UC Merced
Proceedings of the Annual Meeting of the Cognitive Science Society

Title
A Principle-Based Approach To Parsing for Machine Translation

Permalink
https://escholarship.org/uc/item/3vr9f13r

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 9(0)

Author
Dorr, Bonnie J.

Publication Date
1987

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

A Principle-Based Approach To Parsing for Machine Translation

Bonnie J. Dorr
M.I.T. Artificial Intelligence Laboratory
545 Technology Square, Room 810
Cambridge, MA 02139, USA
(617) 253-7836
BONNIE@MIT-PREP.AI.MIT.EDU

Session: Paper
Keywords: Natural Language, Parsing, Principles vs. Rules, Interlingual Translation

Abstract

Many parsing strategies for machine translation systems are based entirely on context-free grammars; to try to capture all natural language phenomena, these systems require an overwhelming number of rules; thus, a translation system either has limited linguistic coverage, or poor performance (due to formidable grammar size). This paper shows how a principle-based "co-routine design" implementation improves the parsing problem for translation. The parser consists of a skeletal structure-building mechanism that operates in conjunction with a linguistically based constraint module, passing control back and forth until underspecified skeletal phrase structure is converted into a fully instantiated parse tree. The modularity of the parsing design accommodates linguistic generalization, reduces the grammar size, enables extendibility, and is compatible with studies of human language processing.[1]

1 Introduction

The problem addressed in this paper is to construct a parsing model that accommodates cross-linguistic uniform machine translation without relying on language-specific context-free rules. Typically parsing systems use grammars that describe language via complicated rules that spell out the details of their application. For example, ATN-based systems (Woods, 1970; Bates, 1978) have several hundred grammar arcs, each with detailed tests and actions; augmented phrase-structure grammars as in Diagram (Robinson, 1982) spell out the type, position, and probability of occurrence of constituents in a given phrase; and the GPSG approach (Gazdar, *et. al.*, 1985) uses a "slash-category" mechanism to incorporate long distance relations directly into the grammar rules.[2] Such systems do not work in the context of translation across several languages: the rules of a given grammar are painstakingly tailored to describe a *single* language, thus forcing a loss of linguistic generalization, limiting the addition of new languages, and inducing inefficiency (due to formidable grammar size).[3]

[1] Frazier 1986 provides recent psycholinguistic evidence that there is a temporal sequence of parsing consistent with the GB-based model presented here. This will be mentioned briefly in section 1, but is not the central focus of this paper.

[2] Barton (1984) describes these rule-based systems in more detail.

[3] For example, Slocum's METAL system (1984, 1985) developed at the Linguistics Research Center at the University of Texas relies on numerous language-specific context-free rules per language solely for parsing. The type of grammar formalism is allowed to vary from language to language. For example, the German

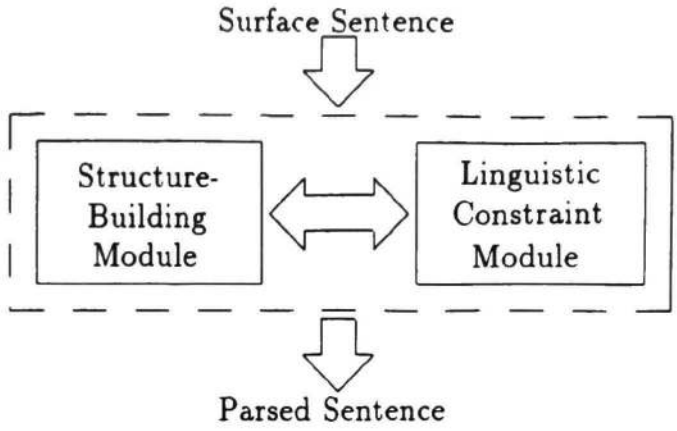

Figure 1: Co-Routine Design of the Parser

Furthermore, these systems fail to preserve the modular organization of new theories of grammar.

In this paper I describe an implementation of a parsing model that is based on subsystems of grammatical principles and parameters.[4] The parser follows a "co-routine design:" the structure-building mechanism operates with simultaneous access to linguistic constraints of Government and Binding (GB) theory as developed by Chomsky (1981, 1982). (See figure 1.) The structure-building module assigns a skeletal syntactic structure to a sentence, and then this structure is eliminated or modified according to the principles of GB. This design is consistent with recent psycholinguistic studies (see Frazier, 1986) that indicate that the human processor initially assigns a (potentially ambiguous or underspecified) structural analysis to a sentence, leaving lexical and semantic descriptions for subsequent processing. Furthermore, the parser is designed so that it applies uniformly across all languages, allowing the user to modify the parameters of the system to accommodate additional additional languages.

