HUMAN LANGUAGE COMPREHENSION IS NP-COMPLETE

Eric Sven Ristad

CS-TR-337-91

June 1991

Human Language Comprehension is NP-complete

Eric Sven Ristad*
Department of Computer Science
Princeton University

*The author is indebted to James Higginbotham and Alec Marantz for insights into anaphora. Thanks also to Robert C. Berwick, Sandiway Fong, Shafi Goldwasser, Ron Greenberg, Ray Hirschfeld, Andras Kornai, Ed Stabler, Ken Steiglitz, and an anonymous *Information and Computation* reviewer for their help in clarifying my presentation. The author was supported by a generous fellowship from IBM while working on much of the material reported in this essay.

Abstract:

Consider the computational problem of understanding the utterances of a human language that contain pronouns. In order to completely understand such utterances, the language user must determine the intended reference of each pronoun in a given utterance. For example, in order to comprehend the English sentence *Jocasta loved her son*, the hearer might determine that the possessive pronoun *her* refers to Jocasta, the queen of Thebes. We prove that two different models of this broad computational problem are NP-hard, and argue that the problem remains in $\mathcal{NP}$ even if our language models account for the computationally complex phenomenon of syntactic ellipsis. These complexity results are based directly on human linguistic knowledge, and are invariant across linguistic theories. No knowledge of linguistic theory is needed to understand the analysis, only knowledge of English.

1 Introduction to the problem

This essay investigates the computational complexity of a fundamental problem in human language, that of determining the antecedents of anaphoric elements. This computational problem has not been precisely defined before, nor has its computational complexity been analyzed previously.

An *anaphoric element* is a noun that does not itself refer to an object or concept in the world. Rather, anaphoric elements refer indirectly, through their antecedents. Examples of anaphoric elements include pronouns (eg., *he*, *they*), reflexives (eg., *himself*, *themselves*), and reciprocals (eg., *each other*). A *referring-expression* is a noun phrase that has intrinsic reference, such as a proper noun (eg., *Ann*, *Bob*) or a definite noun phrase (eg., *the man across the street*, *my sister*). In the sentence *Narcissus loves himself*, we would say that the proper noun *Narcissus* must be the antecedent of the reflexive *himself*, and that *himself* refers to whatever person the referring-expression *Narcissus* refers to.

The question of how anaphoric elements find their antecedents has been a central topic in the study of language from the time of antiquity. In order to understand an utterance containing anaphoric elements, the language user must compute the intended antecedent of each anaphoric element in that utterance; failing that, the language user has failed to completely comprehend the utterance. This is the broad computational problem posed by anaphora in human language comprehension. For example, in order to comprehend the English sentence *Oedipus loved his mother*, the language user might determine that the intended antecedent of the possessive pronoun *his* is proper noun *Oedipus*, and that both refer to Oedipus, prince of Thebes.

In this essay, we construct four precise yet realistic models of this anaphora problem, and analyze their computational complexities.

Our results may be summarized as follows. Section 3 proves the NP-hardness of the comprehension problem for human language under two different models of the anaphora problem. Each model is based directly on well-understood, uncontroversial empirical arguments from linguistics. For this reason, we may believe with confidence that these NP-hardness results apply to all adequate linguistic theories, and that they therefore establish the NP-hardness of the language comprehension problem as a whole. Section 4 shows how a widely-accepted analysis of the linguistic phenomenon of ellipsis leads to a proof that the anaphora problem is PSPACE-hard. Next, we falsify this analysis with counterexamples, and sketch a descriptively superior account of ellipsis that reduces the complexity of elliptical anaphora to $\mathcal{NP}$. This constitutes evidence for an $\mathcal{NP}$ upper bound on the anaphora problem, because ellipsis is arguably the most computationally complex phenomenon in the syntax of human languages.

2 Preliminaries

This work is part of a larger research program whose goal is to characterize the complexity of human language, using only empirical facts of linguistic knowledge (Chomsky 1956; Ristad 1990b). Such a complexity thesis, with tight upper and lower bounds, provides a useful *a priori* design criterion. It objectively measures the significance of a particular design decision, in terms of its effect on the complexity of the language model. It tells us when a language model is too restrictive, and what is a reasonable empirical generalization. Language models that do not satisfy the lower bound do not have an adequate account of complex linguistic phenomena. (Any particular design decision or empirical generalization may of course violate the lower bound, which is relevant only to the model in its entirety.) In the absence of overwhelming counterevidence, a reasonable empirical generalization is one that conforms to the upper bound.

