. In nursing, a Central School for SD PP
Nursing and Related Professions has been set up which will
gradually replace the various schools of nursing.[1]

Aunque hay suficientes médicos bien-calificados disponibles, continúa siendo una necesidad de personal calificado de medio nivel, así como para personal específico de nivel alto, como administradores de salud, ingenieros sanitarios, estadísticos, y entomólogos. La Escuela de Auxiliares Dentales está continuando produciendo graduados que están comenzando a trabajar en áreas rurales en cooperación con los servicios de salud regionales. En enfermería, una Escuela Central para Enfermería y Profesiones Relacionadas ha sido instalar que gradualmente reemplazará las diversas escuelas de enfermería.

1.6 Intersectoral Cooperation - The health sector OK OK
participates in programs and intersectoral development
projects mainly with the Departments of Education, Social
Affairs and Labor.

1.6 Cooperación Intersectorial - El sector salud participa en programas y proyectos de desarrollo intersectoriales principalmente con los Departamentos de Educación, Asuntos Sociales y Mano de Obra.

1.7 International Cooperation - PAHO/WHO technical OK OK
cooperation is provided in the following areas: OK
entomological surveillance, malaria eradication, Aedes
aegypti eradication, control of leprosy, occupational
health, planning and information systems, dental health,
environmental sanitation, and veterinary public health.

1.7 Cooperación Internacional - La cooperación técnica de la OPS/OMS se proporciona en las áreas siguientes: la vigilancia entomológica, erradicación de malaria, erradicación de Aedes aegypti, control de lepra, salud ocupacional, planificación y sistemas de información, la salud dental, el saneamiento ambiental, y salud pública veterinaria.

2. Health Status OK

2. Estado de Salud

2.1 Protection and Promotion of Health OK

2.1 Protección y Promoción de Salud

Food and Nutrition - Data from a recent nutritional OK OK
survey in 5 to 8 year olds do not show signs of an acute
malnutrition problem, although there may have been a
nutritional problem in earlier years. Plans for a OK
nutritional survey in a younger age group are underway. OK
Other nutrition-related diseases such as diabetes,
hypertension, and obesity are highly prevalent. No special OK
programs currently exist in this area.

Nutrición y Alimentación - Datos de una encuesta reciente nutricional en 5 a 8 años de edad no muestran signos de un problema +de malnutrición+ aguda, aunque puede haber habido un problema nutricional en años anteriores. Los planes para una encuesta nutricional en un grupo de edad más joven están en desarrollo. Otras enfermedades relacionadas con nutrición como diabetes, hipertensión, y obesidad son altamente prevalentes. Ningunos programas especiales actualmente existen en esta área.

Oral Health - Oral health, especially in school-age OK OK
children, is a serious reason for concern. A special PP
program to train dental auxiliaries for youth dental health
was started several years ago.[5]

Salud Oral - La salud oral, especialmente en niños en edad escolar, es una razón seria para inquietud. Un programa especial para adiestrar auxiliares dentales para salud dental de juventud se comenzó varios años atrás.[5]

At the central level, it has worked well, but outreach activities suffer from logistical and organizational problems. OK

Al nivel central, ha trabajado bien, pero actividades de extensión sufren de problemas logísticos y de organización.

Accidents - Traffic accidents are on the increase and are taking a high death toll especially among young adults. Accident prevention programs are being implemented by the Police Department. OK OK TU TU OK

Accidentes - Los accidentes de tránsito son en el aumento y están haciendo una mortalidad alta especialmente entre adultos jóvenes. Programas de prevención de accidentes están siendo instrumentados por el Departamento Policial.

Mental Health - Various mental health programs have been identified. Alcohol and drug abuse cause serious concern, although no exact data are available as to their extent. OK OK OK

Salud Mental - Se han identificado diversos programas de salud mental. Uso indebido de alcohol y drogas causa inquietud seria, aunque ningunos datos exactos están disponibles en cuanto a su grado.

2.2 Environmental Health OK

2.2 Salud Ambiental

Drinking Water - The National Program for Drinking Water Supply has the goal to expand to full coverage the already existing distribution system.[1] Only part of the water supply is being totally disinfected. OK PP OK

Agua Potable - El Programa Nacional para Abastecimiento de Agua Potable tiene la meta para ampliar a cobertura total el sistema de distribución ya existente.[1] Se está totalmente desinfectando sólo parte del abastecimiento de agua.

Excreta Disposal - Excreta disposal programs are offered at a lower level of coverage, but also receive full attention from the Government. OK OK

Eliminación de Excretas - Programas de eliminación de excretas se ofrecen a un nivel inferior de cobertura, pero también reciben atención total del Gobierno.

Solid Waste - The Solid Waste Program is currently being handled by several different Ministries. Coverage is high for collecting and disposing of solid waste. OK OK OK

Desechos Sólidos - El Programa de Desechos Sólidos actualmente está siendo manejado por varios Ministerios diferentes. La cobertura es alta para recolección y eliminación de desechos sólidos.

