MASSACHUSETTS INSTITUTE OF TECHNOLOGY
ARTIFICIAL INTELLIGENCE LABORATORY

A.I. Memo 964

January, 1988
Revised December, 1988

Complexity of Human Language Comprehension

Eric Sven Ristad

Abstract:

The goal of this research has been to understand the computational structure of principle-and-parameter linguistic theories: what computational problems do these theories pose and what is the underlying structure of those computations? To do this, I have analyzed the computational problem of human language comprehension: what linguistic representation is assigned to a given sound? This *language comprehension problem* may be factored into smaller, interrelated (but independently statable) problems defined on partial phonological, morphological, and syntactic representations. For example, in order to understand a given sound, the listener must assign a phonetic form to the sound; determine the morphemes that compose the words in the sound; and calculate the linguistic antecedent of every pronoun in the utterance. I prove that these and some other subproblems are all NP-hard, and that language comprehension is itself PSPACE-hard, according to current linguistic theory.

Acknowledgements: This work has benefited from the guidance, criticism, and concern of Professors Robert C. Berwick, Noam Chomsky, Peter Elias, Morris Halle, James Higginbotham, Alec Marantz, and Michael Sipser.

The author is supported by a graduate fellowship from the IBM Corporation. This report describes research done in part 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 a grant from the Kapor Family Foundation, in part by NSF Grant DCR-85552543 under a Presidential Young Investigator Award to Professor Robert C. Berwick, and in part by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research contract N00014-85-K-0124. This report is a preliminary draft of the formal results of my (incomplete) doctoral research, portions of which have been submitted for publication. Theorems 2 and 4 have already appeared in the abstracts of the 11th GLOW Colloquim on Locality, Budapest, March 28–30, 1988.

Contents

1 Introduction to the Reductions

If linguistic theory is a theory of the human language ability, as rational scientists must claim, then linguistic theory must have empirical consequences. Some of those consequences will be computational. At the very least, language users must use the linguistic representations extensionally characterized by the generative grammars of modern linguistics. This in and of itself states a class of computational problems: what linguistic representation is assigned to a given sound, what articulations express a given underlying linguistic representation, and what is the internal state of a language learner after a given sound? Each of these problems—comprehension, production, and acquisition—may be profitably factored into smaller, interrelated (but independently statable) problems defined on partial linguistic representations. The language comprehension problem includes as sub-problems the phonological problems of determining stress, tone, and syllable, morphological, and articulatory structure; the syntactic problems of detecting empty categories, disambiguating words, determining phrase and argument structures, and computing the linguistic antecedents of anaphora. More abstractly, these sub-problems of comprehension consist of completing an independently defined portion of an incomplete linguistic representation.

The goal of my doctoral research has been to determine the computational structure of current linguistic theory: what problems does the theory pose and what is their underlying structure? To do this, I have studied abstract problems in language comprehension with computational complexity theory, a theory of the absolute difficulty of solving computational problems, in terms of their natural parameters. The major technical result of the thesis is that these computational problems are all intractable and cannot be solved in practice.

The greatest difficulty of the research lies in creating precise linguistic models from the inconsistent, incomplete, and imprecise linguistics theories that are available today. These models must be sufficiently abstract to capture the core ideas of modern linguistics in order to remain relevant despite rapid change in the field. My approach has been to reduce the instances of some known problem P into a class of linguistic phenomena such that understanding that phenomena corresponds to solving P. As far as I know, the reductions hold for all current linguistic theories that explain these phenomena. This places these proofs among the strongest formal results achievable in an empirical science.

This manuscript contains the first complexity results for modern (1980s) generative linguistic theory, a task thought by many linguists to be impossible. Moreover, the results cover the entire spectrum of linguistics: phonology, morphology, and each level of syntactic representation (D-structure, S-structure, and Logical Form).

2 The Complex Structure of Sounds and Words

Words are built from units of meaning called morphemes. In the more familiar languages of the world, such as Romance languages, morphemes are concatenated to form words: this process is called affixation. In some languages, such as Semitic languages where vowels are morphemes, a morpheme may appear more that once *inside* another morpheme (this is called infixation). For example, the Arabic word *katab*, meaning 'he wrote', is formed from the active perfective morpheme *a* infixed to the *ktb* morpheme.

In addition to morphemes, words consist of syllables. Each syllable contains one or more vowels V (its nucleus) that may be preceded or followed by consonants C. For example, *katab* consists of two syllables, the two-segment syllable CV and the three-segment closed syllable CVC.

Syllable structure can interact with morphological structure. For example, all Arabic words must end in the closed syllable CVC, and therefore the *a* morpheme must be infixed to the *ktb* morpheme in order to form a word with permissible syllable structure. It cannot be prefixed, because Arabic does not permit syllables of the form VCCC; nor can it be suffixed, because CCCV syllables are not permitted. These interactions appear simple, but they can become extremely complex when morphemes are ambiguous or when segments are underspecified as in distorted speech. In fact, determining the morphological and syllable structure of an underspecified sound is like solving the satisfiability problem for Boolean formulas, where one-segment morphemes are Boolean variables, vowels true literals, and consonants false literals. Then three-segment syllables would correspond to satisfied 3-CNF clauses, and words to entire 3SAT instances.

Among the world's languages, assimilation is one of the most common phonological processes. Assimilation is the process whereby some segment comes to share properties of another segment in the course of the derivation of surface forms from underlying forms. For example, in English, consonant nasality assimilates to immediately preceding vowels; assimilation also occurs across morpheme boundaries, as the varied surface forms of the prefix */in-/* demonstrate: */in-/+/logical/* → *[illogical]* and */in-/+/probable/* → *[improbable]*. In other languages, assimilation is unbounded and can affect nonadjacent segments: these assimilation processes are called harmony systems. In the Turkic languages all suffix vowels assimilate the backnesss feature of the last stem vowel; in Capanahua, vowels and glides that precede a word-final deleted nasal (an underlying nasal segment absent from the surface form) are all nasalized.

Harmony systems are powerful computational mechanisms. In fact, we can virtually duplicate the preceding reduction idea merely by replacing morpheme infixation with harmonic features. In the second reduction then, harmonic features will represent Boolean variables, and harmony processes will ensure consistent truth assignments to the Boolean variables; the other details are essentially unchanged.

2.1 Word Recognition in the Nonlinear Model

In phonological theory, segregated representations obey simple principles; this section examines the complexity of these interacting representations (planes and tiers) and morpheme combination (affixation and infixation). In order to do this, we pose the word recognition problem:

> **Word Recognition Problem (WORD)**
> Given an underspecified timing slot vector V, and a morphological model M, is V a permissible word according to M?

The timing slot vector is an abstract representation of a speech sound in terms of the articulations (phonemes) necessary to produce that sound. It is underspecified when articulations necessary to produce a word of the language are missing from the representation.

Theorem 1 *WORD is NP-hard in the rule-free nonlinear morphological model.*

Proof. The proof is by a reduction from 3SAT to WORD. The reduction input is a 3-CNF formula F containing clauses $C_1, C_2, \ldots, C_p$ in the Boolean variables $x_1, x_2, \ldots, x_n$. Each clause contains exactly three distinct literals labeled by $C_i = (a_i \vee b_i \vee c_i)$. The output is a rule-free nonlinear model M (list of morphemes, feature geometry, and syllable templates) and an underspecified timing slot vector V, such that V is a permissible word according to M if and only if F is satisfiable.

Each variable x_i is represented by a one-segment morpheme μ_i, and each clause C_j is represented by a three-segment syllable σ_j. The reduction constructs the morphemes $\mu_1\mu_2\ldots\mu_n$, plus two features: [$\pm$TRUE] to represent the truth value assigned to a variable, and [$\pm$NEG] to represent whether a literal is negated or not. Segments specified with the features [+NEG,-TRUE] or [-NEG,+TRUE] represent true literals, which correspond to vowels. For example, the false literal $\overline{x_4}$ is represented by the morpheme μ_4 with the feature vector [+NEG,+TRUE].

Syllable templates ensure that each 3-CNF clause contains one or more true literals by requiring syllables to span three segments and contain one or more true literals (FFT, FTF, FTT, TFF, TFT, TTF, TTT). In other words, we are simply excluding FFF, one of eight possible three-segment 'truth templates,' which corresponds to positing a language where syllables span exactly three segments and must include a vowel.

The timing slot vector V contains the string of formula literals, where each formula literal is represented by a morpheme and the appropriate setting of the NEG feature.

$1+n$ distinct planes are needed: a single one-tiered plane for each morpheme *qua* variable, and one two-tiered plane for syllables *qua* satisfied 3-CNF clauses. The reduction requires $O(n)$ time and space to specify the nonlinear model and does not use any morphological rules, which strongly suggests that the nonlinear model is more powerful than $\mathcal{NP}$.$\square$

2.2 Sound Recognition in the Nonlinear Model

Sound Recognition Problem (SOUND)
Given an underspecified timing slot vector V, and a phonological model M, is V a permissible phonological representation according to M?

Theorem 2 *SOUND is NP-hard in the rule-free nonlinear phonological model.*

Proof. The proof is by reduction from 3SAT to SOUND, and follows the preceding proof of theorem 1 in its idea. Phonological harmony will replace morpheme infixation as the linguistic process that ensures consistent Boolean truth assignments. As before, the reduction input is a 3-CNF

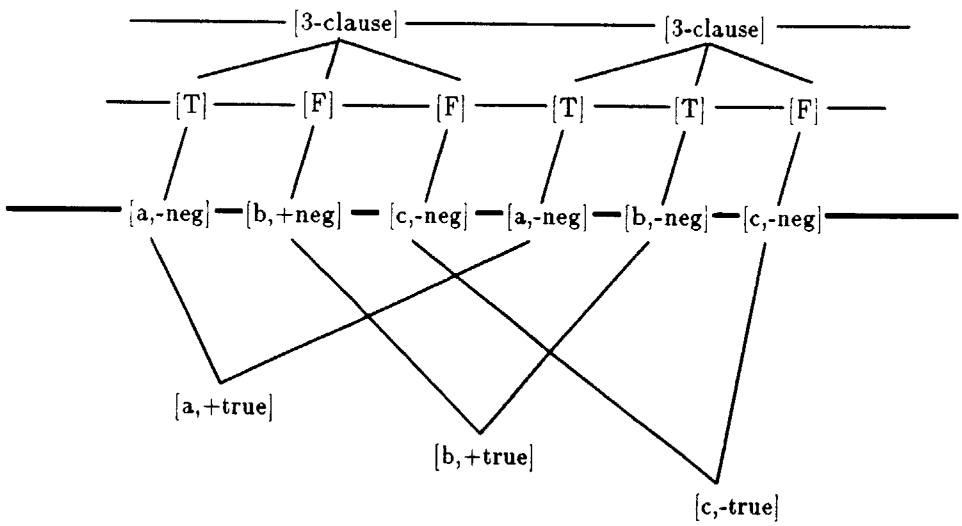
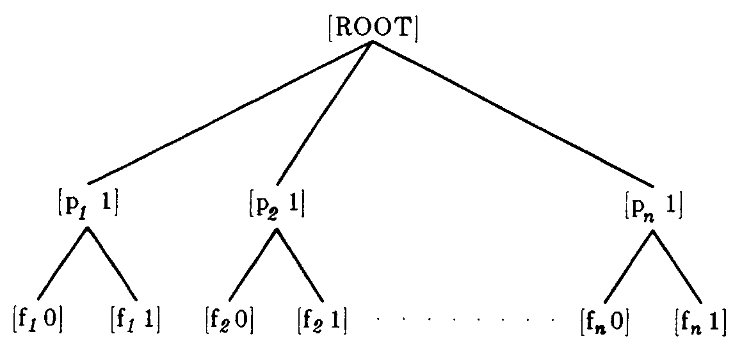


Figure 1: Nonlinear morpho-phonological representation of the 3SAT formula $f = (a \vee \bar{b} \vee c) \wedge (\bar{a} \vee b \vee c)$, satisfied by a, b true and c false. The timing slot vector, drawn horizontally in the center of the figure with a heavy line, represents the string of formula literals. Each of the three one-segment morphemes, drawn below the timing slot vector with noncrossing association lines, resides on its own plane and represents a formula variable. The final two tiers, drawn above the timing slot vector, represent satisfied 3-CNF clauses: one T,F-tier for literals and a second tier for satisfied 3-CNF clauses.

formula F and the output is a (partial) rule-free nonlinear phonological model M (rooted feature geometry, syllable templates, and list of harmony processes) and an underspecified timing slot vector V, such that V is a permissible phonological representation according to M if and only if F is satisfiable.