2.1 A complexity thesis for language

In *Three Models for the Description of Language*, Chomsky (1956) argued that human language has a finite—but no finite-state—characterization, and that the simplest and most revealing characterization is given by a class of unrestricted rewriting systems.

Ristad (1990b) argues human language comprehension and production are NP-complete. This brings Chomsky's 1956 complexity thesis up-to-date, with the advantages of a better understanding of language (the result of thirty years of productive research in linguistics) and a more precise theory of structural complexity, based on computational resources rather than on the format of grammars or automata. The resulting thesis is also much stronger, providing tight upper and lower bounds, and therefore is a truly *constructive* complexity thesis for human language.

An NP-complete problem is hard to solve because the input to the problem is missing some crucial information (the efficient witness), but once the efficient witness is found, it is easily verified to be correct. As stressed by the great nineteenth century linguist Wilhelm von Humboldt, every sound uttered as language is assigned a complete meaning and linguistic representation in the mind of the producer.[1] It is the task of comprehension to find the intended linguistic representation, given only the utterance. When the utterance is missing crucial disambiguating information, and there are global dependencies in the representation, then the task of finding the intended representation quickly becomes very difficult. Yet we know comprehension cannot be too difficult, simply because there is always an efficient witness, namely the linguistic representation from which the utterance was produced.

The central empirical consequence of this complexity thesis is that language models based directly on scientifically adequate theories of linguistic knowledge must be NP-complete, under appropriate idealizations. If such a linguistic system is outside $\mathcal{NP}$, say PSPACE-hard, then the thesis predicts that the system is unnaturally powerful, perhaps because it overgeneralizes from the empirical evidence or misanalyzes some linguistic phenomena. Such a system must be capable of describing unnatural languages. If, however, a complete system is easier than NP-hard, and assuming $\mathcal{P} \neq \mathcal{NP}$, then the system is predicted to be unnaturally weak, most likely because it does not adequately account for some complex linguistic phenomena. Such a system will not be able to describe all human languages. Otherwise the system is NP-complete and is potentially adequate, pending the outcome of more exacting tests of scientific adequacy.

This complexity thesis makes strong predictions, because many proposed linguistic theories violate it, and because the $\mathcal{NP}$ lower bound is in sharp contrast to the prevailing belief that language is efficient, which is held by many students of human language. In the body of this article, we first prove the NP-hard lower bound using two distinct linguistic models. Next we demonstrate the utility of the $\mathcal{NP}$ upper bound by using it to guide the revision of the copy theory of syntactic ellipsis.

[1] "The sentence is not to be constructed, is not to be gradually built up of components, but is to be expressed all at once in a form compressed to unity.... Man inwardly relates a complete meaning with every sound emitted as language: that is, for him it is a complete utterance. Man does not intentionally emit merely an isolated word, even though his statement according to our viewpoint may only contain such an entity." (von Humboldt, 1836:110–111) (The upper bound of Chomsky's 1956 complexity thesis, that language has a finite description, also appears to be motivated by an observation due to von Humboldt, that language is the "infinite use of finite means.")

2.2 Defining the language comprehension problem

The central technical obstacle that we face in studying human language is the incomplete nature of our understanding. Any comprehensive formal model of human language will therefore be the outcome of many arbitrary and ad-hoc decisions that could not be justified scientifically. Consequently, no meaningful formalization is possible, and mathematical analysis of such formal systems would have little or no relevance to language itself. Prior mathematical analyses of language have all been of this nature, proving properties of formal systems within which theories of linguistic knowledge may be represented. These proofs are based on the ad-hoc particulars of the formal system, and only very tenuously (if at all) on empirical facts. Indirect analyses of human language may be found in Peters and Ritchie (1973) and Barton, Berwick and Ristad (1987). In contrast, our proofs are based only on empirically correct (but incomplete) formal models.

Lacking a comprehensive formal model, it is not possible to prove an upper bound on the complexity of human language. Nor can we establish a lower bound without a precise statement of the language comprehension (LC) problem. We may overcome these obstacles as follows.