Food Protection - Food protection is a long-standing priority of the Ministry of Health with a special section devoted to this problem. A new program is now being organized to correct current problems which includes training new people and expanding into new fields such as coverage of food processing plants.[1] OK OK PP

Protección Alimentaria - La protección alimentaria es una prioridad de larga data del Ministerio de Salud con una sección especial dedicada a este problema. Un programa nuevo ahora se está organizando a problemas de corriente correcta que incluye personas nuevas de adiestramiento y en expansión en campos nuevos como cobertura de plantas de procesamiento de alimentos.[1]

Physical and Chemical Pollution - A number of major agricultural development plants have been announced for the rural areas. It is not known which health effects they may have and the health sector is not actively participating in them.[1] OK OK NO TU

Contaminación Física y Química - Varias plantas principales agrícolas de desarrollo se han anunciado para las áreas rurales. No es sabido que la salud efectúa pueden tener y el sector salud no activamente está participando en ellos.[1]

Veterinary Public Health - OK

Salud Pública Veterinaria -

The major zoonoses are salmonellosis and leptospirosis. OK NO
Present also are brucellosis and rabies.[1] Since the OK
exact magnitude in distribution is not well known, the
Veterinary Public Health Program is developing the
necessary epidemiological skills to implement the total
program. An excellent Veterinary Public Health Laboratory OK
with adequate equipment can then support the necessary
studies. A Division of Environmental Health is being OK
created within the Ministry of Health to coordinate this
program.

2.3 Disease Prevention and Control OK

Diseases Preventable by Vaccination - Since 1982, a major OK TU OK
initiative was taken to increase coverage at age 1 against
five diseases: polio, diphtheria, pertussis, tetanus, and OK
measles. Tetanus has increased dramatically from less than OK
30% to over 80% in 1983. The Immunization Program is being OK
expanded to include other high risk groups and other
diseases such as rubella and yellow fever.

Other Infectious and Parasitic Diseases of National OK
Significance - In 1984, a diarrheal disease control program OK
was launched. Diarrheal disease accounts for approximately OK
15% of infant mortality and 20% of mortality in 1-4 year
olds. Malaria is a problem in the interior of the country. OK
Operations have been improved in coverage and efficiency PP TU
by transferring the control activities to the Medical
Mission, a primary health care agency. Aedes aegypti is OK
abundant and several dengue outbreaks have occurred in the
past 5 years. Leprosy has a high incidence and prevalence OK
rate. Complication rates have been coming down, however, PP TU
over the last years because of early case finding and
treatment through the Dermatological Service.

13.6 DIARRHEAL DISEASES OK

Objectives: OK

Las zoonosis principales son salmonelosis y leptospirosis.
Presente también son brucelosis y rabia.[1] Ya que la
magnitud exacta en distribución no es bien conocida, el
Programa de Salud Pública Veterinaria está desarrollando
las especializaciones epidemiológicas necesarias para
instrumentar el programa total. Un Laboratorio excelente
de Salud Pública Veterinaria con equipo adecuado entonces
puede apoyar los estudios necesarios. Una División de
Salud Ambiental se está creando dentro del Ministerio de
Salud para coordinar este programa.

2.3 Prevención y Control de Enfermedades

Enfermedades Prevenibles mediante Vacunación - Desde
1982, una iniciativa principal se tomó para aumentar
cobertura a edad 1 contra cinco enfermedades: la polio, la
difteria, la pertussis, el tétanos, y sarampión. El
tétanos ha aumentado singularmente de menos de 30% a más de
80% en 1983. El Programa de Inmunización se está ampliando
para incluir a otros grupos de alto riesgo y otras
enfermedades como rubéola y fiebre amarilla.

Otras Enfermedades Infecciosas y Parasitarias de
Significado Nacional - En 1984, se lanzó un programa de
control de enfermedad diarreica. La enfermedad diarreica
representa aproximadamente 15% de mortalidad infantil y un
20% de mortalidad en 1-4 años de edad. La malaria es un
problema en el interior del país. Las operaciones se han
mejorado en cobertura y eficiencia mediante transferencia
de las actividades de control a la Misión Médica, una
agencia de atención primaria de salud. Aedes aegypti es
abundante y varios brotes de dengue han ocurrido en los
últimos 5 años. La lepra tiene una incidencia y tasa de
prevalencia alta. Las tasas de complicación han estado
bajando, sin embargo, en los últimos años por causa de
detección de casos y tratamiento prematuro a través del
Servicio Dermatológico.

13.6 ENFERMEDADES DIARREICAS

Objetivos:

1. The objectives of the Regional Program for the Control of Diarrheal Diseases (*CDD) are aimed at reducing: morbidity and mortality from acute diarrheal diseases as well as reducing the diarrhea associated with ill effects, particularly malnutrition in infants and young children. SD OK TU PP

2. The program seeks the promotion of national self-reliance in the delivery of health and the supply of oral rehydration salts (ORS) for the control of diarrheal diseases, through the collaboration with national authorities in charge of *CDD activities. SD OK

Targets: OK

3. Planning and Operations OK

a. by 1987 90% of the countries will have established operational control programs; OK

b. adequate supplies of oral rehydration salts (ORS) will be assured to all countries through the establishment of 15 subregional and national production centers; OK

c. oral rehydration therapy (ORT) will be established in 80% of PHC services in the countries, and access to ORT will be assured for at least 40% of childhood diarrhea cases in all countries; OK

d. *CDD control measures, other than ORT, will be implemented in at least 50% of the countries. SD OK

5. Evaluation OK

It is expected that by 1987: NO

a. 90% of the countries will have included ongoing monitoring and evaluation as integral parts of their country programs; OK

b. 90% of the countries will have had at least a periodic comprehensive review for their national programs. OK

c. 80% will be periodically reporting morbidity and mortality data to assess the impact of *CDD strategies. SD OK

Situation Analysis: OK

Research OK

1. Los objetivos del Programa Regional para el Control de Enfermedades Diarreicas (CDD) se dirigen a reducción: la morbilidad y mortalidad de enfermedades diarreicas agudas así como reducción de la diarrea asociada con efectos adversos, particularmente +la malnutrición+ en lactantes y de preescolares.