Each variable x_i is represented by a binary place-of-articulation feature p_i; the truth assignment to x_i is encoded in the binary harmonic articulatory feature f_i; and each clause C_j is represented by a syllable σ_j spanning three segments. The reduction also constructs the binary feature [$\pm$NEG] to represent whether a literal is negated or not. Each harmonic feature f_i is immediately dominated by a place-of-articulation feature p_i, which is connected to the root:



Segments specified with the features [p_i1, f_i0, +NEG] or [p_i1, f_i1, -NEG] represent true literals of the variable x_i, and will correspond to vowels. For example, the false literal $\overline{x_4}$ is represented by the feature matrix [$p_4$1, $f_4$1, +NEG].

Being harmonic, each feature f_i will receive the same value throughout the timing slot vector. As before, syllable templates ensure that each 3-CNF clause contains one or more true literals by requiring syllables to span three segments and contain a true literal. The timing slot vector V contains the string of formula literals, where each formula literal is represented by its place of articulation feature p_i and the appropriate setting of the NEG feature. For example, the negated literal $\overline{x_4}$ is represented by the underspecified feature matrix [$p_4$1, +NEG].$\square$

Comments. The reductions to the nonlinear model are blocked:

- If planar interaction is limited to superposition of tier features. In this highly restricted nonlinear model, each articulatory feature can be associated with at most one plane: planes are truly computationally independent because they interact only in that they each may affect the acoustic signal. But this nonlinear superposition model is far too impoverished for natural language, where all phonological processes have access to one small fixed set of articulators and hence must make maximal use of their representational possibilities (McCarthy, 1981). For example, consonants and vowels are segregated into distinct c,v-planes and are representationally distinguished by the non-phonetic [consonant] feature. But consonants and vowels must both use the same set of phonetic features in articulation. They are distinguished in articulation only by the obstruction or nonobstruction of the air flowing from the lungs to the lips. (Consonants require an articulator to contact the stationary part of the vocal tract.) But "this striking difference between the production of vowels and consonants must not be allowed to obscure the obvious fact that for the production of both types of sounds speakers have at their disposal only a single piece of anatomical machinery, the vocal tract with its six articulators. . ." (Halle, 1987).

  (Nonetheless, the idea of the second reduction is problematic, because every segment's [consonant] feature is directly affected by some harmonic feature in that segment. I do not know of any harmony processes with this character.)

- In the reductions, the timing slot vector, which is an abstract representation of the acoustic signal, contains sufficient information to identify each literal/morpheme/place-of-articulation, while systematically omitting the feature that encodes truth assignment. At first glance this is plausible because, as is well known, human speech is fantastically resistant to distortion, no matter how underspecified the acoustic signal may be.

  One way to block the reductions would be to require the nonlinear representation to be stable with small changes in perceived phonetic features. As things stand, this is not a property of the preceding reductions, which would fail because we could not ensure segregation of formula variables or consistency of truth assignments in the presence of noise. We might meet this stability requirement by using more phonetic features, to redundantly identify variables and their truth values.

- We might disallow repeated morpheme infixation or universally bound the number of harmony processes by a small constant. Both infixation and harmony are bounded in known languages: no language is thought to have more than two harmony processes, and morpheme infixation appears to be limited to infixing one vowel per morpheme. Unfortunately, there is no elegant way to explain boundedness beyond mere stipulation in an approach to linguistic theory that distinguishes knowledge and ability, as generative linguistics does.

2.3 Postmortem

The aim of the preceding analysis was to determine the complexity of planar interactions, harmony processes, and morpheme combination in the non-linear model. The reduction very clearly shows that the simplest planar interactions are complex, regardless of (the potential complexity of) both interplanar computations, such as stress shift, and determining the correspondence between surface and underlying forms. With respect to morpheme infixation and phonological harmony, we can see that both processes are costly when coupled with articulatory underspecification.

3 Structure of transformational models

The number of levels of linguistic representation, their qualities, and the rules relating them are central concerns in generative linguistics. Within the principles and parameters approach, syntactic theories have been broadly classified both by the levels of representation they employ (S-structure and PF, and sometimes D-structure, NP-structure, and/or LF) and by the rule determining the relation between a trace and its antecedent (move-alpha versus 'rules of construal,' such as Rizzi's chain formation algorithm, Kayne's g-projection, or Koster's dynasty formation).

Transformational theories include D-structure and S-structure representations, and hold that S-structure is derived from D-structure by successive move-alpha transformations, where each application of move-alpha uniquely relates a trace to its antecedent, and where all relations between traces and their antecedents determined by such a derivation are represented at S-structure.[1]

3.1 Empirical properties of move-α

It has been difficult to find empirical arguments favoring the transformational approach over competing 'representational' approaches. Arguments based on exhibiting natural properties of D-structure that are difficult to state at S-structure have not been entirely persuasive because S-structure (redundantly) represents all information represented at D-structure, due to trace theory and the projection principle. Thus, natural properties of D-structure will also be naturally statable at S-structure in terms of traces.

Arguments for transformational theories based directly on empirical properties of move-alpha derivations have been the most persuasive, simply because theories lacking move-alpha derivations will lack those empirical properties. Such arguments exhibit properties of move-alpha or entire derivations that are difficult to state as S-structure rules of construal between traces and

[1]Modern transformational theories are often misleadingly named 'derivational' in contrast to the competing 'representational' approaches. These widely-used appelations are misleading because all generative linguistic theories admit explicitly specified representations, that is, representations *derived* according to some effective procedure. The only distinquishing issues are the role of move-alpha in the derivation of syntactic representations, and the status of D-structure, NP-structure, and LF in the theory.

their potential binders. For example, in his Fall 1987 class lectures Chomsky argued, with empirical evidence, that derivations must obey a 'least effort' principle: that is, the number of derivation steps required to relate a trace to its S-structure binder must be minimal. Such global properties of derivations are thought, without proof or supporting argument, to be difficult to express as S-structure rules of construal.

Another powerful argument for move-alpha is that the conditions on move-alpha can be satisfied neither at D-structure nor S-structure, but only at an intermediate stage of the derivation. This obtains whenever an application of move-alpha disturbs the obligatory structural relation (typically, c-command) between a previously moved constituent and its trace, as when a constituent β containing a trace t_i is moved outside the c-command domain of its antecedent α_i, as in (1).

(1)
a. $...[\ \alpha_i ... [_\beta ... t_i ...]]$
b. $[[_\beta ... t_i ...]_j ... [\ \alpha_i ... t_j\ ]]$

Then the relation between the trace t_i and its binder α_i will not obey the same structural conditions at S-structure as other traces and their binders do. In (1b), t_i is neither c-commanded nor bound by its antecedent at S-structure, yet the resulting S-structure is surely permissible, as in passive VP topicalization structures:

(2)
a. John$_i$ has never been $[_{VP}$ arrested t_i by the FBI$]$

$[_{VP}$ arrested t_i by the FBI $_j$ John$_i$ has never been $t_j]$

(The complexity proofs below will rely on this type of "rolling movement" construction.) The choice between transformational and alternate theories ultimately reduces to the necessity of representations intermediate between D-structure and S-structure: if intermediate representations are necessary, as has been argued in the literature, then syntactic representations must include ordered transformational derivations.

3.2 Conceptual objection to transformations

Transformational grammars (TGs) seek to explicitly relate linguistic structures via a sequence of transformations. For example, in transformational theories, a transformation relates an active sentence to its passive variant. But there is a major conceptual problem with this approach to linguistic theory: TGs do not construct explicit representations. In a TG, linguistic relations are implicitly encoded both in the underlying and surface forms and in the series of transformations required to derive the surface form from its underlying form. In effect, transformations serve double duty, both as a linguistic representation and as the derivational rules of a model of computation.

For example, in transformational grammars elements are nonlocally related in a derivation of syntactic forms using the unbounded one-to-one move-α transformation, locally bounded by linguistic principles such as the empty category principle (ECP) and chain condition. The representations assigned by these theories (chains) include intermediate constituents (traces) whose only apparent role is to allow the iterated local representation of nonlocal relations. Intermediate traces represent only the history of the computation of nonlocal relations.[2]

As a consequence of this failure to separate representation and the process whereby it is constructed, important linguistic relations, such as the nonlocal antecedent-trace relation, are obscured by the transformational system. This is also the reason why the transformational component of current linguistic theories (move-alpha) bears so no resemblance to the rest of the theory (that is, to phrase and argument structure; to case, control, binding, and theta relations; or to the lexicon).

Why should a linguistic theory seek to hide the structure of linguistic representations in order to explicitly encode the derivational relations between those representations? This corresponds to statement of the 3SAT problem in terms of truth assignments and rules that derived all possible 3SAT instances satisfied by a common truth assignment from that given truth assignment. What would be the point? The best statement of a problem

[2] These intermediate traces are the result of a particular conception of the ECP as a local bound on move-α. I would be more persuaded of their existence if they interacted with other grammatical components. For example, they might have binding or phonological effects. Adjunct traces may satisfy the ECP only via antecedent government at LF; as a consequence, adjunct extraction results in intermediate traces that may not be deleted at S-structure. Thus, the only intermediate traces required at S-structure are the traces of adjunct extraction, but these non-case-marked traces do not block *want+to* $\rightarrow$ *wanna* contraction, which is only blocked by case-marked elements (Chomsky 1986a:162). For example:

(3) how do you wanna solve the problem?

As expected, the intermediate traces in specifier of CP and adjoined to VP do not block phonological contraction. Neither do these intermediate $\overline{\text{A}}$-traces affect binding relations, whose domain is NPs in A-positions:

(4) [which woman]$_i$ did John [$_{\text{VP}'}$ dream
$$[_{\text{CP}}\ t_i^2\ [_{\text{IP}}\ \text{Bill}\ [\ \text{I}^0\ [_{\text{VP}}\ t_i^1\ [_{\text{VP}}\ \text{saw} \left\{ \begin{array}{c} \text{Joan} \\ *\ \text{herself} \end{array} \right\} [_{\text{PP}}\ \text{with}\ t_i]]]]]]]$$

The governing category of the direct object is IP (the complete functional complex), and therefore the c-commanding trace t_i^1 adjoined to VP could bind the anaphor in object position within its governing category, if the trace were in an A-position. But, as expected, *herself* is in fact unbound, which strongly suggests that t_i^1 is only relevant to the computation of nonlocal $\overline{\text{A}}$-movement as constrained by the ECP. The precise formulation of the ECP, and the existence of the intermediate traces it requires, is the topic of much active research and debate. But the fact that these intermediate traces do not enter into other syntactic relations casts doubt on their explicit syntactic representation, at least in my mind.

is the most direct characterization of problem instances; relations among problem instances are of secondary concern. I believe that the primary goal of linguistic theory should be explicit representation. Relations among those representations, which are of secondary concern, will be self-evident through the structures they share (by accident, as it were).