In order to define the LC problem, we first select a natural class of utterances, and use the scientific methods of linguistics to determine what knowledge language users in fact have about those utterances. Next, we construct the simplest theory of that knowledge, under an appropriate idealization to unbounded utterances. Finally, we pose the abstract problem of computing that knowledge for a given utterance. These problems are natural subproblems of language comprehension, because in order to comprehend an utterance in that class, the language user must compute that knowledge. Therefore, the complexity of such subproblems are lower bounds on the complexity of the complete LC problem, by the principle of sufficient reason.

How then shall we define the anaphora problem? As stated above, a language user must determine the intended antecedent of every anaphoric element in a given utterance. However, there is no known satisfactory characterization of what is "the intended antecedent of an anaphoric element." It cannot be literally "the antecedent intended by the producer," simply because the comprehender does not have direct access to the producer's intentions. Nor can it be *any* antecedent, because this results in the trivial language *mis*comprehension problem. Such a problem statement would allow the trivial (non)solution where every pronoun in the utterance is assigned a distinct antecedent, none of which were previously mentioned in the discourse.

In order to overcome this difficulty without making any unjustified or unnecessarily strong assumptions, we require that solutions to the anaphora problem introduce no new information. That is to say, the antecedent of each anaphoric element must be drawn from the set of available antecedents, in the current utterance and in previous utterances, produced earlier in the discourse. The anaphora problem must also be defined in terms of linguistic representations, and not strings of discrete terminal symbols (that is, linguistic expressions) in order to prevent the trivial (non)solution where an expression is assigned the null linguistic representation, and no anaphoric elements are understood to be present.

For these reasons the broad *anaphora problem* is defined to be: Given a linguistic representation R lacking only relations of referential dependency, and a set A of available antecedents, decide if all the anaphoric elements in R can find their antecedents in A. The set of available antecedents models the discourse context in which the utterance is produced. In the next section, we examine two models of this problem, one based on the fact that anaphoric elements must agree in certain respects with their antecedents, the other based on the fact that pronouns must be disjoint in reference from certain potential antecedents.

Although we cannot prove an upper bound on the complexity of anaphora, we can still accumulate empirical evidence for one. One way to confirm a thesis is to confirm its predictions. An upper bound makes the following prediction: if an analysis of a linguistic phenomena leads to complexity above the upper bound, then the analysis is in error, and the phenomena has an empirically superior analysis whose complexity is below the upper bound. Therefore, every time that we improve a linguistic analysis while reducing its complexity from outside a complexity class $\mathcal{C}$ to inside $\mathcal{C}$, we accumulate additional empirical evidence for an upper bound of $\mathcal{C}$.[2] In section 4, we pro-

[2]One piece of evidence for an $\mathcal{NP}$ upper bound may be found in Ristad (1990a), which proves that the universal recognition problem (URP) for generalized phrase structure grammars (GPSG) is EXPPOLY-hard, and shows how to construct an empirically superior "Revised GPSG," whose URP is NP-complete. Another piece of evidence is the analysis of phonological theories in chapter 3 of Ristad (1990b) , whose complexity is reduced from undecidable to inside $\mathcal{NP}$. To my knowledge, no other linguistic theory has been proved to have a complexity outside of $\mathcal{NP}$. The work of Peters and Ritchie (1973), who proved that a formal model of their own design was undecidable, and Rounds (1975), who proved that a restricted version of the Peters-Ritchie model was exponential time, is not relevant because their formal model was not independently proposed by linguists, or ever defended

vide evidence along these lines for an $\mathcal{NP}$ upper bound on the complexity of the anaphora problem.

It is not to be expected that the utterances constructed below are easy to comprehend, any more than we expect to actually build the physical devices used to prove lower bounds on the complexity of problems in robot motion planning (Reif 1979; Reif and Sharir 1985; Canny 1988). Certainly, it is not possible to build physical devices of such intricacy, any more than it is possible for a language user to comprehend the utterances we construct below. Yet the practical questions of what physical devices can and cannot be built, or of what linguistic expressions can and cannot be easily understood do not concern us here. We want to understand the theoretical structure of abstract computational problems, and use complexity analysis to better reveal this structure.

3 Two models of the anaphora problem

In this section, we use basic facts about the language user's knowledge of referential dependencies to construct two distinct models for the anaphora problem, and then prove that each is NP-hard.

3.1 The agreement model

What is it that language users know about anaphora?