2. El programa busca la promoción de auto-suficiencia nacional en la entrega de salud y el suministro de sales de rehidratación oral (SRO) para el control de enfermedades diarreicas, a través de la colaboración con autoridades nacionales a cargo de actividades de CDD.

Metas:

3. Planificación y Operaciones

a. para 1987 un 90% de los países habrá establecido programas de control operativos;

b. suministros adecuados de sales de rehidratación oral (SRO) se asegurarán a todos países a través del establecimiento de 15 centros de producción subregionales y nacionales;

c. la terapia de rehidratación oral (TRO) se establecerá en un 80% de servicios de APS en los países, y el acceso a TRO se asegurará para por lo menos 40% de casos de diarrea de niñez en todos países;

d. Medidas de control de CDD, con excepción de TRO, se instrumentarán en por lo menos 50% de los países.

5. Evaluación

Se espera que para 1987:

a. un 90% de los países habrá incluido monitoría y evaluación continua como partes integrales de sus programas de país;

b. un 90% de los países habrá tenido por lo menos una revisión integral periódica para sus programas nacionales.

c. un 80% estará informando periódicamente datos de morbilidad y de mortalidad para evaluar la repercusión de estrategias de CDD.

Análisis de situaciones:

Investigaciones

6. It is expected that by 1987 50% of the countries will SD OK
be conducting operational research projects to help support
*CDD activities.

7. At the end of 1984, the ORT is being actively promoted OK
in all countries, although in seven countries there are not
well developed plans of action. The administration OK TU TU
training course in Portuguese language was held for the
first time in Brazil to train health personnel as a measure
to replicate the national program in all the states. Nine OK
physicians from the African Portuguese-speaking countries
participated. A similar course with the same purpose was OK TU
held in Mexico. The course on supervisory skills was OK TU
conducted in four countries for the first time; and two OK TU
countries conducted this course for the second time. OK
Technical training activities on clinical handling of the
child with diarrhea are routinely carried out by the
national authorities of each country; the Regional Program OK TU TU
contributes to these activities by providing technical
instructional materials and in some occasions by the
participation of an international expert.

8. Comprehensive program reviews were conducted in two OK
countries; two countries conducted diarrheal disease OK
morbidity and mortality surveys; and two more countries OK
made plans to carry out a similar survey at the beginning
of 1985.

9. In the area of operational research support was OK
provided to nine projects in seven countries.

Proposed Program Activities: OK

10. The Regional Program will collaborate with the Member OK
Countries in the areas mentioned below.

11. Planning and Operation OK

- Development of country plans of operation within the OK
context of overall country health plans by:

a. provision of written guidelines; OK

b. provision of consultants. OK

6. Se espera que para 1987 un 50% de los países estará realizando proyectos de investigaciones operativas para ayudar a apoyar actividades de CDD.

7. Al final de 1984, la TRO se está activamente promoviendo en todos países, aunque en siete países no hay bien planes desarrollados de acción. El curso de adiestramiento de administración en idioma portugués se llevó a cabo por la primera vez en Brasil para adiestrar a personal de salud como una medida para reproducir el programa nacional en todos los estados. Participaron nueve médicos de los países de habla portuguesa africanos. Un curso similar con el mismo fin se llevó a cabo en México. El curso sobre especializaciones supervisoras se realizó en cuatro países por la primera vez; y dos países realizaron este curso por la segunda vez. Actividades de adiestramiento técnico en manejo clínico del niño con diarrea de rutina son llevadas a cabo por las autoridades nacionales de cada país; el Programa Regional contribuye a estas actividades mediante provisión de materiales instruccionales técnicos y en algunas ocasiones mediante la participación de un experto internacional.

8. Revisiones integrales de programa se realizaron en dos países; dos países realizaron encuestas diarreicas de enfermedad de morbilidad y de mortalidad; y dos más países hicieron planes para llevar a cabo una encuesta similar al comienzo de 1985.

9. En el área de investigaciones operativas apoyo se proporcionó a nueve proyectos en siete países.

Actividades propuestas de Programa:

10. El Programa Regional colaborará con los Países Miembros en las áreas mencionadas abajo.

11. Planificación y Operación

- Desarrollo de planes de país de operación dentro del contexto de salud giobal de país planifica por:

a. la provisión de normas escritas;

b. la provisión de consultores.

- Development of improved communications support, including promotional and educational materials to support control strategies, especially those directed towards mothers	PP	- Desarrollo de apoyo mejorado de comunicaciones, incluyendo materiales promocionales y educativos para apoyar estrategias de control, especialmente aquellos dirigieron hacia madres	
- Improvement in control strategies through development of cost-effectiveness guidelines for both morbidity and mortality reduction	OK	- Mejora en estrategias de control a través de desarrollo de normas de eficacia en función de los costos para reducción de morbilidad y de mortalidad	
- Development of complementary programs with other primary health care (PHC) elements, particularly immunization, nutrition, family health, and environmental health, within the framework of PHC	PP	- Desarrollo de programas complementarios con otros elementos de atención primaria de salud (APS), particularmente la inmunización, nutrición, salud familiar, y salud ambiental, dentro del marco de APS	
- Ensuring, on a country, regional, and global basis, adequate supplies of oral rehydration salts (ORS) through:	PP	- Asegurar, en una base de país, regional, y global, suministros adecuados de rehidratación oral sala (SRO) a través de:	
a. identification of country needs and available sources of bilateral or multilateral support;	OK	a. identificación de necesidades de país y fuentes disponibles de apoyo bilateral o multilateral;	
b. support for the development of subregional and national ORS production centers;	OK	b. apoyo para el desarrollo de centros de producción de SRO subregionales y nacionales;	80
c. development of alternative methods of preparation of ORS.	OK	c. desarrollo de métodos alternativos de preparación de SRO.	
- Collaboration in the control of epidemics	OK	- Colaboración en el control de epidemias	
12. Training	OK	12. Adiestramiento	
- Management training:	OK	- El adiestramiento gerencial:	
a. a course for national program administrators on determining overall communicable disease priorities, setting objectives and targets, selection of delivery and logistics systems, and evaluation;	PP	a. un curso para administradores nacionales de programa al determinar prioridades globales de enfermedad transmisible, fijando objetivos y metas, selección de entrega y sistemas logísticos, y evaluación;	
b. a supervisor's course on community involvement, setting local targets, treatment of diarrhea, training, and monitoring and evaluation.	PP	b. el curso de un supervisor sobre participación comunitaria, ambiente metas locales, tratamiento de diarrea, adiestramiento, y monitoría y evaluación.	
- Technical training:	OK	- El adiestramiento técnico:	
a. strengthening of national training centers;	OK	a. el fortalecimiento de centros nacionales de adiestramiento;	
b. providing mechanisms for TCDC, especially in ORS supplies and technical training.	OK	b. provisión de mecanismos para CTPD, especialmente en suministros de SRO y adiestramiento técnico.	