4 The Complex Structure of Sentences

The syntactic structure underlying even simple constructions can be quite intricate. For example, according to current syntactic theory, the simple passive sentence *John was seen* is assigned a complex syntactic structure, whose phrase structure component follows:

(5) $[_{\text{IP}}\ [_{\text{NP}}\ \text{John}]_i\ [\text{was}\ [_{\text{VP}}\ \text{see+en}\ t_i]]]$

Most interestingly, each word interacts strongly with every other word in the sentence. The passive verb *see+en* selects the underlying object t_i and assigns it a 'patient' thematic role; the underlying object appears as the surface subject *John*; the subject agrees with the auxiliary verb *was*, which assigns it nominative case; and, to complete the circle of interactions, the auxiliary verb selects the passive verb. These properties of words, such as case-marking, thematic role assignment, selection and agreement, are all independent, not directly deducible from the phonological form of the words, and potentially correlated in the lexicon.

It is easy to see that interactions among the words in a sentence can potentially become extremely complex. For example, imagine if the lexicon contained two slightly different verbs, see_1 and see_2, with the same phonological form (homophones) but different selectional restrictions. Then verb phrases could encode satisfied 2-CNF clauses: see_1 would be false and select a true argument; conversely, see_2 would be true and select either a true or false argument. The consequence is that any verb phrase headed by *see* would contain a word representing a true literal. We could even get two literals of the same variable to agree on truth value by moving one to the subject position of the other, where they must agree, exactly as in the passive construction: the underlying object moves to the subject position, where it must agree with the auxiliary verb. Then if words were Boolean literals, it might even be possible to encode 3SAT instances in sentences.

5 Complexity of Move-α

Transformational derivations have played a central historical role in generative linguistics. In each case, transformational models have been proven intractable (Peters and Ritchie, 1973; Rounds, 1975). To investigate whether current transformational models, based on the generalized move-α transformation, share the computational characteristics of their ancestors, we define a natural problem posed by our language ability and determine its complexity according to the linguistic model. Consider the lexical resolution problem, a subproblem of language perception:

> **Lexical Resolution Problem (LRP)**
> Given an S-structure S with ambiguous or underspecified words, and lexicon L containing ambiguous words, can the words in S be found in the lexicon L?

Theorem 3 *The LRP is NP-hard in transformational models with move-α*

Proof. The proof is by a reduction from 3SAT to LRP; it begins with a sketch of the proof idea along with a simple example. I then state the essential characteristics of the transformational models necessary for the reduction to succeed.

The input to the reduction is a 3-CNF formula F containing clauses $C_1, C_2, \ldots, C_p$ in the Boolean variables $x_1, x_2, \ldots, x_n$. Each clause contains exactly three distinct literals labeled by $C_i = (a_i \vee b_i \vee c_i)$. The output is a lexicon L and an S-structure S containing underspecified words such that the words in S can be resolved found in L if and only if F is satisfiable.

The first step is to create a D-structure that represents F, where the ith constituent represents the ith formula literal and selects the $i+1$th constituent (see figure 2). The selectional properties of constituents will ensure that each 3-clause contains at least one true literal, although lexical ambiguity will prevent us from knowing which literals in the 3-clause are true. To do this, the first literal a_i of a 3-clause C_i must promise to make the 3-clause true, either by being true itself or by selecting a literal that promises to make the 3-clause true; to fulfill its promise, the second literal b_i must either be true or select a true literal:

a_i **true**
a_i **false** — b_i **true**
a_i **false** — b_i **false** — c_i **true**

(If a_i is true, it selects the next literal b_i with any truth value.)

Given variable truth values, affixes listed in the lexicon determine literal truth values; they will negate or preserve variable truth values, according to F (see below).

Then, scanning from right to left, each literal moves to the specifier position of the closest literal of the same variable, either by long movement or by successive cyclic movement (see figures 3, 4).

In the resulting S-structure, the specifier position of the ith occurence of a variable contains the $i+1$th occurence of the same variable, and agrees with it by specifier-head agreement: all but the first occurence of a variable are contained in the specifier position of the first occurence of that variable (see figure 4).

Now, by specifier-head agreement at S-structure, all variables have consistent truth assignments, and by D-structure selection all clauses contain a true literal, where negation is performed by affixes. The formula is satisfiable if and only if the corresponding D- and S-structure are well-formed.$\square$

Requirements on the transformational model. Each formula literal is represented by a constituent with the following characteristics:

1. *Transparent to extraction.* The construction must permit successive cyclic movement (typically adjunction) between bounding nodes in order to satisfy the subjaceny condition of bounding theory.
2. *Contains a landing site that agrees with the constituent head.* The constituent will contain a specifier position; the head of the constituent will agree with the specifier position and assign case to it; the landing site will be limited to literals of the same variable by identifying agreement features on the head.
3. *Undergoes obligatory movement.* The constituent will be a properly governed argument assigned a theta-role but no case; correspondingly, it must contain a properly governed caseless position that it selects and assigns a theta-role.
4. *Selectional properties are correlated with agreement features.* Each constituent will contain an element that selects the constituent representing the next literal. Local affixation rules will morphologically merge the head of the construction with this element, thereby correlating selectional properties with agreement features in the lexicon.

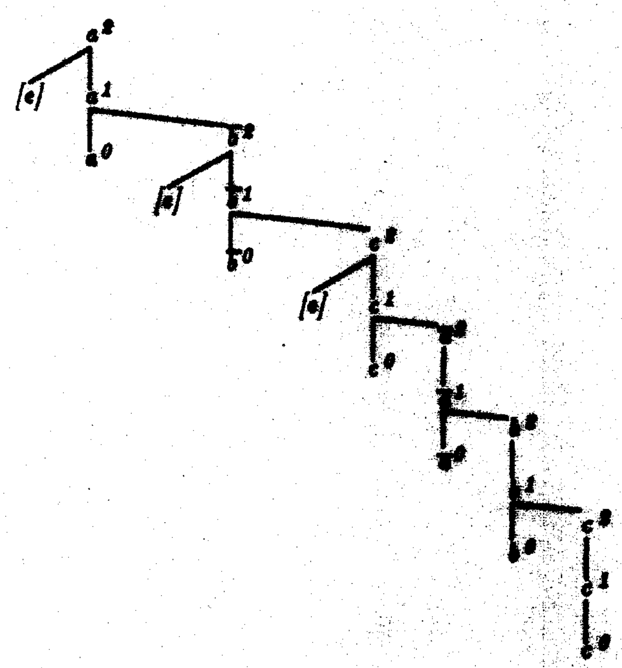


Figure 2: On the input 3SAT instance $F = (a \vee \bar{b} \vee c), (\bar{a} \vee b \vee c)$, a D-structure is created to represent F, where each literal is a constituent that selects the literal immediately to its right.

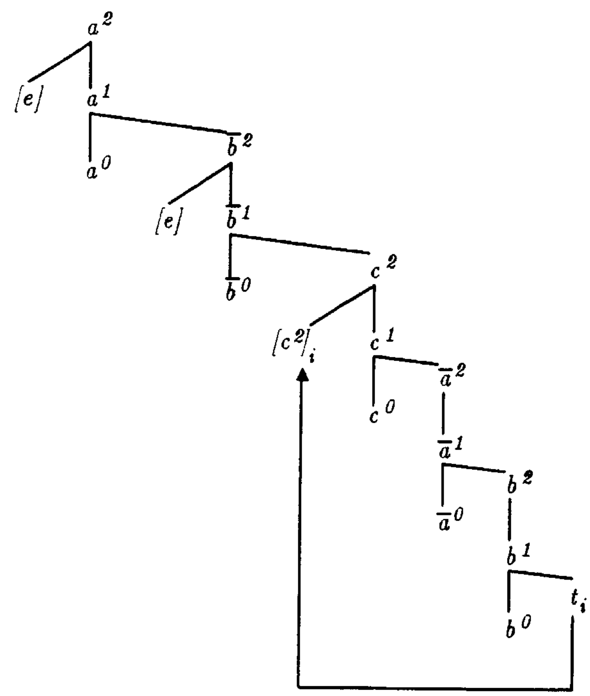


Figure 3: Then the sixth constituent (c^2) moves to the specifier position of the third constituent (c^2), leaving behind a trace t_i. This transformation relates the c^2 literal of the second clause in the 3-CNF formula to the c^2 literal of the first clause. (This example assumes constituents transparent to movement.) Now both constituents agree, by specifier-head agreement; therefore, the corresponding literals of the formula variable c will be assigned the same truth value, even though they appear in different clauses.

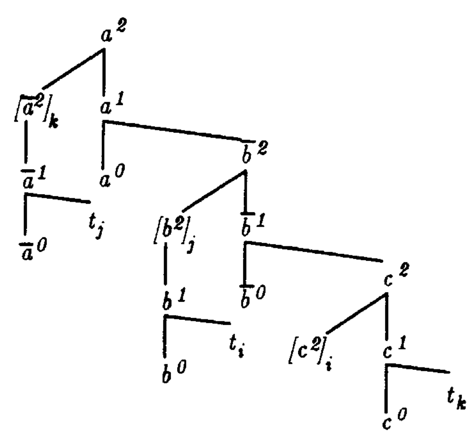


Figure 4: The end result is an S-structure where specifier-head agreement ensures that all instances of a variable receive the same truth value.

The reduction succeeds for any transformational theory that allows such a (potentially intricate) constituent. This provides additional evidence to support the speculation of Ristad (1986) that all linguistic models accounting for syntactic agreement and lexical ambiguity, as all descriptively adequate models must, will give rise to intractability.

6 Complexity of Move-α with Trace-Deletion

A central question for current transformational theories of syntax, such as the trace-deletion approaches of Lasnik and Saito (1984) and Chomsky (1986), is what are the consequences of interacting agreement relations, such as specifier-head agreement, head-head agreement, head-projection agreement, and the various forms of chain agreement (link, extension, composition)? This question is especially important because although these forms of agreement are explicated in distinct definitions, they are all performed with the same mechanism of coindexing. Thus, these models claim that there is only one form of agreement, disjunctively determined by the interaction of complex rules and principles.

In the Barriers model of Chomsky (1986), blocking categories (BCs) stop unbounded application of move-α. Informally, a BC is a category not θ-marked by a lexical X^0. For example, matrix verb phrases are BCs because they are selected by the nonlexical category I^0 (inflection) without being assigned a θ-role. Unbounded A-movement becomes possible when a category is moved local steps, adjoining to intermediate nonargument positions before moving on (adjunction is typically to BCs). As in GPSG, this notion of strict phrase structure locality does not suffice to prevent computational intractability unless local agreement (specifier-head, head-head) can be decoupled from nonlocal agreement (chains), which may be impossible.