For one, language users know that an anaphoric element α may inherit the reference of an argument β, as in *Todd*$_1$ *hurt himself*$_1$, where the reflexive *himself* is understood as referring to the proper noun *Todd*, or *Todd*$_1$ *said Mary liked him*$_1$, which could mean that 'Todd said Mary liked Todd'. The judgement of coreference between α and its antecedent β is depicted here by assigning α and β the same subscript. (Careful attention must be paid to the the intended interpretation of the anaphoric elements in the examples below, as indicated by the subscripts.)

Language users also know that an anaphoric element must agree with its antecedent in certain respects. Examples (1) are possible only if *Chris* is masculine, whereas (2) are possible only if *Chris* is feminine.

as a remotely plausible linguistic theory. It amounts to little more than a very loose interpretation of the statement "a grammar maps deep structures to surface structures by the repeated application of rewriting rules."

(1)
a. Chris_1 liked himself_1.
b. Chris_1 thought Bill liked him_1.

(2)
a. Chris_1 liked herself_1.
b. Chris_1 thought Bill liked her_1.

This condition on agreement is transitive, as illustrated by the paradigm in (3), where *the student* can be masculine (3a) or feminine (3b), but not both simultaneously (3c).

(3)
a. The student_1 prepared her_1 breakfast
b. The student_1 did his_1 homework.
c. *The student_1 prepared her_1 breakfast after doing his_1 homework.

The asterisk '*' is used in (3c) to indicate the impossibility of the depicted interpretation.

Knowledge of coreference and the agreement condition must be represented in the brain of the language user, and hence by every scientifically adequate linguistic theory. The simplest, and least controversial, representation is to postulate an abstract linguistic relation of antecedence between an anaphoric element and its antecedent and an agreement condition defined over the set of syntactically-relevant features.

More precisely, let a *feature* be a [*feature-name, feature-value*] pair, such as, [`gender, masculine`]. Features are a formalization of the notion "relevant syntactic distinction." Every linguistic element is associated with a set of such features. Two elements are *nondistinct* if and only if their feature sets do not disagree on any common feature (that is, they may be unified). The facts of interpretation illustrated in (1-3) above are commonly generalized to state: all anaphoric elements that share an antecedent must be nondistinct from it and from each other.

Pronouns in different languages are marked for a wide range of distinctions in person, gender, number, animacy, case, social class, kinship, reference, antecedent noun class, grammatical function, thematic role, and so on (cf., Wiesemann 1986; Sells 1987). It is true that every particular language contains a fixed number of agreement features. However, linguistic theory idealizes to an unbounded number of agreement features because these features and the range of their possible values varies considerably from language to language and does not seem to be restricted in principle (see Pullum (1983) and Ristad (1990b), appendix A).

We may also formalize the notion of antecedence, by postulating an assymetric $\mathsf{link}(\alpha, \beta)$ relation between an anaphoric element α and its immediate antecedent β, subject to the agreement condition (following Higginbotham 1983). Let link^+ be the positive transitive closure of the directed link relation. The anaphora problem in this agreement model is: Given a linguistic expression containing sets of anaphoric elements E and available antecedents A, Find a set of links $L \subseteq E \times A$ s.t. (i) every $e \in E$ is linked to some $a \in A$; and (ii) $\mathsf{link}^+(e_1, a) \in L^+$ and $\mathsf{link}^+(e_2, a) \in L^+$ only if e_1, e_2, and a are nondistinct.

The fact that all anaphoric elements that share an antecedent must agree with it and with each other, in combination with the fact that anaphoric elements must have antecedents, suffices to establish the following lemma.

Lemma 3.1 *Graph k-coloring* $\propto$ *Anaphoric agreement.*

Proof. On input k colors and a graph $G = (V, E)$ with vertices $V = \{v_1, v_2, \ldots, v_n\}$ and edges E, we construct a natural utterance U containing $|V|$ pronouns and k available antecedents such that G is k-colorable if and only if the pronouns in U can be linked to the k available antecedents without violating the agreement condition. Available antecedents correspond to colors; pronouns in U correspond to vertices in G; and disagreement between the pronouns in U corresponds to edges in G.