- Inclusion of *CDD material in curricula of medical and nursing schools, schools of public health and other health worker training institutions	SD OK	- Inclusión de material CDD en programas de escuelas médicas y de enferme ía, escuelas de salud pública y otras instituciones de adiestramiento de obreros de salud
- Development of locally appropriate training and health education materials on prevention and treatment strategies	OK	- Desarrollo de adiestramiento y materiales de educación de salud localmente apropiados en prevención y estrategias de tratamiento
13. Evaluation	OK	13. Evaluación
- Use of ongoing information systems to meet specific country needs for monitoring *CDD program progress:	SD OK TU	- Uso de sistemas de información continua para satisfacer necesidades específicas de país de monitoría de progreso de programa de CDD:
a. review and analysis of specific data to be collected on *CDD *morbi/mortality, accessibility and coverage of ORT to determine impact on *CDD strategies at the country level;	SD SD SD NO	a. la revisión y análisis de datos específicos a recoger en morbi/mortalidad de CDD, accesibilidad y cobertura de TRO para determinar repercusión en estrategias de CDD al nivel de país;
b. strengthen existing national *DD surveillance systems.	SD OK	b. reforzar sistemas existentes nacionales de vigilancia de DD.
- Improvement and development of regional and global information systems integrated with other priority information needs to measure achievement of targets and objectives	OK	- Mejora y desarrollo de sistemas de información regionales y globales integrados con otras necesidades de información de prioridad para medir logro de metas y objetivos
- Use of comprehensive reviews to provide periodic formal assessment of program accomplishments and constraints	OK	- Uso de revisiones integrales para proporcionar evaluación formal periódica de logros de programa y limitaciones
14. Operational Research	OK	14. Investigaciones Operativas
- Oral rehydration therapy:	OK	- La terapia de rehidratación oral:
a. study of appropriate methods for its delivery to the community and family within the context of primary health care, including studies on acceptability, effectiveness of various health education methods for teaching ORT, and effectiveness of simple methods to assure the early home treatment of diarrhea.	PP	a. el estudio de métodos apropiados para su entrega a la comunidad y familia dentro del contexto de atención primaria de salud, incluyendo estudios en aceptabilidad, eficacia de diversos métodos de educación de salud para enseñanza de la TRO, y de eficacia de métodos simples para asegurar el tratamiento prematuro en el domilicio de diarrea.
- Control measures, other than ORT:	OK	- Medidas de control, con excepción de TRO:

a. studies of the effect of breast-feeding, personal and OK
domestic hygiene on the incidence and severity of diarrheal
diseases, and traditional beliefs and practices concerning
these interventions;

b. study of methods of achieving accurate, representative OK
reporting of diarrheal diseases.

13.7 ACUTE RESPIRATORY INFECTIONS OK

Objective: OK

1. To reduce the mortality from acute respiratory OK
infections (ARI) in children in developing countries
through effective case management at primary health care
and referral levels and health education. In the long OK
term, the aim is to prevent severe respiratory infections.

Targets: OK

2. By 1987, a prototype national program for intervention NO
will have been developed, adapted to national
characteristics and applied in 33% of the countries in the
Region, as an integrated part of child care within primary
health care.[5]

Proposed Program Activities: OK

3. National Program Development OK

Advice on norms and on organization of program activities OK TU
will be provided by permanent staff and consultants. Close OK
coordination with other child health programs and with
UNICEF will facilitate an integrated approach and rational
use of the limited funds for this program. Priorities for OK
implementation are for the countries in continental Middle
America and the Andean Region.

4. Training and educational materials will be developed OK
and adapted to the needs of each country.

a. estudios del efecto de lactancia, higiene personal y
doméstica en la incidencia y gravedad de enfermedades
diarreicas, y creencias y prácticas tradicionales en cuanto
a estas intervenciones;

b. el estudio de métodos de logro de notificación exacta,
representativa de enfermedades diarreicas.

13.7 INFECCIONES RESPIRATORIAS AGUDAS

Objetivo:

1. A reducir la mortalidad de infecciones respiratorias
agudas (IRA) en niños en países en vías de desarrollo a
través de gerencia efectiva de casos a atención primaria de
salud y niveles de referencia y salud la educación. A la
larga, el objetivo es prevenir infecciones graves
respiratorias.