Theorem 4 *The LRP is NP-hard in the Barriers model.*

Proof. The idea of the proof is from theorem 3 above; although the idea is simple, the actual details are quite complex. On input 3-CNF formula F, the reduction will create a lexicon L and a D-structure that represents F, then apply move-α, and finally insert underspecified lexical forms to derive an incomplete S-structure that can be completed according to the lexicon L if and only if F is satisfiable. The crux of the reduction is to represent each literal with the noun complement structure $[\ N\ [\ I\ VP\ ]]$ in (1) (for example, *desire [to visit places]*):

$$[_{NP_1} \cdots [_{NP_i}\ N\ [_{IP_i}\ [e]\ I\ [_{VP_i}\ V\ NP_{i+1}\ ]]]] \qquad (1)$$

The Barriers model endows this construction with the required characteristics:

1. *Transparent to extraction.* NP_{i+1} can be moved out of NP_i in the structure (1). VP is a BC and barrier for NP_{i+1} because it is not L-marked, but NP_{i+1} can adjoin to the nonargument VP and void its barrierhood because nonarguments may be freely adjoined to. Both NP_i and IP_i are L-marked, and therefore are neither BCs nor barriers for further NP_{i+1} raising. Thus, NP_{i+1} can be A-moved to any c-commanding specifier-of-IP position *[e]* without violating the ECP because all traces are properly governed (both θ-governed by the verb V that selects NP_{i+1}, *and* γ-marked (antecedent-governed) by the deleted trace adjoined to VP).

   Reinhart (personal communication) suggests a similar, albeit marginal, natural example where an NP containing an argument trace is topicalized to CP specifier from an L-marked position:[3]

   (7) ? [What burning $t_{i+1}]_i$ did John say [of what book$]_{i+1}$ [t_i would be magnificent]

2. *Contains a landing site that agrees with the constituent head.* The internal IP_i contains a specifier position (landing site) that will agree with I by specifier-head agreement in nonlexical categories; the specifier position will also agree with N (the constituent head), by predication. Alternately, head movement from V to I to N can create an inflected noun $[_N\ [_I\ V\ I]\ N]$ in the X^0 position of NP_i that will agree with the landing site. Although I cannot find a natural example of such an inflected noun, no arguments or analyses exclude it in principle. A close natural example is noun incorporation in Mohawk verbs (Baker 1985:139).

3. *Undergoes obligatory movement.* V assigns a θ-role but no case to the NP_{i+1} position, requiring NP_{i+1} to move. This is possible if V has lost its ability to assign case (passive morphology) or if NP_{i+1} is the underlying subject of VP_i, as in many current accounts (Sportiche 1986, Fukui 1986, Larson 1987, et cetera).

---

[3] Chomsky (pc) suggests that the correct analysis of (7) is

(6) [what burning$]_i$ did John say [[t_i of what book] would be magnificent]

and that a better topicalization example might be *What burning did John say (that) of that book, Mary thought would be magnificent.*

4. *Selectional properties are correlated with φ-features.* V undergoes obligatory head movement to the affix I, creating an inflected verb in the head of IP. As noted above, the φ-features will appear on the inflected verb by specifier-head agreement, where they may be systematically correlated with the verb's selectional properties in the lexicon.

Details. The input to the reduction is a 3-CNF formula

$$f = (u_{11}u_{12}u_{13}), (u_{21}u_{22}u_{23}), \ldots, (u_{n1}u_{n2}u_{n3})$$

with n clauses, $3n$ literals, and variables $q_1 \ldots q_m$. As before, the reduction will create a D-structure that represents F, and then apply move-α and lexical insertion to derive the S-structure S. The reduction output will consist of an S-structure S containing underspecified words and a lexicon L, such that the words in S can be resolved according to L if and only if F is satisfiable.

The reduction algorithm consists of five steps:

1. *Create the two φ-features* `VBL` *and* `TRUE`

   The m-ary `VBL` feature identifies the formula variables: the formula variable q_k is associated with the feature `[VBL` k`]`. The final binary φ-feature `TRUE` denotes the truth assignment to the given variable.

2. *Create the D-structure representation of f.*

   For each literal u_{ij} in the i^{th} clause, create the D-structure noun complement construction:

   (8) $[_{NP_{ij}}\ N_{ij}\ [_{IP_{ij}}\ [e_{ij}]\ I_{ij}\ [_{VP_{ij}}\ V_{ij}\ NP_{ij+1}\ ]]\,]$

   where NP_{ij} represents the subformula $u_{ij} \ldots u_{n3}$; NP_{ij+1} represents the subformula $u_{ij+1} \ldots u_{n3}$; and N^0_{ij} bears the φ-features of the variable q_k that corresponds to the literal u_{ij}.

3. *Apply head-movement within noun complements.*

   Move V_{ij} to I_{ij}, and then move the inflected verb $[\mathrm{V}_{ij}\ \mathrm{I}_{ij}]$ to N_{ij}:

   (9) $[_{NP_{ij}}\ N_{ij}\ [_{IP_{ij}}\ [e_{ij}]\ I_{ij}\ [_{VP_{ij}}\ V_{ij}\ NP_{ij+1}\ ]]\,]$

(10) $[_{NP_{ij}} [[V\ I]\ N_{ij}]\ [_{IP_{ij}}\ [e_{ij}]\ t\ [_{VP_{ij}}\ t\ NP_{ij+1}\ ]]\ ]$

As a result, all X^0 positions other than N^0 positions have traces in them; N^0 positions are filled with inflected nouns of the form [[V I] N]. (The goal of this operation is to force agreement between the verb and noun; this could also be done with predication.)

4. *Apply long distance NP-movement across noun complements.*

   Starting with the rightmost, innermost NP (NP_{n3}) and scanning left and up, move NP_{ij} to the specifier position $[e_{lm}]$ of the closest inflection I_{lm} that agrees with it in φ-features:

(11) $\ldots [_{NP_{lm}} [[V\ I]\ N_{lm}]\ [_{IP_{lm}}\ [e_{lm}]\ t\ [_{VP_{lm}}\ t\ [_{NP_{lm+1}} \cdots [_{NP_{ij}} [[V\ I]\ N_{ij}]\ [_{IP_{ij}}\ [e_{ij}]\ t\ [_{VP_{ij}}\ t\ NP_{ij+1}\ ]]\ ]]\ ]]\ ]$

(12) $\ldots [_{NP_{lm}} [[V\ I]\ N_{lm}]\ [_{IP_{lm}}\ [NP_{ij}]\ t\ [_{VP_{lm}}\ t\ [_{NP_{lm+1}} \cdots t_{ij}\ ]\ ]]\ ]$

The leftmost literal of each formula variable is assigned case in situ by the formula literal that selects it; all other literals of the variable receive a theta role but no case, which is what forces them to move. All literals of a variable but the rightmost one assign case to their specifier position, so that a literal will be able to land there. Our third ally in making the movement obligatory is the extended projection principle, which requires specifier of IP to be filled.

Movement is by adjunction to intermediate VPs, deleting intermediate traces. This movement satisfies the ECP because all traces are properly governed and therefore γ-marked at S-structure (both θ-governed by the verb V_{ij} that selects them, *and* antecedent-governed by the deleted trace adjoined to VP). The movement satisfies minimality trivially because the deleted trace is not excluded by VP. Because NPs are only moved to specifier positions that c-command their bound traces, the resulting structure also satisfies the chain condition and binding theory when the chain is created. Note that the movement is permissible independent of whether we express the antecedence relation induced by move-α using indices or links (cf. Higginbotham, 1983).

Now NP_{ij} agrees with its c-commanding NP_{lm} on the **TRUE** feature by specifier-head agreement in IP and head-head agreement between I_{lm} and the head of NP_{lm}.

5. *Perform (underspecified) lexical insertion.*

   Insert the morphological form $\alpha_j \beta_j \gamma \varphi_k$ for the inflected noun $[[\mathrm{V\ I}]\ \mathrm{N}_{ij}]$ when $u_{ij} = q_k$ (that is, u_{ij} is not a negated variable); when $u_{ij} = \overline{q_k}$, insert the morphological form $\alpha_j \beta_j \overline{\gamma} \varphi_k$ instead.

   (13) $\ldots [_{NP_{lm}}\ [[V\ I]\ N_{lm}]\ [_{IP_{lm}}\ [NP_{ij}]\ t\ [_{VP_{lm}}\ t\ NP_{lm+1}\ ]]\ ]$

   (14) $\ldots [_{NP_{lm}}\ [\alpha_l \beta_l \gamma \varphi_k]\ [_{IP_{lm}}\ [\alpha_i \beta_i \overline{\gamma} \varphi_t]\ t\ [_{VP_{lm}}\ t\ NP_{lm+1}\ ]]\ ]$

   Note that each '$\alpha_j \beta_j$' is a listed word; that γ and $\overline{\gamma}$ are suffixes; and that φ_k is the morphological realization of the φ-feature `[VBL` k`]`.

   Further note that the surface structure is a string of inflected words. For example, on input $f = q_1 \overline{q_2} q_3 q_2 \overline{q_1} q_3$ the reduction algorithm would yield the surface string or linguistic expression:

   $$\alpha_1\beta_1\gamma\varphi_1\ \alpha_2\beta_2\overline{\gamma}\varphi_1\ \alpha_2\beta_2\overline{\gamma}\varphi_2\ \alpha_1\beta_1\gamma\varphi_2\ \alpha_3\beta_3\gamma\varphi_3\ \alpha_3\beta_3\gamma\varphi_3$$

We now turn to morphological and lexical details of the reduction, namely how we ensure that each clause has one true literal, that literals and variables may be true or false, and that negation changes truth values.

The *lexicon* contains ambiguous inflected nouns *[[V I] N]* that have undergone verbal incorporation. We define the morphological primitives and lexicon to ensure that the N-complement structure NP_{i1} for the first literal in the i^{th}clause u_{i1} either (a) corresponds to a true literal, or (b) selects a true literal (as represented by) NP_{i2}, or (c) selects a false literal NP_{i2} that will select a true literal NP_{i3}. Similarly, the N-complement structure NP_{i2} for the second literal in the i^{th}clause u_{i2} must either (a) correspond to a true literal, (b) correspond to a false literal, or (c) correspond to a false literal and select a true literal NP_{i3}. Finally, the structure NP_{i3} for u_{i3}, the third and final literal in the i^{th}clause, may correspond to a true or false literal.

This lexical system requires one φ-feature `TRUE` on nouns to encode variable truth assignments; one non-φ-feature `LITERAL` to encode literal truth values on both nouns and verbs; and one non-φ-feature `T` on nouns that verbs can select for. Recall that we are assuming that the m-ary φ-feature `VBL` will be morphologically realized on the nouns in the lexicon, and will distinguish the formula variables: N_{ij} and N_{lm} bear the same `VBL` value if and only if u_{ij} and u_{lm} are literals of the same variable.

The *morphological system* is constructed assuming the relativized head definition of DiScuillo and Williams (1987). There are six ambiguous morphological primitives: the prefixes $\alpha_1, \alpha_2, \alpha_3$ and roots $\beta_1, \beta_2, \beta_3$. Together they define three ambiguous listed words $\alpha_1\beta_1$, $\alpha_2\beta_2$, and $\alpha_3\beta_3$, which may each combine with either of the two nominal suffixes $\gamma, \overline{\gamma}$.

Listed Word	Prefix (α_i) features	Root (β_i) features
$\alpha_1\beta_1$	N[LITERAL 1]	V[LITERAL 1]
	V[LITERAL 0,Select T]	N[LITERAL 0]
$\alpha_2\beta_2$	N[LITERAL 1,T]	V[LITERAL 1]
	V[LITERAL 0]	N[LITERAL 0]
	V[LITERAL 0,Select T]	N[LITERAL 0,T]
$\alpha_3\beta_3$	N[LITERAL 1,T]	V[LITERAL 1]
	V[LITERAL 0]	N[LITERAL 0]

The suffixes relate variable truth assignments to literal truth values. For example, $\overline{\gamma}$ will only be lexically inserted to represent a negated variable, and therefore inversely relates the truth value of the variable and literal.