To do this we need the n binary agreement features $\varphi_1, \varphi_2, \ldots, \varphi_n$; the pronouns $p_1, p_2, \ldots, p_n$; and the available antecedents $R_1, R_2, \ldots, R_k$. Each R_i is a referring-expression, such as a noun phrase, that is not specified for any of the n agreement features. Pronoun p_i represents vertex v_i: for each edge $(v_i, v_j) \in E$ attached to v_i, pronoun p_i has $\varphi_i = 0$ and pronoun p_j has $\varphi_i = 1$. It does not matter how the pronouns and referring-expressions are arranged in the expression U, provided that every R_i is a possible antecedent for each p_j, and that no other linguistic constraints interfere with the disagreement relations we are constructing. It is always trivial to quickly construct such a sentence, as we did in example (3). (Simply make all the p_i and R_j arguments of different verbs, which will make them nonlocal, and then make sure that none of the p_i c-commands any of the R_j. This is discussed in greater detail below.)

In order to be understood, every pronoun must refer to one of the k available antecedents. If there is an edge between two vertices in the input graph G, then those two corresponding pronouns cannot share an antecedent in the utterance U without disagreeing on some agreement feature. Therefore each permissible interpretation of U exactly corresponds to a k-coloring of the input graph G. $\square$

The reduction uses n binary agreement features, one for each vertex in the graph. The feature system is used to represent subsets of the n vertices, and therefore must be capable of making an exponential number of distinctions. (In terms of the input length $m = |G|$, this feature system is capable of making $2^{m^{1/2}}$ distinctions.)

3.2 The referential dependence model

We have seen that anaphoric elements must have antecedents, subject to an agreement condition. A second component of linguistic knowledge is that that pronouns must be disjoint in reference from certain arguments. For example, every English speaker knows that *Todd*$_1$ *hurt him*$_{*1}$ cannot mean that 'Todd hurt Todd'. This judgement of disjoint reference, that α cannot refer to β, is depicted here by assigning α the subscript of β, preceded by an asterisk.

For the same reasons that knowledge of coreference must be represented in the brain of the language user, and hence by every scientifically adequate linguistic theory, so must knowledge of disjoint reference. In particular, we are led to postulate a second abstract linguistic relation: a symmetric **obviate**(α, β) relation that holds between two arguments α and β that cannot share any referential values (Higginbotham, 1985).

Like the agreement condition, the prohibition against sharing referential values is enforced globally, as shown by the paradigm in (4).

(4)
a. John$_1$ said that [Bill$_2$ liked him$_{1/*2}$].
b. John$_1$ said that [he$_{1/*2}$ liked Bill$_2$].
c. *John$_1$ said that [he$_1$ liked him$_1$].

Him can refer to *John* in (4a); *he* can refer to *John* in (4b); but *he* and *him* cannot both refer to *John* in (4c), because *he* locally c-commands *him* and hence they are globally obviative.

Every linguistic representation, then, includes an undirected graph of **obviate** relations whose vertices are the arguments in the linguistic representation and whose undirected edges represent the obligatory nonoverlapping reference of two arguments. Let the combined graph of referential dependencies, which consists of the **obviate** and **link** relations of a given linguistic representation, be called the *RDG*.

The anaphora problem in this model of referential dependence is: Given a natural utterance whose linguistic representation contains the sets of available antecedents A, anaphoric elements E, and obviation relations $O \subseteq (E \cup A) \times (E \cup A)$, Find a set of links $L \subseteq E \times (E \cup A)$ s.t. (i) every anaphoric element $e \in E$ is linked to some $a \in (A \cup E)$; and (ii) $\mathsf{link}^+(e_1, a) \in L^+$ and $\mathsf{link}^+(e_2, a) \in L^+$ only if $\mathsf{obviate}(e_1, e_2) \notin O$. (Recall that link^+ is the positive transitive closure of the assymetric **link** relation.)

Our goal is to prove the NP-hardness of the anaphora problem without using features or the agreement condition. The idea of the proof will be to reduce some NP-complete problem to the problem of computing the **link** relations in a way that satisfies the **obviate** relations. Let us therefore examine a range of syntactic configurations, in order to better understand the distribution of **link** and **obviate** relations in linguistic representations.

3.2.1 Local c-command configuration

We say α *c-commands* β in a phrase structure tree if and only if all branching nodes that dominate α in the tree also dominate β.[3] In particular, the direct object c-commands the indirect object, and the subject of a clause c-commands both direct and indirect objects. To a first approximation, we say two elements are *local* if they are co-arguments, that is, arguments of the same verb or noun.