The reason that parsing uniformly across languages is difficult is the parser appears to require a massive amount of "knowledge" in order to parse all possible phenomena (and their interaction effects) in any given language without allowing ill-formed sentences to also be parsed. Consider (1):

(1) Le quiere a Juan
'(She) loves John'

Although (1) appears to be simple, it is not simple from the point of view of uniform parsing since the equivalent sentence parses differently in other languages. The Spanish and

parser is based on phrase-structure grammar, augmented by procedures for transformations; by contrast, the English parser employs a modified GPSG approach with no transformations. Regardless of the type of grammar formalism, each parser is nevertheless based on hundreds context-free rules of a language-specific nature. Consequently, each parser operates unilingually and has an increased running time over parsers that access smaller grammars. (As noted in Barton (1984), the Earley algorithm (1970) for context-free language parsing can quadruple its running time when the grammar size is doubled.)

[4]For example, there is a "constituent order" parameter associated with a universal principle that requires there to be a language-dependent ordering of constituents with respect to a phrase; the parameter is set by the user to be *head-initial* for a language like English, but *head-final* for a language like Japanese. This is discussed in section 2.1.

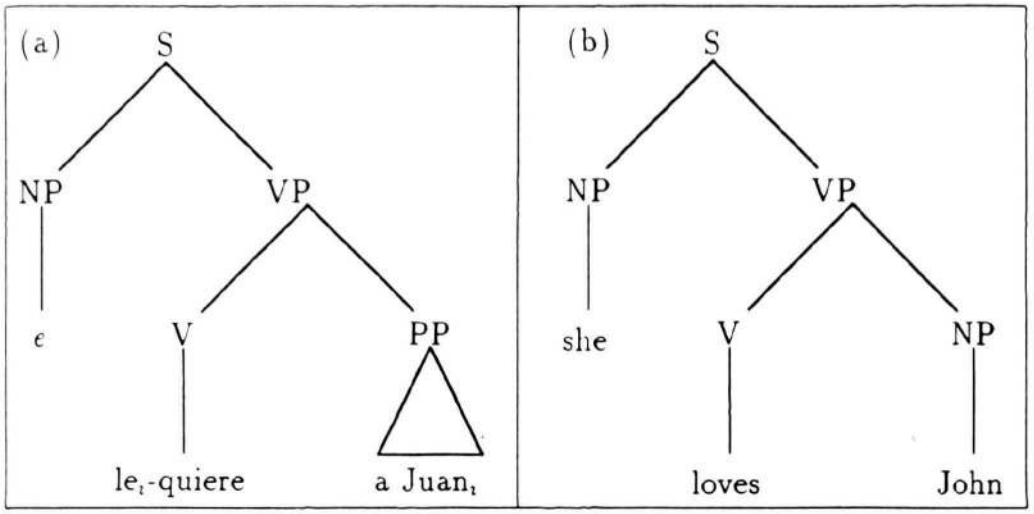

Figure 2: Spanish and English Parse Trees for an Equivalent Sentence

English parse trees for (1) are in figure 2.[5] Literally, the English translation for (1) is (2), which is ungrammatical:

(2) him *e* loves to John

The *e* stands for a null subject that is realized as *she* in English.[6] The parsing implementation presented here rules out sentence (2) without sacrificing the ability to parse (1).

The co-routine design differs from other GB parsing/translation systems (*e.g.*, Sharp, 1985) in that the GB principles are used for "on line" verification during parsing rather than as well-formedness conditions on output. Furthermore, in Sharp's system, context-free rules (set up for English-like languages) are hardwired into the code rather than generated on the fly using principles of GB; thus, languages (like German or Japanese) that do not have the same order of constituents as English cannot be handled by the system. The primary factor that introduces this malady in Sharp's system is that the user has limited access to the principles of the system. The system described here allows the user to specify parameter values to the principles, thus modifying the effect of the principles from language to language. There are two classes of GB principles used by the system: those that are applied on line (*i.e.*, at processing time) and those that are applied off line (*i.e.*, at precompilation time).[7]

[5]Subscripts are used for co-referring elements. Thus *le* (= him) refers to *Juan*.