Suffix	Features	Selects
γ	N[TRUE 1]	[LITERAL 1]
	N[TRUE 0]	[LITERAL 0]
$\overline{\gamma}$	N[TRUE 1]	[LITERAL 0]
	N[TRUE 0]	[LITERAL 1]

Finally, there are m listed inflectional affixes corresponding to the m possible values of the VBL φ-feature. This completes the presentation of the reduction algorithm.$\square$

Comments. The proof exploits a flaw in the A-system of local θ-relations described in Barriers that arises because distinct linguistic relations (head chains, NP-chains, and specifier-head agreement) are conflated into the relation of A-chain coindexing. Abstracting from the details of the preceding reduction, it is very clear that complex, undesirable interactions can arise when all agreement is performed by coindexing, as it is in both trace-deletion models considered above. Linguistically, this corresponds to conflation of 'local' relations, such as specifier-head agreement, and 'nonlocal' relations, such as long A-movement.

The lexicon in the reduction above strongly correlates φ-features with selectional restrictions. In effect, words can indirectly select the selectional restrictions of their complements. In my view, this is perfectly reasonable: it is well-known that the properties of words in natural lexicons are not statistically independent, and statistical correlation is perfectly natural, even expected, in any listing of idiosyncratic properties (that is, a lexicon). For example, it is possible to imagine a natural language where the φ-features on the subject are correlated with the object's φ-features. In brief, the syntactic and lexical models permit this kind of correlation, it may be used in complexity proofs.

The most serious question is whether the complexity of the lexical resolution problem accurately reflects on the complexity of language perception as a whole. The S-structure built by the reduction is very complex and I have not established that it is uniquely determined by its phonological form. If not uniquely determined, then hearers would be free to perceive some simpler S-structure with the same phonological form as the intractable S-structure built by the reduction. If this were possible, then the lexical resolution problem would no longer be a realistic subproblem of linguistic perception.

Theorem 5 *The LRP is NP-hard in the Lasnik-Saito model.*

Proof. The preceding proof proceeds without alteration in the Lasnik-Saito model because in that model, θ-government suffices for proper government, and traces may be deleted at any level after γ-marking. $\square$

6.1 Blocking the Barriers reduction is difficult

How might we change the Barriers model in order to block the preceding reduction? The Barriers model appears to crucially rely on the assumed atomicity of the two nonlexical categories C and I, which together form a barrier to argument movement. If C assigns nominative case to specifier-of-IP, as some authors have proposed (Bennis 1980, Dasgupta 1985), then CP and IP form an atomic unit, a barrier to antecedent government. Although this would stop the exact construction used in the reduction, it would also leave the structure of the infinitival noun complement construction unexplained. A second, theory-internal, problem with this change is that now both C and I must govern the specifier of IP (C in order to assign case, and I for specifier-head agreement), which violates the minimality condition.

The preceding proof relies on long movement of the NP complement of a verb (in a noun complement construction), which is precisely what Barriers strives to prevent by reducing proper government to antecedent government, using the Lasnik-Saito γ-marking mechanism. (The commitment to eliminating θ-government from proper government is tentative at best. The strongest position taken is "Possibly, a verb does not properly govern its θ-marked complement," p.79). In the Barriers reduction, an argument undergoes long movement by adjoining the argument NP to VP, γ-marking its trace, and then deleting the intermediate $\overline{\text{A}}$-trace at LF.

This is the exact derivational sequence (adjoin, γ-mark, delete adjoined trace) used in Barriers (pp.21-22) to move a wh-subject from a θ-marked CP complement to a specifier of CP, provided the wh- phrase is licensed at LF. Barriers attempts to exclude similar long movement of an NP from a similar (but caseless) subject position by appeal to Binding Theory condition C at S-structure: the NP trace in subject position would be an $\overline{\text{A}}$-bound R-expression A-bound in the domain of the head its chain (p.93, fn.20). (Barriers differentiates the two constructions solely by the nature of their traces: wh- traces are not R-expressions, while NP-traces are.) Crucially, Chomsky's argument holds only if the trace deleting ability of affect-alpha is restricted to LF, contrary to the Lasnik-Saito model: otherwise, adjoined traces could be deleted before causing an S-structure binding violation.

But trace deletion cannot be restricted solely to LF. If it were, then any ECP violation created by LF-movement may be avoided, simply by deleting offending intermediate traces after they have done their γ-marking duty. This can be done because adjoined $\overline{\text{A}}$-traces are not required by the extended projection principle at LF.

Even if this conundrum could be resolved, another one awaits us: the long movement used in the proof is applied cyclically, so that the trace of the argument NP is no longer c-commanded by the argument NP once all movement has applied, and hence not A-bound by the head of its chain at S-structure. This brazen violation of the c-command condition on chain links is standard in the literature, and hence does not raise any special problems here. Examples include the topicalization example (7) above, antecedent-contained ellipsis,

(15) [Everyone that Max wants to $e_2]_1$ John will [kiss $e_1]_2$

and passive VP topicalization in English:

(16) $[_{\text{VP}}$ Arrested t_i by the FBI$]_j$ John$_i$ has never been $t_j]$

Furthermore, even if trace deletion were disallowed entirely, long movement would still be possible from θ-marked noun complements, and the proof would proceed, because θ-government cannot be eliminated without disastrous consequences in the rest of the theory. Proper government can be reduced to antecedent government only if antecedent government suffices for NP-movement (eg. passive and raising) in accordance with the chain extension operation. This fails because only the terminus of an (extended) A-chain may theta-mark or case-mark, in order to obtain the CED effect (Condition on Extraction Domains, see Barriers, p.72). Therefore, in passive constructions, where the A-chain headed by the subject NP must be extended to include the verb and inflection and thereby achieve antecedent government of the NP-trace at S-structure, the inflection will simultaneously loose its ability to case-mark the subject position. The direct consequence is that Barriers incorrectly predicts both passives in (17) violate the case filter and are ungrammatical, whereas only (17a) is ill-formed.

(17)
a. * [e] was killed John
b. John was killed t

In short, the chain extension required to satisfy the ECP without theta-government will prevent the subject NP from receiving case, and thereby violate the case filter. This very serious problem may be remedied by abandoning either (1) the case filter, which would without question be disastrous for the theory, (2) the Barriers analysis of CED effects, which would reduce empirical coverage, or (3) the coindexing/chain extension analysis of NP-movement, which will have the direct consequence that proper government cannot be reduced to antecedent government.

Additional evidence against the coindexing/chain extension analysis of NP-movement comes from gapping and VP-deletion, which raises a contradiction for the Barriers chain extension account of simple passives: (1) each passive conjunct in (18) contains a distinct R-expression subject that therefore must head distinct A-chains with differing indices, but (2) both extended A-chains share the same verb, and therefore must share the same index.

(18)
a. John$_i$ was given t_i records and Sue$_j$ [e] books
b. John$_i$ was [killed t_i]$_k$ but Tom$_j$ wasn't [e]$_k$

The possibility of long argument movement by adjunction to intermediate positions remains in Chomsky's most recent theory of derivation, where he proposes that derivations be subject to a 'least effort principle,' with the following provisions. LF permits only the following five elements: arguments, adjuncts, lexical elements, predicates, and operator–variable constructions. Affect-alpha must apply at LF to each illegitimate object to yield one of these five legitimate elements. Chomsky (1988b:20) urges us to "consider successive-cyclic A-bar movement from an argument position. This will yield a chain that is not a legitimate object, and that can become a legitimate object, namely an operator-variable construction, only by eliminating intermediate A-bar traces. We conclude, then, that these must be deleted by LF." A parallel consequence of this theory of derivations is that successive-cyclic A-bar movement from a theta-marked argument position *to a case-marked argument position* will also yield an illegitimate object, that can become a legitimate object, namely an A-chain, only by eliminating intermediate A-bar traces by LF (that is, *before* LF).

7 Controlling complexity in systems of knowledge

The construction used in the preceding reductions may appear to arise more from inconsistancies in the transformational model than from human language. But, as I argued above, no simple change to Barriers can block the reduction while maintaining the theory's intended empirical coverage, because every property of the reduction construction is independently justified in the theory. More accurately then, the reduction construction is simply a complex, unnatural combination of the properties of actual natural constructions.

In order to explain complex phenomenon, generative linguists invariably construct complex theories that are difficult to change and computationally intractable. In my opinion, this problem is inherent in the attempt to construct a system of knowledge. If you know about long movement, agreement, and head movement, then you also know all possible combinations of those linguistic processes. Nor is a cost assigned to employing those combinations or determining the consequences of knowledge. Theories of knowledge are inherently non-modular and maximally interactive—they lack techniques for controlling complexity—and this is why it is so difficult to construct a complex system of knowledge. This is the same critical flaw that doomed cybernetics.

7.1 Taming the complexity of simple operations

Technically, the preceding proofs have shown that simple local linguistic operations in transformational grammars (including local agreement and move-α bounded by the ECP) can have complex consequences. How might the complexity of simple linguistic operations be reduced? Interactions among operations must be tamed. As things currently stand, this will be difficult to do for the transformational models considered above because many conceptually distinct relations (forms of syntactic agreement) are uniformly represented via coindexing, and this maximizes interactions among principles. Therefore, the first step is to segregate distinct relations in representation. This approach has led phonologists from the solely segmental representations of SPE to nonlinear models (autosegmental, prosodic). In syntax, the potentially distinct anaphora-antecedent, chain link, and specifier-head agreement relations would be best represented by distinct relations, rather than by the uniform co-indexing mechanism. Representational theories of movement (eg., g-projection, dynasty formation, chain formation algorithm) might strive to decouple principles on the verge of orthogonality, such as binding and government in Rizzi's relativized minimality approach.

The second step is to more clearly distinguish linguistic representations from the abstract process whereby they are specified. This necessity of this conceptual distinction has been clear since Descartes first argued for it in introductory philosophy. In generalized phrase structure grammar, the local tree representation is sharply distinguished from the process of specifying the local trees (metarule finite closure, and ID rule projection according to universal feature instantiation, et cetera). The distinction is similarly clear in representational approaches to syntactic movement: chain link representations are specified by an abstract process of iteration (g-projection, dynasty formation, or chain formation algorithm) whose intermediate states are distinct from its output. But in transformational approaches, the fundamental distinction between representation and process is blurred, because the chain link representation is simply the history of the process of applying the move-α transformation. In such a model, it is unclear if the difficulty of constructing the linguistic representation is inherent, or merely an artifact of an unnecessarily complex specification process.

Neither representational segregation nor a sharper distinction between process and representation entails a reduction in the complexity of simple principles. But both steps would certainly clarify the fundamental computational structure of linguistic theory; hopefully, this clarification will lead to an improved understanding of the inherent complexity of simple linguistic principles.

8 Syntactic Binding

Every child knows that pronouns stand for nouns, and that speech would soon become extremely difficult without them. We would soon grow tired of repeating the nouns pronouns stand for over and over. Nor could we easily introduce ourselves to fellow travelers and proceed to discuss our respective destinations.

Consider the problem of determining pronoun antecedents. Two pronouns can share an antecedent only if they agree and they are not too close. Thus, the pronoun *her* cannot stand for *Bill* in the sentence *Bill saw her mother* because *her* is not masculine gender. And the sentence *Bill knew he liked him* cannot mean 'Bill knew Bill liked Bill,' unlike the sentence *Bill knew he liked himself*, which could mean that. In effect, pronouns are competing for antecedents: in some cases, two pronouns may share an antecedent, and in the other 'disjoint reference' cases they may not. Exactly how hard is this task?