The first syntactic configuration with consequences for referential dependencies is "local c-command": reflexives (and reciprocals) must link to some locally c-commanding β and pronouns must obviate all such β. In addition, each anaphoric element obviates all referring-expressions that it c-commands, regardless of whether they are local or not.[4] The exact definition of local c-command does not matter for our purposes. All that matters here is the fact that antecedence and disjoint reference are understood to be possible or necessary in some configurations, and not in others.

A reflexive must be linked to a unique c-commanding argument, and this argument must be local. This is illustrated by the paradigm in (5), where the domain of locality is indicated by square brackets.

[3] A phrase structure tree is a kind of derivation tree for a natural language expression. See Winston (1984), chapter 9, or Hopcroft and Ullman (1979), section 4.3.

[4] The requirement that reflexives and reciprocals have local antecedents is called "condition A" and the requirement that pronouns be locally obviative is called "condition B" in the linguistics literature. The third requirement is called "condition C".

(5)
a. [John_1 shot himself_1]
b. [John_1 introduced Bill_2 to $\text{himself}_{1/2}$]
c. John_1 thought Bill_2 said [Mary liked $\text{himself}_{*1/*2}$]

Example (5b) shows that a reflexive in the indirect object position can take any c-commanding antecedent inside its local domain; example (5c) shows that a reflexive must have some antecedent inside its local domain. *Mary* is not a possible antecedent for *himself* in (5c) because they disagree on gender.

Pronouns are locally obviative: a pronoun cannot share referential values with any argument that c-commands it in its local domain. (The domain of locality is roughly the same as for reflexives; again, all is needed for the proofs below is that there exist configurations that result in obviation.) This is illustrated by the paradigm in (6).

(6)
a. [John_1 shot him_{*1}]
b. [John_1 introduced Bill_2 to $\text{him}_{*1/*2}$]
c. John_1 thought Bill_2 said [Mary liked $\text{him}_{1/2}$]

Example (6b) shows that a pronoun is disjoint from all locally c-commanding arguments; example (6c) shows that a pronoun can link to any argument outside its local domain.

Pronouns (and other anaphoric elements) also obviate all referring-expressions that they c-command, regardless of whether they are local or not. This is illustrated in (7).

(7)
a. [He_{*1} shot John_1]
b. [$\text{He}_{*1/*2}$ introduced John_1 to Bill_2]
c. Mark told $\text{himself}_{*1/*2/*3/*4}$ that [$\text{he}_{*1/*2/*3/*4}$ said that [John_1 thought [Bill_2 wanted [Sam_3 to like Fred_4]]]]

Obviation applies equally to all linguistic coreference, including the intra- and inter-sentential linking of pronouns, because obviation cannot be violated, even if a pronoun and its antecedent are in different sentences. Without loss of generality then, all linkings in this essay will be intrasentential.[5]

[5]Pronouns may also find their antecedents extralinguistically in (at least) two ways, neither of which prejudices the following discussion. First, the antecedent of a pronoun may be 'demonstrated' extralinguistically, as when the speaker points to an antecedent seen by the hearer. For example, if a speaker were to say *Bill saw HIS mother* while stressing the pronoun and vigorously pointing to Jack, then the hearer might understand *HIS*

The other local c-command configuration is "exceptional case-marking" (ECM). In an ECM configuration, the subject of an ECM verb locally c-commands the subject of its infinitival complement. This is illustrated by the paradigm in (8), with the ECM verbs *want* and *expect*.[6]

(8)
a. John_1 wants [himself_1 to shoot Bill]
b. John_1 expects [him_{*1} to shoot Bill]

Examples (8) demonstrate that the subject *John* of an ECM verb locally c-commands the subject α of the infinitival complement [α *to shoot Bill*], for both reflexives and pronouns.

Our goal is to prove that the anaphora problem is NP-hard, without using any agreement features. Let us therefore pause to consider how such a proof might work.