[6]Section 2.3 discusses the null subject phenomenon in Spanish.

[7]Experiments are currently underway to determine the "optimal" balance of principle clustering between the precompilation and processing phases. In order for the GB constraints to be applicable, a structure must first be created. The question under investigation is how much structure must be generated at precompilation time in order to perform on line verification of GB constraints efficiently. On the one hand, incorporating a large number of constraints into the precompilation phase causes the grammar size to become explosive, thus slowing down grammar search time; on the other hand, eliminating a large number of constraints from precompilation forces a high cost at constraint verification time. Frazier (1986) suggests that all phrase structure possibilities get multiplied out, leaving only a small subset of GB constraints to apply at processing time. In the parser presented here, a relatively small number of GB constraints (those concerning skeletal phrase structures and empty noun phrases) are accessed at precompilation time, leaving many of the GB constraints to apply at processing time. Time tests have shown this clustering of principles the most promising for efficient parsing using the co-routine design.

Both classes include parameters of variation.

The modularity imposed by the GB framework is an improvement over context-free based systems for several reasons: (a) properties common to all languages are not specified directly in rules, but are abstracted into modularized principles, thus allowing linguistic generalization to be captured; (b) multiplicative effects of linguistic constraints are not spelled out in the form of grammar rules, thus reducing grammar size (hence processing time); and (c) a separate description is not required for each language, thus the parser is easily extendible to additional languages.

2 Underlying Linguistic Theory

In order to arrive at the modules that form the basis of the structure-building and GB components of figure 1, we must separate underlying subsystems of grammar that interact to gain the effects of complicated rules systems. This section describes parameters of variation associated with the principles of three GB subtheories ($\overline{\text{X}}$-Theory, θ-Theory and Trace Theory), and discusses the relevance of these parameters within the context of the parsing model. The goal is to incorporate the parameterized principles of GB (in the form of modular subsystems of structural and well-formedness constraints) into a single, cross-linguistically uniform parsing system.

2.1 $\overline{\text{X}}$-Theory Parameters: Choice of Specifiers and Constituent Order

The central idea of $\overline{\text{X}}$-Theory is that the dictionary (henceforth *lexicon*) specifies subcategorization frames for lexical items (*e.g.*, the frame for the verb *put* includes two arguments, one that is a noun phrase, and another that is a prepositional phrase, as in *put the car in the garage*), and phrase-structures are projections of a lexical head X (= N, V, P or A).[8] $\overline{\text{X}}$-Theory assumes that phrase structures for English are derived by rules of the form:

(3) $\text{X}^{max} \Rightarrow$ (Specifier) X (Complement)

where X^{max} is the maximal projection (more commonly called XP) of the lexical head X. The Specifier of X is determined by a parameter setting associated with the $\overline{\text{X}}$ module, and the complement of X is determined by the subcategorization frame of the verb. For example, if X is a noun, X^{max} is NP, a possible Specifier is a determiner, and a possible complement is a prepositional phrase (depending on whether this is specified in the lexical entry for the noun).

English requires that specifiers of all lexical categories occur before the lexical head, and complements follow the lexical head. However, this rule does not apply to all languages (*e.g.*, Navajo, German, Japanese, *etc.*). For example, consider the following Navajo sentence:

[8]The lexical representation used in the parser presented here is based on the input representation required by the morphological analyzer. It includes the root forms of words and pointers to applicable affixes. Root verbs are stored with their argument structure specifications and θ-role assignment possibilities. The lexicon is discussed in Dorr (forthcoming), but will not be emphasized in this paper.

(4) ashkii at'ééd yiyiiłtsą́
'the boy saw the girl'

This sentence literally translates as *the boy the girl saw* since Navajo requires the complement to precede the head.[9] It is assumed that the constituent order of a language is determined by a parameter of variation. Thus, before parsing begins, $\overline{X}$ rules are set up according to the constituent order of the language being parsed. This is crucial in the parsing model since many of the principles of other GB subtheories cannot apply until a valid licensed structure (with predetermined ordering restrictions) has first been built, *i.e.*, $\overline{X}$-Theory provides basic templates to which remaining parsing constraints can apply.