The interactions among pronouns can become impossibly complex when there are many pronouns and only a few possible antecedents. In fact, determining pronoun antecedents is like coloring the nodes of a graph, where possible antecedents are colors and pronouns are the nodes of a graph whose edges represent disjoint reference conditions. This correspondence proves that the pronoun antecedent problem, like the NP-complete graph coloring problem, cannot be solved in practice by any known method. How then can we understand sentences containing pronouns?

8.1 Binding at S-structure

Binding theory is concerned with the permissible linguistic antecedents of anaphora, such as pronouns, reflexives, and reciprocals. Binding theory states conditions on binding relations between A-positions (positions assigned a grammatical function, such as SUBJECT or OBJECT) at S-structure and LF. Chomsky (1981:188) formulates the binding theory as follows:

(19)
a. An anaphor is bound in its governing category.
b. A pronominal is free in its governing category.
c. An R-expression is free everywhere.

(Arguments are bound by c-commanding antecedents, else free; a governing category for α is the minimal NP or S containing α and a governor of α.) Two additional binding conditions are (1) anaphora must find their antecedents in some level of linguistic representation, such as S-structure, LF, or discourse structure, and (2) anaphora that share an antecedent must agree with each other and with that antecedent on agreement features.[4] Occasionally, a third condition, the *i-within-i condition*, is cited to prohibit categories from being antecedents of anaphora they dominate.

A *syntactic binding* is a set of anaphor-antecedent relations that satisfies the binding theory, such that all anaphora have S-structure antecedents. The binding problem is:

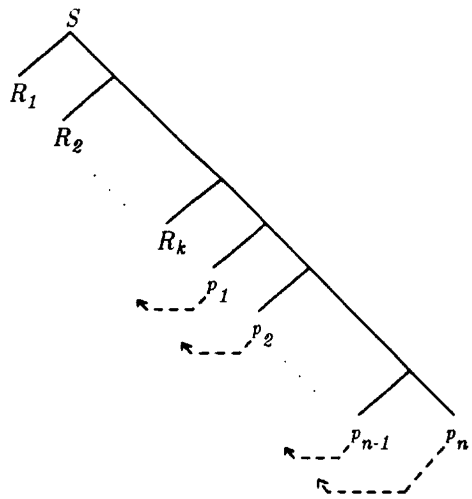
> **Binding Problem (BINDING)**
> Given a syntactic representation S lacking only binding relations, construct a permissible syntactic binding for S, according to binding theory.

The corresponding decision problem is to decide if a given syntactic representation has a permissible syntactic binding. The *all bindings problem* is to enumerate all permissible syntactic bindings for a given syntactic representation.

Theorem 6 *The binding decision problem is NP-hard.*

Proof. The proof will be by a trivial reduction from GRAPH k-COLORABILITY to BINDING. On input graph $G = (V, E)$ with vertices $v_1 \ldots v_n$, we need n binary agreement features, n pronouns, and k R-expressions. Pronoun p_i represents vertex v_i: for each edge (v_i, v_j) from v_i, pronoun p_i has $\varphi_i = 0$ and pronoun p_j has $\varphi_i = 1$. In the S-structure, the ith R-expression R_i c-commands the $i+1$th R-expression R_{i+1}, and each pronoun is free in its governing category, where it is c-commanded by all R-expressions, which are its potential antecedents.

[4]The transitivity of agreement is trivially demonstrated by the ungrammaticality of *Every student$_i$ prepared her$_i$ plan and did his$_i$ homework*, where each pronoun agrees with its operator antecedent, but the two pronouns disagree with each other, and though neither pronoun is the other's antecedent, the fact that they disagree through their common antecedent nonetheless excludes the example.



Each pronoun must ultimately have an R-expression antecedent (to satisfy the binding theory), which will c-command it (by construction). All R-expressions have disjoint reference by binding condition C, and therefore there are exactly k distinct possible antecedents. If there is an edge between two vertices in the graph, then the two corresponding pronouns cannot share an antecedent without disagreeing on some agreement feature; there are only k possible antecedents (the R-expressions), and therefore each permissible binding for the S-structure corresponds to a k-coloring of the graph G.$\square$

Comments. Crucially, the reduction proves that it is the combination of disagreement, binding conditions B and C, and prohibitions against circular dependencies that creates inherently complex disjoint reference graphs. (The disjoint reference graph in the preceding reduction was created entirely via feature disagreement simply because that was the simplest reduction.)

Although the reduction's extensive use of agreement features does not increase our understanding of natural language, it does offer greater insight into linguistic theory. It should be clear that agreement features, a much neglected part of linguistic theory, are surprisingly powerful. A reduction based on the other disjoint reference conditions (binding conditions B and C, and i-within-i), with perhaps a constant number of agreement features, might offer better insights into binding phenomenon.

The other weakness is that the reduction requires an unnaturally arbitrary specification of agreement features on pronouns (see below). Functionally, pronouns and anaphors serve to simplify linguistic expressions, and therefore one would expect, contrary to the reduction construction, that pronouns and anaphors specified for different agreement features would also have different antecedents. But this weakness is ultimately attributable to a failure of current linguistic theory to view anaphora from this perspective.

The great strength of this reduction is that it is based directly on a natural syntactic construction, whose pronouns must find antecedents in order for the construction to be understood. Therefore, the binding problem is a necessary subproblem of language comprehension. It also raises a significant computational question for the binding theory, namely, is binding really so much like graph coloring, a problem which is hyper-exponentially ambiguous and cannot be efficiently performed by any known method, even approximately (Lawler, 1976)?

8.2 Limiting disagreement with the transparency condition

The most unnatural characteristic of the binding construction used in the preceding reduction from graph coloring is that two pronouns with entirely different features could nonetheless share an antecedent, provided that they did not disagree on any common feature. That is, two pronouns could share an antecedent if their agreement features could be unified: they need not agree exactly, but they must not have different values for the same feature. In natural language, it seems that an argument may bind anaphora that are virtually identical, modulo case and affix (for example, the anaphoric suffix *-self*). Let us say that two anaphora are *transparently identical* if they have the same root morpheme and agree exactly on agreement features, although they may have different cases or affixes. This relation partitions anaphora into equivalence classes, which we may call transparency classes (t-classes). For example, *he*, *him*, and *himself* are transparently identical, and form one t-class in the English lexicon.

In light of these observations, I propose an additional binding condition, which limits interactions among t-identical anaphora.

> **Transparency condition**
> All anaphora with the same antecedent must be t-identical.

The transparency condition explains why, in some dialects of English, a masculine singular antecedent can independently bind the pronouns *his* and *their*, but cannot simultaneously bind both pronouns in the same structure:

(20) a. Every boy$_i$ did his$_i$ homework.

b. Every boy$_i$ did their$_i$ homework.

c. * Every boy$_i$ did his$_i$ homework and prepared their$_i$ lesson plan.

(21) a. Sue saw the dog$_i$ when it$_i$ barked.

b. Sue saw the dog$_i$ after he$_i$ was kicked.

c. * Sue saw the dog$_i$ when it$_i$ barked after he$_i$ was kicked.

(Note how poor the second example is with the pronouns *he* and *it* interchanged.) The transparency condition can also explain why disjoint reference conditions only appear to hold within a t-class:

(22) Bill$_i$ reminded Sue$_j$ that [$_{\mathrm{gc}(their)}$ he$_i$ paid for their$_{\{i,j\}}$ tickets]

A compound noun may be the antecedent of two t-classes, in apparent violation of the transparency condition. (Of course, it is always possible to make up a story where the compound noun posses two indices, as in *[baby brother$_j$]$_i$*.)

(23) Suzie's [baby brother]$_i$ cried when she left it$_i$, but soon thereafter he$_i$ went to sleep.

Unfortunately, although this transparency condition successfully prevents the particular details of the preceding reduction, it does not in general suffice to limit binding interactions. Now t-classes, rather than individual pronouns, compete for antecedents in the presense of disagreement conditions. In the next section I will present a pair of binding constructions that satisfy the transparency condition, and then use those constructions to simulate QUANTIFIED 3SAT. Moreover, the disjoint reference graph arising from binding conditions B and C remains unaffected.

9 Binding and Copying in Syntax

Language users inwardly relate a complete meaning to every sound uttered as language: that is, for a language user, every sound is a complete utterance. This gives language tremendous expressive power, especially when taken in combination with language's intricate underlying structure.

For example, the sentence *the men ate but the women didn't* is understood to mean 'the men ate but the women didn't eat.' In so-called VP-deletion constructions, the second verb phrase is an invisible copy of the first verb phrase: understood, but not spoken. If the verb phrase contains a pronoun, that pronoun may be understood as though it were also copied, because the pronoun and its copy may have different antecedents. For example, the sentence *the men ate their dinner but the women didn't* can mean 'the men ate the men's dinner but the women didn't eat the women's dinner.'

If the twin subjects of a VP-deletion construction disagreed on some feature (gender in the preceding examples) and the pronoun in the deleted verb phrase represented a variable in a Boolean formula, then such a so-called sloppy identity construction would represent a universal quantifier over that feature: the construction would be permissible only if the pronoun could could be independently bound by both subjects. This is possible only if the pronoun can agree with both (contrary) feature settings.

A slightly different binding construction mimics an existential quantifier. In the sentence *the men knew the women ate their dinner*, either subject may bind the pronoun *their*. Continuing with the idea of pronouns as variables and antecedents as truth values, this binding construction represents a existential quantifier over some feature, where the pronoun can be freely bound by either of the two antecedents that disagree on that feature.

A third binding construction can represent 3SAT instances. The sentence *he introduced him to him* is necessarily about three people: no two of the three pronouns can share an antecedent. Similarly, at least one of the three pronouns in the sentence *John thought Peter said he introduced him to him* must refer to a third, unnamed male person. This construction can be used to mimic a 3-CNF clause, where at least one of the pronouns must be bound outside the construction, to an antecedent representing a quantified variable. This pronoun represents a true literal in the 3-CNF clause.

Finally, if the VP-deletion and binding constructions are put together properly, they would represent quantified Boolean formulas. Therefore, determining the meaning of such linguistic utterances can be as difficult as playing board games, such as Checkers, Chess, and Go.

9.1 Linking and Copying at LF

Every linguistic account of sloppy identity in VP-deletion constructions requires that the overt VP in effect be copied to the empty VP position, and that pronominal binding occur both before and after that copying is performed. Some representative accounts include: copying or 'reconstruction' of the VP directly at S-structure (Koster, 1987; Ross, 1967); copying in the S-structure to LF mapping (May, 1985); copying and interpretation in syntax (Wasow, 1972); or in two stages—copying an empty VP structure at S-structure and interpreting the antecedent VP at LF, in Sentence Grammar, and then copying the interpreted antecedent VP after LF, in Discourse Grammar (Williams, 1977).[5] Although these theories differ in the exact formulation of the copying rule, at what level of representation the rule applies, and the rule's name, real differences are extremely minor, and do not affect the subsequent discussion or the proof at all.

Without loss of generality, I will restrict my attention to binding and copying at LF in the T-model (for example, May 1985). In the T-model anaphora-antecedent relations may be determined both at S-structure and LF, while

[5] Williams (1977) proposes that Sentence Grammar contains an optional Pronoun Rule that replaces pronouns inside a VP with variables bound by lambda operators whose scope is the entire VP. For example, *John saw his mother* is assigned the LF *[John [λx (x saw [x's mother])]]* when the Pronoun Rule applies, and *[John [λx (x saw [his mother])]]* when it does not. (Note that Williams claims both LF representations have the same meaning, but his failure to explicitly relate the pronoun *his* to its antecedent in the second LF representation is incoherent.) Discourse Grammar strictly follows Sentence Grammar, and contains a VP Rule that copies entire VP-lambda expressions. Sloppy identity readings occur when VP copying applies to the output of the Pronoun Rule; alternate readings occur when the optional Pronoun Rule does not apply. Therefore, Williams' account would seem to preclude free pronoun binding after VP copying. But, by Williams' own arguments, the Pronoun Rule must belong to Discourse Grammar because it operates across sentence boundaries:

(24) a. Frank loves Sue.
b. No, he only thinks he loves her.