Metas:

2. Para 1987, un programa nacional de prototipo para
intervención se habrá desarrollado, adaptado a
características nacionales y aplicado en 33% de los países
en la Región, como una parte integrada de atención infantil
dentro de atención primaria de salud.[5]

Actividades propuestas de Programa:

3. Desarrollo Nacional de Programas

El consejo sobre normas y en organización de actividades
de programa será proporcionado por personal y consultores
permanentes. La coordinación cercana con otra salud
infantil programa y con UNICEF facilitará un enfoque y uso
racional integrado de los fondos limitados para este
programa. Las prioridades para instrumentación son para
los países en Mesoamérica continental y la Región andina.

4. El adiestramiento y materiales educativos se
desarrollarán y adaptarán a las necesidades de cada país.

5. The implementation of control activities to reduce severe disease and mortality will be accompanied by a system of evaluation of operation and impact, both through regular reporting (mortality, coverage of health services and volume of cases managed) and clinical results by severity, through special operational studies or sentinel health units. PP

Manpower Development OK

6. At the managerial and supervisory levels, national courses will utilize the basic diarrheal disease modules for programming, plus a technical module on ARI. Whenever possible, the training will be combined with other programs of prevention of child mortality. OK OK

7. At peripheral and community health levels, the countries will develop their own training system with emphasis on training auxiliary personnel as well as during their supervision. OK

Technology Transfer OK

8. The development and reinforcement of technology for ARI in the laboratories of developing countries will receive special attention. National and subregional courses and seminars and coordination with reference centers in developed countries will be the main strategies. OK OK

9. In clinical and management aspects, model research protocols will be prepared to be applied in studies in selected centers of developing countries. OK

13.8 TUBERCULOSIS OK

Objective: OK

1. To cooperate with all Member Countries of the Region to reduce suffering, disability and death from tuberculosis (TB), and to eliminate the disease as a public health problem by strengthening diagnostic, treatment and vaccination services. PP

Targets: OK

2. By 1987: NO

5. La instrumentación de actividades de control para reducir enfermedad y mortalidad grave será acompañado por un sistema de evaluación de operación y repercusión, ambos a través de regular informando (mortalidad, cobertura de servicios de salud y volumen de casos administrados) y resultados clínicos por gravedad, a través de estudios operativos o unidades de salud de centinela especiales.

Desarrollo de Recursos Humanos

6. A los niveles gerenciales y supervisores, cursos nacionales utilizarán los módulos básicos de enfermedad diarreica para programación, más un módulo técnico en IRA. Cuando posible, el adiestramiento se combinará con otros programas de prevención de mortalidad infantil.

7. A niveles periféricos y de salud comunitaria, los países desarrollarán su propio sistema de adiestramiento con énfasis en adiestramiento de personal auxiliar así como durante su supervisión.

Transferencia de Tecnologías

8. El desarrollo y potencialización de tecnología para IRA en los laboratorios de países en vías de desarrollo recibirán atención especial. Los cursos nacionales y subregionales y seminarios y coordinación con centros de referencia en países desarrollados serán las estrategias principales.

9. En aspectos clínicos y de gerencia, protocolos de investigaciones de modelo se prepararán para aplicarse en estudios en centros seleccionados de países en vías de desarrollo.

13.8 TUBERCULOSIS

Objetivo:

1. A cooperar con todos Países Miembros de la Región para reducir sufrimiento, incapacidad y muerte de tuberculosis (TBC), y a eliminar la enfermedad como un problema de salud pública al reforzar de diagnóstico, tratamiento y servicios de vacunación.

Metas:

2. Para 1987:

- All countries in the Region will have case finding and notification systems with full coverage of health services to evaluate the impact of the program through decreases in incidence.	OK	- Todos países en la Región tendrán detección de casos y sistemas de notificación con cobertura total de servicios de salud para evaluar la repercusión del programa a través de descensos en incidencia.
- At least 60% of the countries with 80% of the population will show a decrease in the estimated risk of infection or incidence in bacillary cases larger than 5% per year.	PP	- Por lo menos 60% de los países con un 80% de la población mostrará un descenso en el riesgo estimado de infección o incidencia en casos bacilares más grandes que 5% por año.
- An average reduction of over 5% per year will be achieved for TB mortality.	OK	- Una reducción media de más de 5% por año se logrará para mortalidad de TBC.
Situation Analysis:	OK	Análisis de situaciones:
3. All countries have a policy of integration of activities and a national program with a responsible officer at the Ministry of Health.	OK	3. Todos países tienen una política de integración de actividades y un programa nacional con un oficial responsable en el Ministerio de Salud.
4. Around 250,000 new cases are notified per year (only 9% in North America). The problem is being reduced at nearly 5% per year, but the number of cases notified increases yearly because of better case finding, registration, availability and coverage of diagnostic facilities and an increase in population in developing countries.	OK OK	4. Alrededor de 250,000 casos nuevos se notifican por año (sólo 9% en América del Norte). El problema se está reduciendo a casi 5% por año, pero el número de casos notificó aumentos anualmentemente por causa de detección de casos mejor, registro, disponibilidad y cobertura de instalaciones de diagnóstico y un aumento en población en países en vías de desarrollo.
5. The main problems are low accessibility of the population to free of charge diagnosis and treatment, reduced funds from the Governments for provision of drugs and insufficient supervision.	OK	5. Los problemas principales son accesibilidad baja de la población a diagnóstico y tratamiento gratuito, fondos reducidos de los Gobiernos para provisión de medicamentos y supervisión insuficiente.
Proposed Program Activities:	OK	Actividades propuestas de Programa:
6. Program planning and implementation:	OK	6. La planificación e instrumentación de programas:
a. Technical advice on norms and procedures and cooperation on evaluation of results and epidemiological impact;	OK TU	a. El consejo técnico sobre normas y procedimientos y cooperación en evaluación de resultados y repercusión epidemiológica;
b. Support for training through fellowships and national and international courses on program management and bacteriology of tuberculosis;	OK	b. Apoyo para adiestramiento a través de becas y cursos nacionales e internacionales sobre gerencia de programas y bacteriología de tuberculosis;
c. Assistance in the procurement of drugs and materials;	OK	c. La asistencia en las adquisiciones de medicamentos y materiales;
d. Provision of technical information for program officers;	OK	d. La provisión de información técnica para oficiales de programa;