2.2 θ-Theory Parameters: Clitic Doubling

θ-Theory is the theory of thematic (or semantic) roles. A principle of this theory is the θ-Criterion which states that each noun phrase argument of a verb is uniquely assigned a semantic role (*e.g.*, agent, patient, *etc.*) and each semantic role is uniquely assigned to an argument. In order for a semantic role (henceforth θ-role) to be assigned, there is a principle of θ-role transmission that maps arguments in the dictionary entry of the verb to their corresponding θ-roles.

In Spanish, the phenomenon of clitic doubling is relevant to parametric variation of the θ-role transmission principle. A clitic is a pronominal constituent that is associated with a verbal object. For example, the clitic *le* in the following sentence is a clitic associated with *Juan*, the object of the verb *regalé*:

(5) Le regalé un libro a Juan.
'I gave a book to John.'

The phenomenon of clitic doubling is defined in terms of the pair <*clitic, lexical NP*> where the clitic must agree in number, person and gender with the lexical NP. In (5) the clitic *le* actually stands for an NP that does not yet have a θ-role (namely, *Juan*). Thus, in order to satisfy the θ-Criterion, a parameter of variation is required for θ-role transmission. Jaeggli (1981) proposes that clitics supply θ-roles to object NPs that are doubled via a θ-role transmission rule:

(6) $[\text{CL } +case_i \ +\theta_j] \ldots [\text{NP } +case_i] \Rightarrow [\text{CL } +case_i \ +\theta_j] \ldots [\text{NP } +case_i \ +\theta_j]$

This rule allows a doubled NP object to receive θ-role as long as the clitic and NP must have the same case.[10] If a clitic is not present, a θ-role is assigned in the usual fashion, (*i.e.*, from the verb that contains the argument in its dictionary entry). Thus, for languages that allow clitics, clitic doubling must be available as a parameter of variation to the θ-role transmission principle of θ-Theory. The θ-Criterion can then be used as a well-formedness condition during parsing so that clitic doubling constructions will be ruled out unless (6) is allowed to fire. This is important in a parsing model since languages that allow clitics could not be analyzed uniformly without such a parameter of variation.

[9] Hale (1973) describes how this and several other phenomena in Navajo reveal parametric variation to GB principles.

[10] A description of Case Theory is not given here. See Chomsky (1981).

2.3 Trace Theory Parameters: Choice of Traces and Pro-drop

Trace theory is another subtheory of GB that is important for uniform parsing across languages, in particular because it provides an explanation for the distinctions between languages that allow null subjects (like Spanish) and other languages. A trace is an empty sentence position that is either base-generated or left behind when a constituent has moved. The choice of traces for a language is specified as a parameter setting to the trace module.

According to the analysis of the null subject (or pro-drop) parameter introduced by van Riemsdijk and Williams (1986), the choice of whether a language requires a sentential subject is allowed to vary from language to language. In Spanish, as in Italian, Greek and Hebrew, morphology is rich enough to make the subject pronouns redundant and recoverable. Thus, we can have the sentence:

(7) Hablé con ella.
'(I) spoke with her.'

Since the inflection on the verb is first person singular, the subject pronoun *yo* (=I) need not be used.

The formulation of the *pro-drop parameter* by van Riemsdijk and Williams is motivated by the observation that subjects are missing in a variety of constructions, not just in cases like (7). These constructions do not appear in many other languages (*e.g.*, English, *etc.*); thus, there must be a parameter that will account for the distinction between pro-drop and non-pro-drop languages. The *pro-drop parameter*, then, is a minimal binary difference that does or does not allow empty noun phrases to occupy subject position. (For details on the pro-drop parameter, see van Riemsdijk and Williams, pp. 298-303.) The parameter setting approach is more desirable than a rule-based approach since it accounts for several types of null subject constructions without requiring several independently motivated rules.[11] The pro-drop parameter is important in the parsing model because it allows uniform analysis of pro-drop and non-pro-drop languages, ensuring that sentence without a subject are ruled out unless the pro-drop parameter is set.

2.4 Principles and Parameters

Table 1 contains a table summarizing the subsystems of principles and parameters (grouped according to subtheory) relevant to the parsing model as presented here.[12] Table 2 summarizes the parameter settings required for parsing Spanish and English.