Therefore, Williams' framework must also allow pronoun binding to occur both before and after VP copying, contrary to his claims.

constituent copying, as required for VP-deletion constructions, is performed in the S-structure to LF mapping. For example, (25) can mean either that the women ate the women's dinner, or that the women ate the men's dinner (see 26).

(25) The men didn't eat their dinner but the women did

(26)
a. The men$_i$ didn't eat their$_i$ dinner but the women$_j$ ate their$_i$ dinner
b. The men$_i$ didn't eat their$_i$ dinner but the women$_j$ ate their$_j$ dinner

In example (26a), binding is performed *before* copying, while in (26b), binding is performed *after* copying. Conversely, example (27) does not have a reading where some women eat some man's dinner.

(27) Which men didn't eat their dinner and which women did

This is explained by the binding theory, which states that pronouns may only link to A-positions, and the LF binding conditions, which require each operator to c-command its variable and all anaphora linked to that variable. Therefore, in this VP-deletion construction, anaphora can only be bound by the subject of the VP that contains them at LF.

I now make crucial use of sloppy identity in VP-deletion constructions to prove that the S-structure to LF mapping is extremely difficult. Informally, if we embed some VP-deletion constructions, and hide a parameterized computation in the innermost VP, we might then have to perform that computation differently an exponential number of times.

The reduction is clearest when presented assuming the linking theory of Higginbotham (1983), which replaces the indices of the standard binding theory with directed edges (links) from anaphor to antecedent, and the stipulatory i-within-i condition with a more natural (semantic) prohibition against circular dependencies. (An element depends on all elements it dominates, and on all elements its antecedents depend on.)

Theorem 7 *The S-structure to LF mapping is PSPACE-hard.*

Proof. By reduction from QUANTIFIED 3SAT. The input Ω is a quantified Boolean formula in prenex 3-CNF, consisting of alternating quantifiers $\forall x_1 \exists x_2 \ldots \forall x_{n-1} \exists x_n$ preceding (and quantifying the literals in) the clauses $C_1, C_2, \ldots, C_p$ in the Boolean variables $x_1, x_2, \ldots, x_n$. Each clause contains exactly three distinct literals labeled by $C_i = (a_i \vee b_i \vee c_i)$. The output is an S-structure S such that S has a permissible LF representation iff Ω is true.

Each quantifier ($\forall x_i$ or $\exists x_i$) is represented by a S-structure construction $\mathbf{QC}_i$, where antecedents in A-positions represent the quantified variables. Each Boolean clause C_j is represented by an S-structure construction $\mathbf{CC}_j$, where pronouns represent formula literals, and at least one pronoun in each $\mathbf{CC}_j$ must link to some c-commanding A-position in one of the dominating $\mathbf{QC}_i$ constructions. Thus, the entire QUANTIFIED 3SAT instance is represented by an S-structure composed of $\mathbf{QC}$ and $\mathbf{CC}$ constructions: each $\mathbf{QC}_i$ immediately contains $\mathbf{QC}_{i+1}$; $\mathbf{QC}_n$ immediately contains $\mathbf{CC}_1$; and each $\mathbf{CC}_j$ immediately contains $\mathbf{CC}_{j+1}$, as shown in figure 5.

A universal quantifier is represented by a VP-deletion construction. In the VP-deletion construction $\mathbf{QC}_i$, the disagreeing operator-variable chains (Op_i, $[\mathrm{e}]_i$) and ($\overline{\mathrm{Op}}_i$, $\overline{[\mathrm{e}]}_i$) represent the quantifier determined truth assignments $x_i = 1$ and $x_i = 0$, respectively. S-structure linking conditions require pronouns to only link to A-positions (in this case, to the syntactic variables in subject position). Feature agreement on the `ID` feature requires pronouns representing literals of x_i to link to either $[\mathrm{e}]_i$ or $\overline{[\mathrm{e}]}_i$, and never to another A-position in another $\mathbf{QC}$ construction. Recall that linking conditions at LF require each operator to c-command its syntactic variable and all anaphora linked to that syntactic variable. Therefore, because each variable in subject position is locally bound by an operator, pronouns contained in the original VP must link to the syntactic variable $[\mathrm{e}]_i$ linked to Op_i; these pronouns will correspond to unnegated literals, which are true iff x_i is true. Similarly, pronouns in the copied VP must link to the syntactic variable $\overline{[\mathrm{e}]}_i$ linked to $\overline{\mathrm{Op}}_i$; these pronouns will correspond to negated literals, which are true iff x_i is false.

An existential quantifier is represented by a linking ambiguity. In the linking ambiguity construction $\mathbf{QC}_{i+1}$, the R-expressions NP_{i+1} and $\overline{\mathrm{NP}}_{i+1}$ represent the two possible truth assignments $x_{i+1} = 1$ and $x_{i+1} = 0$, respectively. Structural conditions (c-command) and feature agreement on the `ID` feature force the pronoun p^*_{i+1} to link to either NP_{i+1} or $\overline{\mathrm{NP}}_{i+1}$. The pronoun p^*_{i+1} bears an additional feature, indicated by '*', that disagrees with a feature

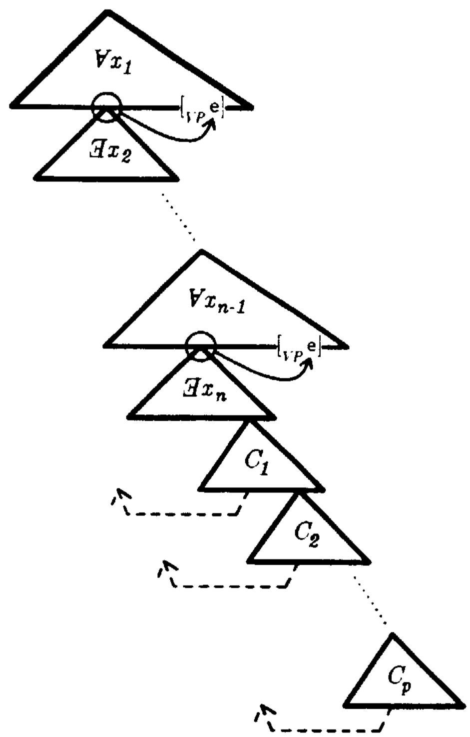


Figure 5: The S-structure in the figure represents the QUANTIFIED 3SAT instance $\forall x_1 \exists x_2 \ldots \forall x_{n-1} \exists x_n C_1, C_2, \ldots C_p$. Each universal quantifier $\forall x_i$ is represented by a VP-deletion construction. At LF, each of the $n/2$ circled overt VPs is copied to its empty VP position $[_{VP}$ e$]$. The resulting LF is exponentially larger than the S-structure shown. Each existential quantifier $\exists x_{i+1}$ is represented by a linking ambiguity construction. Both quantifier constructions contain permissible antecedents that represent possible truth assignments to their quantified variable. Each 3-CNF clause C_j is represented by a $\mathbf{CC}_j$ construction that contains a *selected* pronoun that must link outside that construction, to a permissible antecedent in some dominating $\mathbf{QC}_i$ construction. These obligatory long distance links are drawn with dashed arrows. Each selected pronoun represents a true literal in its 3-CNF clause.

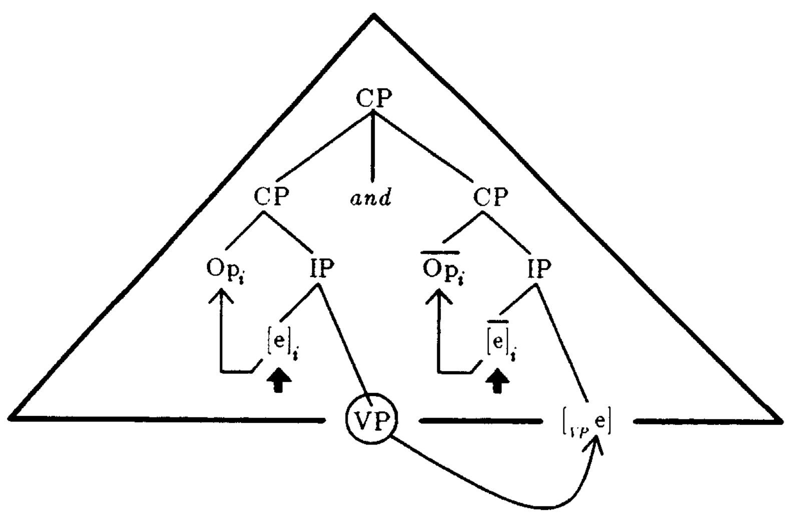


Figure 6: The VP-deletion construction $\mathbf{QC}_i$ represents the universal quantifier $\forall x_i$. The disagreeing operator-variable chains (Op_i, $[\mathrm{e}]_i$) and ($\overline{\mathrm{Op}}_i$, $\overline{[\mathrm{e}]}_i$) represent the truth assignments $x_i = 1$ and $x_i = 0$, respectively. LF linking conditions ensure that pronouns contained in the CC constructions may only link to the subject of the VP that contains them (identified by bold arrows). Therefore, only pronouns in the original VP representing unnegated literals of x_i may link to $[\mathrm{e}]_i$. Similarly, only pronouns in the copied VP representing negated literals of x_i may link $\overline{[\mathrm{e}]}_i$. The circled VP leaf node will immediately dominate the $\mathbf{QC}_{i+1}$ construction representing $\exists x_{i+1}$.

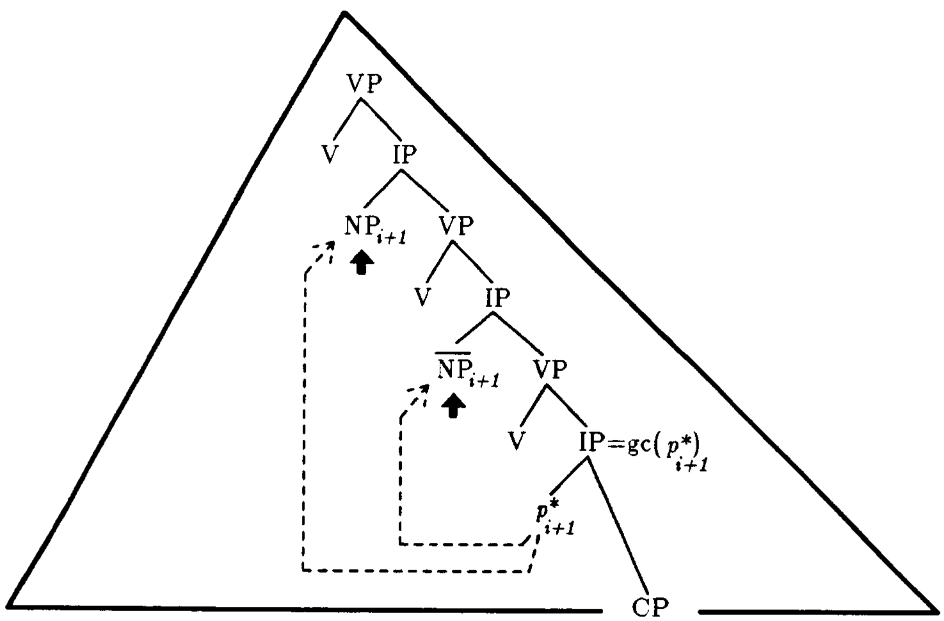


Figure 7: The linking ambiguity construction $\mathbf{QC}_{i+1}$ represents the existential quantifier $\exists x_{i+1}$. The disagreeing R-expressions NP_{i+1} and $\overline{\mathrm{NP}}_{i+1}$ (identified by bold arrows) represent the truth assignments $x_{i+1} = 1$ and $x_{i+1} = 0$, respectively. The pronoun p^*_{i+1} must link to one of these two R-expressions, which it then "disables" as follows. No pronoun contained in a **CC** construction may link to p^*_{i+1} or its antecedent, because of disagreement on the '*' feature. Therefore, all pronouns representing literals of x_{i+1} must link to the same R-expression, either NP_{i+1} or $\overline{\mathrm{NP}}_{i+1}$, which corresponds to assigning a consistent truth value to x_{i+1} everywhere. The CP leaf node will immediately dominate the $\mathbf{QC}_{i+2}$ construction representing $\forall x_{i+2}$.