e. Evaluation of national and regional trends. OK

Research OK

7. Research activities will be promoted and coordinated OK
in cooperation with the WHO Collaborating Centers, with the
following priorities: epidemiological analysis of trends; OK OK
operational results of treatment; quality of bacteriology OK
for case finding; prevalence of bacteriological resistance OK
to drugs; and efficacy of BCG vaccination. OK

PROGRAM ON MICROCOMPUTER APPLICATIONS FOR DEVELOPING OK
COUNTRIES

The Board on Science and Technology for International OK
Development (BOSTID) has received a grant from the U.S.
Agency for International Development to conduct a three-
year study of microcomputer applications useful for
developing countries. A panel of scientists knowledgeable SD PP
on the use of microcomputers in various sectors, chaired by
Dr. William J. Lawless, Jr. of *Cognitronics Corporation,
has been appointed to oversee the program. At its first OK
meeting, the overview panel defined the scope of the
individual seminars to be held under this program. The PP
first, held in Sri Lanka, November 4-9, 1984, considered
applications in the areas of health, agriculture, and
energy. Planning for a second meeting, to be held in OK
Mexico in November 1985, is under way. The second seminar OK
will consider applications in education and training. OK
Future meetings will concentrate on expert systems and
robotics and the policy implications of microcomputer
technology.

The increasing use of microcomputers is recognized as a OK
revolution at least as significant as those in agriculture
and industry that preceded it. Stanford anthropologist PP
Robert Textor noted a major difference, however: the pace OK
of the information revolution is so rapid that people are
aware that their lives are being changed.

e. Evaluación de tendencias nacionales y regionales.

Investigaciones

7. Actividades de investigaciones se promoverán y
coordinarán en cooperación con los Centros Colaboradores
*de la OMS, con las prioridades siguientes: análisis
epidemiológico de tendencias; resultados operativos de
tratamiento; la calidad de bacteriología para detección de
casos; la prevalencia de resistencia bacteriológica a
medicamentos; y la eficacia de vacunación de BCG.

PROGRAMA SOBRE APLICACIONES DE MICROCOMPUTADOR PARA PAISES
EN VIAS DE DESARROLLO

La Junta en Ciencia y Tecnología para Desarrollo
Internacional (BOSTID) ha recibido una subvención de la
Agencia de EE.UU. para Desarrollo Internacional para
realizar un estudio de tres años de aplicaciones de
microcomputador útiles para países en vías de desarrollo.
Un panel de científicos versados en el uso de
microcomputadores en diversos sectores, presididos por Dr.
William J. Lawless, Jr. de Cognitronics Corporation, se ha
nombrado para supervisar el programa. En su primera
reunión, el panel de revisión definió el alcance de los
seminarios individuales a celebrar debajo este programa.
El primer, celebrado en Sri Lanka, 4-9 de noviembre de
1984, aplicaciones consideradas en las áreas de salud,
agricultura, y energía. La planificación para una segunda
reunión, a celebrarse en México en noviembre de 1985, está
en curso. El segundo seminario considerará aplicaciones en
educación y adiestramiento. Reuniones futuras concentrarán
en sistemas expertos y robótica y las consecuencias
políticas de tecnología de microcomputador.

El uso creciente de microcomputadores se reconoce como
una revolución por lo menos tan significativo como aquellos
en agricultura e industria que lo precedió. El antropólogo
de Stanford Robert Textor notó una diferencia principal,
sin embargo: el paso de la revolución de información es
tan rápido que las personas son conscientes que se están
cambiando sus vidas.

"This very awareness is an immensely valuable fact of life," he added, "for it means that people have a cognitive base upon which to conceive and design action aimed at realizing some of the benefits that the Revolution has to offer, and at fending off some of its destructiveness."

Sri Lanka Symposium

The first symposium in the BOSTID Microcomputers for Developing Countries program was concerned with the three sectors of agriculture, health, and energy. In addition to providing information on the ways in which microcomputers are currently being used, the symposium highlighted the policy issues that need to be considered to integrate microcomputer technology into these sectors. The symposium, held near Colombo, Sri Lanka, November 4-9, 1984, was sponsored jointly by the Computer and Information Technology Council of Sri Lanka (*CINTEC) and the BOSTID of the U.S. National Research Council.

Participants included microcomputer users from 16 countries, from research organizations, universities, government ministries, and the private sector. After the first day's plenary session, the participants divided into sectoral workshops to exchange information on specific topics. Each sectoral meeting room was provided with microcomputers for demonstrations, and participants shared their experiences.

In the health sector, for example, papers and demonstrations were presented describing database systems, statistical analysis and evaluation, and models for prediction and resource allocation. *Donald Lauria of the University of North Carolina's School of Public Health, the sectoral chairman, provided a literature search of programs in the health field.

"Esta misma conciencia es un hecho inmensamente valioso de vida," agregó, "para ello medios que personas tienen una base cognoscitiva en el cuál a concebir y diseñar acción orientado a comprender alguno de los beneficios que la Revolución tiene que ofrecer, y a contener alguno de su destructividad."