3 Parsing Implementation

The parser is one of three translation stages in an interlingual translation system, UNI-

[11] A rule-based approach (*e.g.*, GPSG (1985)) would require a separate rule for every possible null subject construction allowed in a pro-drop language including free subject inversion, relative clauses, that-trace constructions, resumptive pronouns, *etc.* (These constructions are not discussed here. See van Riemsdijk and Williams (1986).) The parameter setting approach obviates the need for independent treatment of these closely related phenomena.

[12] Because of space limitations, only those parameters that are relevant to a condensed description of the parser are presented here. The actual implementation currently has 20 parameters.

Theory	*Principles*	*Parameters*
$\overline{X}$	A phrasal projection (X^{max}) has a head (X), a specifier and a complement	Constituent Order, Choice of Specifiers
θ	[CL $+case_i$ $+\theta_j$] ... [NP $+case_i$] $\Rightarrow$ [CL $+case_i$ $+\theta_j$] ... [NP $+case_i$ $+\theta_j$] if language allows clitic doubling	Clitic Doubling
Trace	Null subjects are allowed for pro-drop languages	Pro-drop
	An empty position may occur where traces are allowed	choice of traces

Table 1: Principles and Parameters of GB

Theory	*Parameters*	*Parameter Values*	
		Spanish	English
$\overline{X}$	Constituent Order	spec-head-comp	spec-head-comp
	Choice of Specifiers	V: have-aux; N: det, *etc.*	V: have-aux, do-aux; N: det, *etc.*
θ	Clitic Doubling	applicable and allowed	not applicable
Trace	Pro-drop	yes	no
	Choice of Traces	N^{max}, Wh-phrase, V, P^{max}	N^{max}, Wh-phrase, V, P^{max}

Table 2: Parameter Values for Spanish and English

TRAN (Dorr, forthcoming), which is implemented in Commonlisp and currently translates simple sentences bidirectionally between Spanish and English. In contrast to the transfer approach (*e.g.*, METAL, Slocum, 1984, 1985), the parser (and other translation modules) is uniform across all languages with respect to its theoretical and engineering basis. (See figure 3.) The transfer approach, on the other hand, requires several parsers and a third translation stage (the transfer stage) in which one language-specific representation is mapped into another. Thus, a separate parser must be supplied for each language in the transfer approach, while in the interlingual approach a single parser is used for all languages. The interlingual approach more closely approximates a true universal approach since the principles that apply across all languages are entirely separate from language-specific characteristics expressed by (user-modifiable) parameter settings.[13]

The parameters of table 2 are represented declaratively, and are subject to modification by the user. (See figure 4.) There are two types of procedures (corresponding to the two boxes of figure 1) within the system: the first type includes those procedures that perform structure-building actions (predicting, attaching and scanning), relying primarily on phrase structure templates generated at precompilation time; and the second type consists of constraint verification routines (θ-Criterion, empty NP conditions, *etc.*), performing well-formedness tests on phrase-structures built by structure building procedures.

[13]The approach is "universal" only to the extent that the linguistic theory is "universal." There are some residual phenomena not covered by the theory that are consequently not handled by the system in a principle-based manner. For example, the language-specific English rules of *it-insertion* and *do-insertion* cannot be accounted for by parameterized principles, but must be individually stipulated as idiosyncratic rules of English. Happily, there appear to be only a few such rules per language since the principle-based approach factors out most of the commonalities across languages.