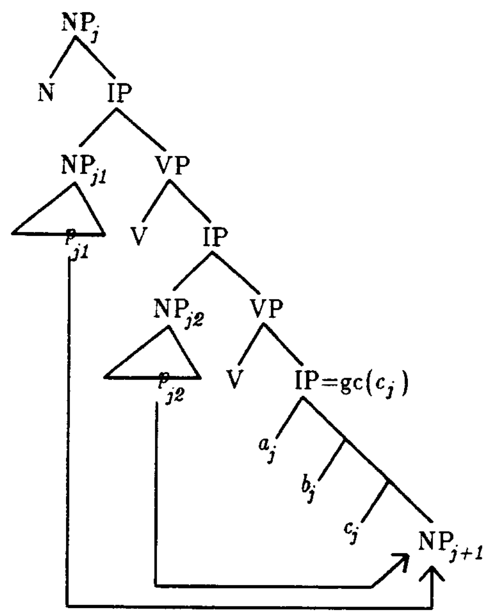


Figure 8: The S-structure construction $\mathbf{CC}_j$ represents the clause C_j of the 3-CNF formula. Only two antecedents (NP_{j1}, NP_{j2}) are available to the three disjoint pronouns a_j, b_j, c_j in the construction. Therefore at least one of the three pronouns must be linked to one of the c-commanding antecedents in some dominating $\mathbf{QC}_i$ construction. The NP_{j+1} leaf will immediately dominate the next $\mathbf{CC}_{j+1}$ construction.

borne by all pronouns in the **CC** constructions. Thus, feature disagreement prevents both p^*_{i+1} and its antecedent (either NP_{i+1} or $\overline{\mathrm{NP}}_{i+1}$) from binding any pronoun (representing a literal of x_{i+1}) contained in any $\mathbf{CC}_j$. Therefore, those pronouns in $\mathbf{CC}_j$ must be bound by the NP subject that p^*_{i+1} does *not* link to. For example, if p^*_{i+1} links to NP_{i+1}, then $\overline{\mathrm{NP}}_{i+1}$ may bind pronouns in any $\mathbf{CC}_j$ that correspond to $\overline{x}_{i+1}$, which is true iff $x_{i+1} = 0$. Otherwise, p^*_{i+1} must link to $\overline{\mathrm{NP}}_{i+1}$, and then NP_{i+1} will bind those pronouns that correspond to x_{i+1}, which is true iff $x_{i+1} = 1$.

3-CNF clauses are represented by linking ambiguities. In the S-structure construction $\mathbf{CC}_j$, each clause C_j of the Boolean formula is represented by a noun phrase NP_j, and each literal of the Boolean clause is represented by a pronoun with a truth value that makes that literal true. (Consequently, pronouns representing negated literals will only agree with $\overline{\mathrm{Op}}_i$ and $\overline{\mathrm{NP}}_{i+1}$ in some dominating $\mathbf{QC}_i$ construction.)

Feature agreement forces pronouns p_{j1} and p_{j2} to both link to NP_{j+1}. Therefore, no pronoun dominated by NP_{j+1} may have NP_{j1}, NP_{j2}, p_{j1}, or p_{j2} as an antecedent without creating an impermissible circular dependence. The three pronouns a_j, b_j, c_j c-command NP_{j+1} and all it contains; therefore, they cannot be linked to any A-position below them. Nor may they be linked to each other or otherwise share an antecedent without violating linking condition B, because they are in the same governing category. (Nothing hinges on being able to stuff three pronouns in one governing category, because disjoint reference can always be accomplished with feature disagreement.) Moreover, because only two distinct antecedents (NP_{j1}, NP_{j2}) are available to the three disjoint pronouns in the entire $\mathbf{CC}_j$ construction, at least one of the three pronouns must be linked to — and agree with — a c-commanding subject outside $\mathbf{CC}_j$, in some dominating $\mathbf{QC}_i$ construction. This pronoun represents a true literal in the Boolean clause C_j: it agrees with the subject it is linked to in some $\mathbf{QC}_i$, and is called the *selected* pronoun. Note that the $\mathbf{CC}$ constructions are contained inside $n/2$ VP-deletion constructions at S-structure, and that therefore the corresponding LF representation will contain $2^{n/2}$ copies of the $\mathbf{CC}_j$ S-structure, each capable of selecting its own pronoun.

The selected pronoun may represent a literal of either a universally or existentially quantified variable x_i. If x_i is universally quantified, then the agreement condition ensures that only pronouns p_i corresponding to unnegated literals x_i and contained in the original VP at LF can link to the subject $[\mathrm{e}]_i$ of the original VP. Conversely, only pronouns $\overline{p_i}$ corresponding to negated literals $\overline{x_i}$ and contained in the copied VP at LF can link to the subject $\overline{[\mathrm{e}]}_i$ of the copied VP.

Otherwise, x_i is existentially quantified, and the agreement condition ensures that pronouns p_i corresponding to unnegated literals x_i can only link to NP_i, that pronouns $\overline{p_i}$ corresponding to negated literals $\overline{x_i}$ can only link to $\overline{\mathrm{NP}}_i$. Recall that if any pronoun in any $\mathbf{CC}_j$ links to NP_i (or $\overline{\mathrm{NP}}_i$), then all pronouns in all $\mathbf{CC}_j$ must link to that antecedent.

The $\mathbf{CC}_j$ construction is permissible iff all of its LF copies are permissible, which is only possible when the Boolean clause C_j contains a true literal for any possible quantifier-determined truth assignment to its literals, as represented by the dominating **QC** constructions. Therefore, the LF representation of the entire S-structure construction is permissible iff the quantified formula Ω is true.$\square$

Note that the binding constructions of this proof can also be used to give a reduction from 3SAT to the binding problem, where each pronoun is no more than four-ways ambiguous.

In our examination of transformational theories of syntax, it was not clear how the complexity of the lexical resolution problem reflected on the complexity of the more general language comprehension problem because it was conceivable, although unlikely, that the phonological form corresponding to the complex S-structure of the reduction could also correspond to a simple S-structure. If that were the case, the hearer could always choose the simpler S-structure, the difficult instances of the LRP would not arise in practice, and the complexity of the LRP would not directly bear on the complexity of language perception.

But in the preceding proof, and in the proof that the S-structure binding decision problem is NP-hard, the phrase structure constructions are trivial: there is no movement, all positions are strongly identified, and all features are explicit on the lexical items. The lack of ambiguity in and the directness of the PF to S-structure mapping for these constructions provides extremely powerful evidence that language perception is intractable.

A central consequence of the PSPACE-hard result is that linguistic representations in the T-model do not have efficient witnesses, unless NP = PSPACE. In another words, it is not possible to efficiently determine the grammaticality of a complete linguistic representation (in our case, a completely specified D-structure, S-structure, LF, and PF). Therefore, using linguistic knowledge in language comprehension can be intractable. (The complexity of language comprehension is measured relative to the size of the phonological form or surface structure.)

Theorem 8 *Using linguistic knowledge is PSPACE-hard.*

Proof. A corollary to theorem 7, given the triviality and directness of the PF to S-structure mapping (hearing) and its inverse (speaking) for the VP-deletion/linking ambiguity construction of theorem 7.$\square$

CS-TR Scanning Project
Document Control Form

Date : 5 / 26 / 95

Report # AIM-964

Each of the following should be identified by a checkmark:
Originating Department:

- [x] Artificial Intellegence Laboratory (AI)
- [ ] Laboratory for Computer Science (LCS)

Document Type:

- [ ] Technical Report (TR)
- [x] Technical Memo (TM)
- [ ] Other: ____________

Document Information

Number of pages: 59 (55-images)
Not to include DOD forms, printer intstructions, etc... original pages only.

Originals are:

- [x] Single-sided or
- [ ] Double-sided

Intended to be printed as :

- [ ] Single-sided or
- [x] Double-sided

Print type:

- [ ] Typewriter
- [ ] Offset Press
- [x] Laser Print
- [ ] InkJet Printer
- [ ] Unknown
- [ ] Other: ____________

Check each if included with document:

- [x] DOD Form (2)
- [ ] Funding Agent Form
- [ ] Cover Page
- [ ] Spine
- [ ] Printers Notes
- [ ] Photo negatives
- [ ] Other: ____________

Page Data:

Blank Pages (by page number): ____________

Photographs/Tonal Material (by page number): ____________

Other (note description/page number):

Description :	Page Number:
(A) IMAGE MAP (1) UN#'ED TITLE PAGE	
(2-49) PAGES #'ED 1-48	
(50-52) SCANCONTROL, DOD(2)	
(53-55) TRGT'S (3)	

(B) CUT & PASTE FIG'S ON PAGE TITLE PAGE (-) MARKS ON TOP OF EACH PA[GE]

Scanning Agent Signoff:

Date Received: 5 / 26 / 95 Date Scanned: 6 / 7 / 95 Date Returned: 6 / 8 / 95

Scanning Agent Signature: Michael W. Cook

Rev 9/94 DS/LCS Document Control Form cstrform.vsd

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

AD# ADA214591

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: AIM 964	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): Complexity of Human Language Comprehension	5. TYPE OF REPORT & PERIOD COVERED: memorandum
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): Eric Sven Ristad	8. CONTRACT OR GRANT NUMBER(s): N00014-85-K-0124
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Artificial Intelligence Laboratory, 545 Technology Square, Cambridge, MA 02139	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS: Advanced Research Projects Agency, 1400 Wilson Blvd., Arlington, VA 22209	12. REPORT DATE: December 1988
	13. NUMBER OF PAGES: 49
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): Office of Naval Research, Information Systems, Arlington, VA 22217	15. SECURITY CLASS. (of this report): UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Distribution is unlimited

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

18. SUPPLEMENTARY NOTES

None

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

linguistic theory
natural language
computational complexity
government-binding
phonology
syntax

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

See reverse side

Block 20

Abstract:

The goal of this research has been to understand the computational structure of principle-and-parameter linguistic theories: what computational problems do these theories pose and what is the underlying structure of those computations? To do this, I have analyzed the computational problem of human language comprehension: what linguistic representation is assigned to a given sound? This *language comprehension problem* may be factored into smaller, interrelated (but independently statable) problems defined on partial phonological, morphological, and syntactic representations. For example, in order to understand a given sound, the listener must assign a phonetic form to the sound; determine the morphemes that compose the words in the sound; and calculate the linguistic antecedent of every pronoun in the utterance. I prove that these and some other subproblems are all NP-hard, and that language comprehension is itself PSPACE-hard, according to current linguistic theory.