Simposio de Sri Lanka

El primer simposio en los Microcomputadores de BOSTID para programa de Países En Vías de Desarrollo fue preocupado con los tres sectores de agricultura, salud, y energía. Además de provisión de información sobre las maneras en el cuál los microcomputadores están actualmente empleándose, el simposio destacó los temas políticos que necesitan considerarse para integrar tecnología de microcomputador en estos sectores. El simposio, celebrado cerca de Colombo, Sri Lanka, 4-9 de noviembre de 1984, se patrocinó conjuntamente por el Computador y Consejo de Tecnología de Información de Sri Lanka (CINTEC) y el BOSTID del Consejo de Investigaciones Nacional de EE.UU..

Participantes incluyeron a usuarios de microcomputador de 16 países, de organizaciones de investigaciones, universidades, ministerios del gobierno, y el sector privado. Después de la sesión plenaria del primer día, los participantes dividido en talleres sectoriales para intercambiar información sobre temas específicos. Cada sala de reunión sectorial se proporcionó con microcomputadores para demostraciones, y participantes compartieron sus experiencias.

En el sector salud, por ejemplo, se presentaron documentos y demostraciones describiendo sistemas de banco de datos, análisis estadístico y evaluación, y modelos para predicción y asignación de recursos. Donald Lauria de la Universidad de North Carolina School de Salud Pública, el presidente sectorial, proporcionó una búsqueda de literatura de programas en el campo de salud.

The majority of the uses found in the literature as well as OK
at the symposium were database systems with some
statistical analysis. For example, rural clinics in OK
Thailand and Malaysia are linked with central ministries
through a microcomputer database system allowing analysis
at various levels of aggregation. In Sri Lanka, reports OK
were presented describing databases used to organize
clinical and surgical caseloads. The individual records in OK
the latter set were analyzed after surgery to determine the
accuracy of prior diagnoses. Models are used primarily for OK
large scale planning in health-related areas such as water
supply and wastewater disposal.

Papers presented in the energy sector used microcomputers OK
for a wide range of approaches in database management,
modeling, and analysis. In Indonesia, for example, a OK
commercial spreadsheet program is used as a flexible tool
to do a quick analysis of energy demand projections. From PP
the Sudan, a complex simulation program was presented which
facilitates energy planning at the national level.[1] OK
complejos
Microcomputers were also used for subsectoral management
tasks such as energy conservation investment analysis in
Sri Lanka and data acquisition from pilot plant facilities
in India It was in the energy sector that perhaps the NO
most sophisticated uses of the microcomputer were found.[1]

A number of policy recommendations for developing country OK
governments and for donor agencies were drawn out of the
deliberations of the symposium. The importance of training OK
was brought out by many speakers.

La mayoría de los usos hallados en la literatura así como
en el simposio fue sistemas de banco e datos con algún
análisis estadístico. Por ejemplo, las clínicas rurales en
Tailandia y Malasia están vinculadas con ministerios
centrales a través de un sistema de banco de datos de
microcomputador que permite análisis a diversos niveles de
agregación. En Sri Lanka, se presentaron informes
describiendo bancos de datos empleados para organizar casos
clínicos y quirúrgicos. Los registros individuales en el
último conjunto se analizaron después de cirugía para
determinar la exactitud de diagnósticos previos. Modelos
se emplean principalmente para planificación en gran escala
en áreas relacionadas con salud como abastecimiento de agua
y desecho de aguas residuales.

Documentos presentados en el sector de energía emplearon
microcomputadores para una variedad amplia de enfoques en
gerencia de bancos de datos, modelado, y análisis. En
Indonesia, por ejemplo, un programa de hoja electrónica
comercial se emplea como una herramienta flexible para
hacer un análisis rápido de proyecciones de demanda de
energía. Del Sudán, un programa de simulación de

se presentó que facilita planificación de energía al nivel
nacional.[1] Microcomputadores también se emplearon para
tareas de gerencia subsectorial como análisis de
inversiones de conservación de energía en Sri Lanka y
adquisición de datos de instalaciones de planta piloto en
la India. Fue en el sector de energía que quizás los usos
más complejos del microcomputador se hallaron.[1]

Varias recomendaciones políticas para los gobiernos de
país en vías de desarrollo y para agencias donantes se
extrajeron de las deliberaciones del simposio. La
importancia de adiestramiento fue hecha resaltar por muchos
expositores.

Training needs include computer literacy for the public and PP TU
for policymakers, a fostering of awareness of possible
computer applications in various technical professionals in
fields such as health, energy, and agriculture, and
technical training at the professional level to create a
corps of computer professionals. It was proposed that a PP TU
study be conducted to provide computer manufacturers with a
recommended set of design features needed to enhance the
ability to use microcomputers under developing country
environmental conditions. To ensure that the microcomputer OK
package chosen is appropriate for its intended use, the
group suggested that donor agencies and outside experts
should confer with local decisionmakers concerning locally
recognized needs.

Las necesidades de adiestramiento incluyen familiarización con computadores para el público y para autoridades, un incentivo de conciencia de aplicaciones posibles de computación en diversos profesionales técnicos en campos como salud, energía, y agricultura, y adiestramiento técnico al nivel profesional para crear un cuerpo de profesionales de computación. Se propuso que un estudio se realizara para proporcionar a fabricantes de computador con un conjunto recomendado de rasgos de diseño necesarios para mejorar la capacidad para emplear microcomputadores bajo condiciones ambientales de país en vías de desarrollo. Para asegurar que el paquete de microcomputador elegido sea apropiado para su uso concebido, el grupo sugirió que agencias donantes y expertos externos deben conferir con autoridades locales en cuanto a necesidades localmente reconocidas.