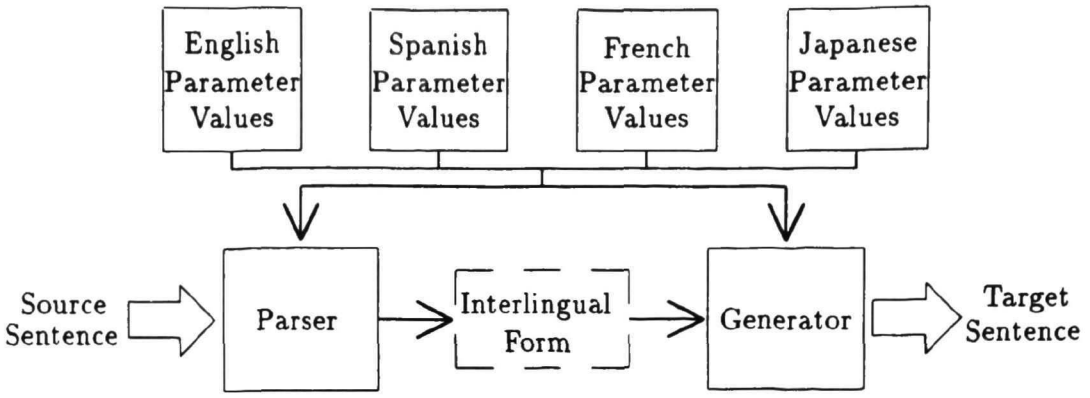

Figure 3: Interlingual Design as found in Dorr 1987

```
(DEF-PARAM CONSTITUENT-ORDER
   :SPANISH (SPEC HEAD COMP) :ENGLISH (SPEC HEAD COMP))

(DEF-PARAM CHOICE-OF-SPEC
   :SPANISH (V (HAVE-AUX) N (DET) I (N-MAX) C (WH-PHRASE))
   :ENGLISH (V (HAVE-AUX DO-AUX) N (DET N-MAX) I (N-MAX) C (WH-PHRASE)))

(DEF-PARAM CLITIC-DOUBLING :SPANISH (T T) :ENGLISH (NIL NIL))

(DEF-PARAM PRO-DROP :SPANISH T :ENGLISH NIL)

(DEF-PARAM CHOICE-OF-TRACES
   :SPANISH (N-MAX WH-PHRASE V P-MAX) :ENGLISH (N-MAX WH-PHRASE V P-MAX))
```

Figure 4: Representation of Parameter Settings for Spanish and English

Before parsing begins, the precompilation stage generates and stores a constant number of underspecified phrase-structure templates per language according to the two $\overline{X}$ parameters of figure 4: constituent order and choice of specifier. When the parser is activated, the structure-building module draws upon these templates, processing each word of input until no more structure-building actions apply. At this time, constraint verification takes place, and the last three parameters of figure 4 are accessed in order to modify or eliminate the structures derived thus far. The parse proceeds in this fashion until all sentence constituents have been successfully scanned, and all constraints have been verified. A sentence is rejected if: (a) there is a constraint violation, or (b) after consulting the constraint module no structure-building actions apply to the remaining input words. A sentence is accepted otherwise. Because the constraint component is available during parsing, the phrase-structure templates accessed by the structure-building module need not be very elaborate; consequently the grammar size need not, and should not, be as large as those found in other parsing systems.[14] Thus, the

[14]In fact, the number of phrase structure templates that are generated per language generally does not exceed 150 since there are a limited number of configurations per language that are allowed by the $\overline{X}$

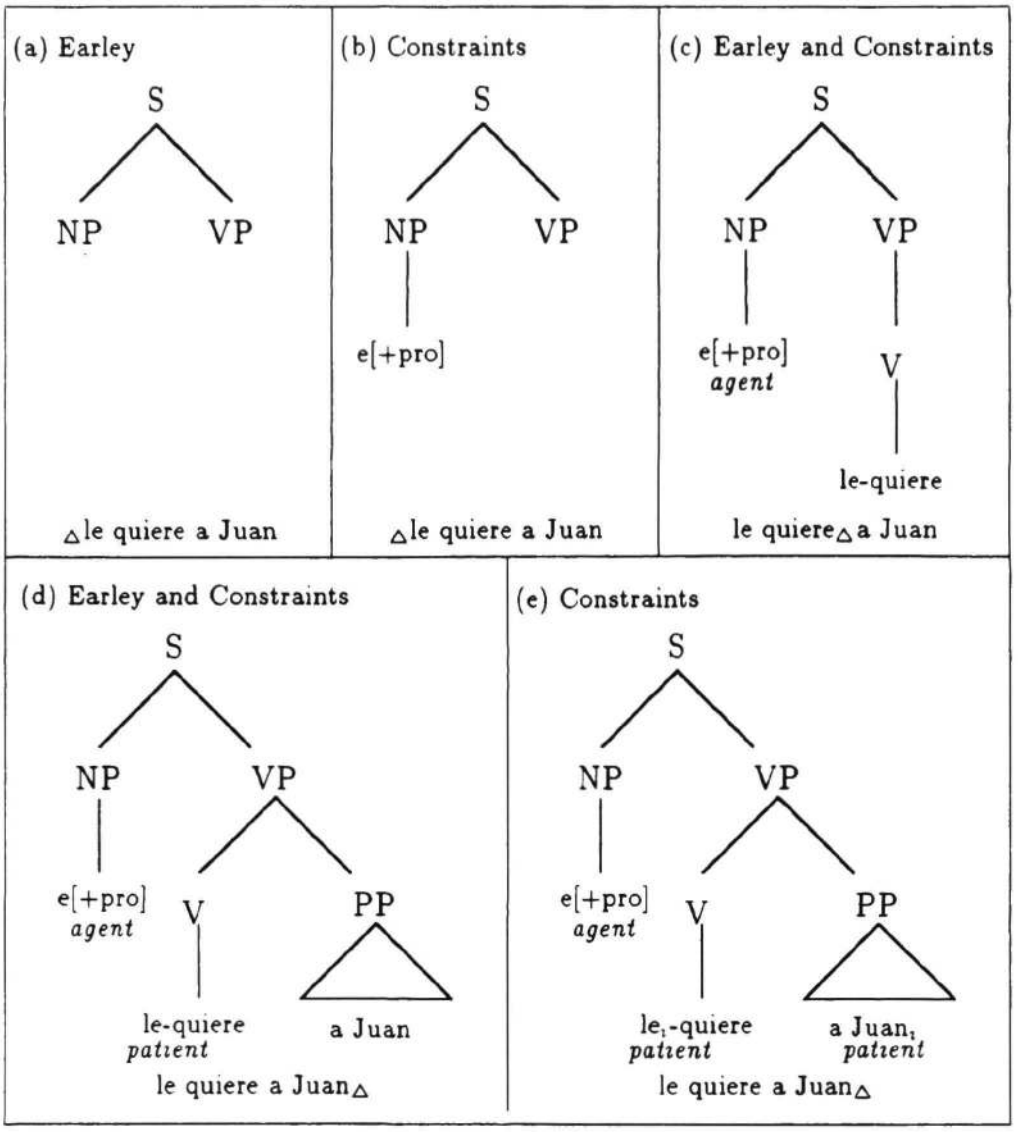

Figure 5: Snapshots of Parser in Action

system avoids high computational costs due to grammar search time.

To clarify the above description of the parsing algorithm, the next section presents and example of how the parsing modules operate.

4 An Example

Consider the problem of parsing (1) repeated here as (8):

(8) Le quiere a Juan
'(She) loves John'

principles accessed at precompilation time. Thus, the running time of the parser is not subject to the same slow-downs that are found in other systems.

Figure 5 gives snapshots of the parser in action. First the Earley structure-building component predicts that the sentence has a noun phrase (NP) and a verb phrase (VP) (see (a)), the order of which is determined by the "constituent order" parameter at precompilation time.[15] The only structures available for prediction by the Earley module are those generated at precompilation time; thus, at this point no further information about the structure is available until the linguistic constraint module takes control.

The constraint module accesses the "null subject" parameter (see section 2.3), which dictates that the empty element attached to NP is a subject. The [+pro] (pronominal) feature is associated with the node (see (b)) so the subject will accommodate both null subject source languages and overt subject source languages.[16]

In snapshot (c), the Earley module expands VP and scans the first two input words *le quiere*.[17] Now the Earley module cannot proceed any further; thus, the constraint module takes over again. First a semantic role (or θ-role, as it is called in GB Theory) of *agent* is assigned to the empty subject of the sentence. This information is determined from the dictionary entry of *quiere* which dictates that this verb requires both an agent (assigned to the subject or *external argument* of the verb) and a patient (assigned to the object or *internal argument* of the verb). The dictionary entry for *querer* (the root form of *quiere*) is encoded as follows: **(querer: [ext: agent] [int: patient] V (english: love) (french: aimer) ...)**

Now the constraint module predicts that a noun phrase (corresponding to the internal argument of *querer*) must be available. Because the clitic-doubling parameter is set to (T T), it is determined that the NP *le* can act as an object of the verb *quiere*; consequently, it receives *patient* θ-role as dictated by the lexical entry of *querer*. The constraint module then "records" the fact that a clitic has been seen, so that the NP corresponding to *le* will have a θ-role transmitted to it later if it appears in the input.[18] Once control passes back to the Earley module, the final two words are scanned, thus completing the PP. Snapshot (d) shows the parse thus far.