It was unanimously agreed that government could best PP
foster the growth of computer literacy in all sectors by a
policy of encouraging the dissemination of microcomputers
rather than by trying to regulate them.[1] Individuals PP
should be permitted to acquire hardware and software that
they believe fulfills their needs, and government can
encourage this individual acquisition by a favorable import
tax policy.[2]

Unánimemente se acordó que el gobierno mejor podía incentivar el crecimiento de familiarización con computadores en todos sectores por una política de alentar la difusión de microcomputadores en lugar de al tratar de reglamentar ellos.[1] Los individuos se deben permitir para adquirir equipo y software que creen cumple sus necesidades, y el gobierno puede alentar esta adquisición individual por una política fiscal de importación favorable.[2]

Service and support are two other areas in which OK TU
government can encourage the spread of computer use. It OK
was recommended that each national government, with funding
from donor agencies, establish a computer center as a
resource for users. The center should offer advice and OK TU
assistance for the selection and installation of both
hardware and software and be available for consultation on
issues of standardization and compatibility without playing
a decisionmaking role. The center should also serve as a OK
clearinghouse for microcomputer applications developed in
other countries with similar needs.

El servicio y apoyo son dos otras áreas en las cuales el gobierno puede alentar la propagación de uso de computaciones. Se recomendó que cada gobierno nacional, con recursos financieros de agencias donantes, estableciera un centro de computación como un recurso para usuarios. El centro debe ofrecer consejo y asistencia para la selección e instalación de equipo y software y estar disponible para consulta en temas de normalización y compatibilidad sin desempeño de una función de toma de decisiones. El centro también debe servir como un centro de distribución de información para aplicaciones de microcomputador desarrolladas en otros países con necesidades similares.

Information and software available at the center would limit the amount of investment the individual microcomputer user would have to make. It was also suggested that user groups including members from developed and developing countries could be coordinated through such a center.

La información y software disponible en el centro limitarían la cantidad de inversión el usuario de microcomputador individual tendría que hacer. También se sugirió que grupos de usuario incluyendo miembros de países desarrollados y en vías de desarrollo se podían coordinar a través de dicho centro.

The Proceedings of the seminar including abstracts of the presentations and a full set of recommendations will be published in June 1985 by *CINTEC, Sri Lanka. A report based on the information presented in the sectoral discussions in agriculture, energy, and health, is scheduled for publication by BOSTID in Fall 1985.

Las Actas del seminario incluyendo resúmenes analíticos de las presentaciones y un conjunto total de recomendaciones se publicarán en junio de 1985 por CINTEC, Sri Lanka. Un informe basado en la información presentada en las discusiones sectoriales en agricultura, energía, y salud, se programa para publicación por BOSTID en *Otoño 1985.

Second Symposium

Segundo Simposio

The Symposium on Microcomputer Applications in Education and Training in Developing Countries, cosponsored by the National Academy of Engineering of Mexico, will be held in Cuernavaca in November 1985. The issues to be addressed include the preparation and availability of courseware, teacher training, the technologies that support and enhance microcomputers in education and training, and policy issues in implementing and managing their use. These issues will be discussed as they apply to (1) elementary, secondary, and literacy education; (2) professional and university education; and (3) non-formal, vocational, and continuing education.

El Simposio en Aplicaciones de Microcomputador en Educación y Adiestramiento en Países En Vías de Desarrollo, copatrocinados por la Academia Nacional de Ingeniería de México, se celebrará en Cuernavaca en noviembre de 1985. Los temas a abordarse incluyen la preparación y disponibilidad de material didáctico, adiestramiento de maestros, las tecnologías que apoyan y mejoran microcomputadores en educación y adiestramiento, y temas políticos en instrumentación y administrando su uso. Estos temas se tratarán según aplican a (1) educación primaria, secundaria, y de alfabetismo; (2) la educación profesional y de universidad; y (3) la educación, vocacional, y permanente no-formal.

Each participant, who should have actual experience in using microcomputers in one of these areas of education or training, will be asked to prepare a demonstration or a paper describing that experience. We request that each potential participant send biographical information and a 100-word abstract of the demonstration or paper, so that we can select those who will make the greatest contribution to a useful exchange of information at the symposium.

A cada participante, que debe tener experiencia concreta en empleo de microcomputadores en uno de estas áreas de educación o adiestramiento, se le pedirá que prepare una demostración o un documento que describe esa experiencia. Solicitamos que cada participante potencial envíe información biográfica y un resumen analítico de 100 palabras de la demostración o documento, para que podamos seleccionar los que hará la contribución mayor a un intercambio útil de información en el simposio.

By bringing together experienced individuals from Latin America and other areas, the symposium is planned to assist in the diffusion of practical information on this relatively new technology. Your willingness to cooperate in the work on this significant topic will contribute greatly to the value of the interaction among participants at the symposium.

OK TU

OK

Mediante reunión de individuos experimentados de América Latina y otras áreas, el simposio se planifica para asistir en la difusión de información práctica sobre esta tecnología relativamente nueva. Su voluntad para cooperar en el trabajo en este tema significativo contribuirá enormemente al valor de la interacción entre participantes en el simposio.

05/28/86
*HDR99S9999999 APPENDIX D TO REPORT

PARSER STATS

	OK	PP	NO	TOTAL
OF SENTENCES	268	34	8	310
% OF TOTAL SENTENCES	86	11	3	
AVERAGE LENGTH	14	30	19	16
COMPLETE SENTENCES	163			197
LONGEST SENTENCE	42			
AVERAGE WORDS PARSED		24		
CURRENT ALLOCATION LIMIT		40		
TOTAL WORDS + PUNC				5266
TOTAL AFTER UNIT LOOKUP				5065