At this point the constraint module attempts to assign θ-role to the NP *Juan*. However, all of the θ-roles from the lexical entry of *querer* have already been assigned; thus, assigning a role from this entry would be a violation of the θ-criterion. On the other hand, leaving *Juan* without a role also violates the θ-criterion. Consequently, the constraint module determines (via the clitic-doubling parameter setting) that the θ-role transmission rule (6) is applicable, and recognizes that the NP *Juan* corresponds to the "recorded" clitic preceding the verb *quiere* (since the two match in person, number and gender). Thus, a θ-role of *patient* is transmitted to *Juan*.[19] As a result of the application of the θ-transmission rule, *le* and *Juan* are coindexed; thus, these two constituents are interpreted as coreferential during the stages following the parse. The final parse is illustrated in snapshot (e).

---

[15]Since Spanish is a *head-initial* language, NP must precede VP; however, this would not be the case for non-*head-initial* languages. (See fn. 5 for a description of the "constituent order" parameter.)

[16]For example, Italian and Hebrew do not require an overt subject, but English and French do; thus, during a later stage (generation), e[pro] will either be left as is, or lexicalized to a pronominal form (*e g.*, *he* or *she* in English) that agrees with the main verb.

[17]Clitic adjunction is generated at precompilation time. The presence or absence of a clitic for a particular language is determined by an adjunction parameter setting associated with $\overline{X}$. This parameter will not be discussed here.

[18]Since clitic doubling is optional, the parse will not be discarded if the corresponding NP does not appear in the input; however, if it does appear (as it does in the above example), it is correctly assigned θ-role.

[19]Note that the θ-role *patient* is assigned the NP *Juan*, not to the PP *a Juan*; in general, the structural entity that is assigned semantic role is an NP, regardless of the type of phrase the containing it.

5 Conclusion

The system described here is based on modular theories of syntax that include systems of principles and parameters rather than complex, language-specific rules. The "co-routine design" allows the structure-building mechanism to operate with user-modifiable principles of current linguistic theory. The user has access to parameters associated with the system principles, thus enabling extension of the system to additional languages. The presence of linguistic constraints allows phrase structure templates to be underspecified (*i.e.*, more general), thus reducing the grammar size of a given language. In summary, the modularity imposed by the GB framework is an improvement over context-free based systems because it facilitates extensions and alterations to the system, simplifies descriptions of natural grammars, and is backed by psycholinguistic evidence (see fn. 3).

Acknowledgements

This report describes research done at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support for the Laboratory's artificial intelligence research has been provided in part by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research contracts N00014-80-C-0505 and N00014-85-K-0124, and also in part by NSF Grant DCR-85552543 under a Presidential Young Investigator's Award to Professor Robert C. Berwick. Useful guidance and commentary were provided by Bob Berwick, Ed Barton, Sandiway Fong and Dave Braunegg.

References

Barton, Edward G. Jr. (1984) "Toward a Principle-Based Parser," MIT AI Memo 788.

Bates, M. (1978) "Natural Language Communication with Computers," Springer-Verlag, 191–254.

Chomsky, Noam A. (1981) *Lectures on Government and Binding*, Foris Publications, Dordrecht.

Chomsky, Noam A. (1982) "Some Concepts and Consequences of the Theory of Government and Binding," MIT Press.

Dorr, Bonnie J. (forthcoming) "UNITRAN: A Principle-Based Approach To Machine Translation," S.M. thesis, Department of Electrical Engineering and Computer Science, MIT.

Earley, Jay (1970) "An Efficient Context-Free Parsing Algorithm," *Communications of the ACM* 14, 453–460.

Frazier, Lyn (1986) "Natural Classes in Language Processing," presented at the *Cognitive Science Seminar, MIT, November*, Cambridge, MA.

Gazdar, G., E. Klein, G. Pullum, and I. Sag (1985) *Generalized Phrase Structure Grammar*, Basil Blackwell, Oxford, England.

Hale, K. (1973) "A Note on Subject-Object Inversion in Navajo," in *Issues in Linguistics: Papers in Honor of Henry and Renee Kahane*, B. Kachrue *et. al.* (eds.), University of Illinois Press, Urbana.

Jaeggli, Osvaldo Adolfo (1981) *Topics in Romance Syntax*, Foris Publications, Dordrecht, Holland/Cinnaminson, USA.

Robinson, J. J. (1982) "DIAGRAM: A Grammar for Dialogues," *Communications of the ACM* 25:1, 27–47.

Sharp, Randall M. (1985) "A Model of Grammar Based on Principles of Government and Binding," M.S. thesis, Department of Computer Science, University of British Columbia.