Imagine that we must color the following four-vertex graph G_4 with three colors:

$$\{(1,2),(2,3),(3,4),(4,2)\}$$

Then our reduction might construct a sentence containing three available antecedents and four pronouns. The first part of the sentence, *Before* $\textit{Mark}_a$, $\textit{Phil}_b$, *and* $\textit{Hal}_c$ *were friends....*, would represent the three colors, where each proper noun corresponds to a different color. The second part of the sentence would have an obviation graph equivalent to G_4, where the pronoun p_i in the sentence corresponds to vertex i in G_4. As expected, it is difficult to understand the resulting sentence:

(9) Before Mark_a, Phil_b, and Hal_c were friends,
[he_1 wanted him_2 to introduce him_3 to him_4].

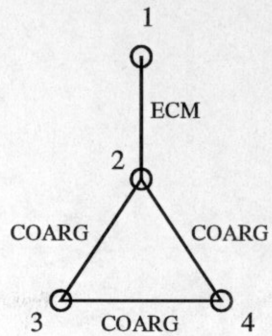
The corresponding obviation graph appears in (10).

---

to extralinguistically refer to "Jack." Second, a pronoun may have no antecedent available, in which case it becomes necessary for the hearer to postulate some new, heretofore unmentioned, antecedent. For example, if we are warned to "Watch out for him," we must postulate the existence of some male individual M that we must watch out for; if we are then told "he's coming to get you," we would link the subject pronoun *he* to the previously postulated individual M. There appears to be a significant cognitive cost to this invention of unmentioned antecedents, because speaker-hearers are very reluctant to do so.

[6]These verbs are called "exceptional case-marking" because, unlike other verbs that take a finite clausal complement, ECM verbs take an infinitival clausal complement and assign abstract case to its subject. Other ECM verbs in English include *believe*, *prefer*, *like*, and related verbs.

(10)



Each vertex is labeled with the numerical index of its corresponding pronoun, and each edge is labeled with the syntactic configuration responsible for the corresponding **obviate** relation. (Recall "coarg" means "argument of same verb or noun," and "ecm" means "exceptional case marking" configuration.)

By carefully grounding the reference of each pronoun in turn, we can confirm that the obviation graph for (9) exactly corresponds to the four-vertex graph G_4. Let *he*$_1$ link to *Mark*$_a$ in the sentence—this corresponds to coloring vertex 1 in G_4 with the color a. Then in the simplified sentence [*Mark wanted him*$_2$ *to introduce him*$_3$ *to him*$_4$] we can clearly see that *Mark* can be the antecedent of any pronoun but *him*$_2$—this corresponds to G_4, where coloring vertex 1 with a given color only prevents us from coloring vertex 2 the same color. Continuing in this fashion, one can convince oneself that the pronouns in such a sentence can find their antecedents in the sentence iff the corresponding graph G_4 is 3-colorable.

The local c-command configurations used in (9) only give rise to simple obviation graphs, and therefore we will extend the referential dependence model to include three additional syntactic configurations: control, strong crossover, and invisible obviation.

3.2.2 Control configuration

In the expression *Sue screamed before jumping*, all English speakers know that *Sue* is the understood subject of the gerund *jumping*, that is, everyone knows that Sue did the jumping. In order to represent this linguistic knowledge—as we must—we may postulate a silent pronoun 'PRO' in the subject position of the adjunct [*before jumping*], and obligatorily link PRO to *Sue*.

(11) Sue$_1$ screamed [before PRO$_1$ jumping]

This is called "subject control" because the reference of PRO is controlled by the subject of the main clause.

Further evidence for the existence of this silent pronoun comes from its interaction with overt anaphoric elements. Observe that *himself* must refer to *Mark* in (12a), and *him* must be disjoint from *Mark* in (12b).

(12)
a. Mark_1 vomited [after PRO_1 getting himself_1 plastered]
b. Mark_1 vomited [after PRO_1 getting him_{*1} plastered]

Without PRO, such facts are a complete mystery. But once the understood subject of the gerund is explicitly represented using PRO, as we have done in (12), the facts are trivially accounted for as canonical configurations of local c-command between silent PRO and an overt anaphoric element. (In any event, the complexity reduction proceeds whether PRO exists or not; all that matters for the reduction is the empirical fact that *himself* must refer to *Mark* in (12a), and that *him* must be disjoint from *Mark* in (12b).)

Another example of subject control appears in (13a) with the verb *promise*. Contrast this to the verb *persuade* in (13b), which is an object control verb.

(13)
a. Tom_1 promised Mary_2 [$\text{PRO}_{1/*2}$ to attend school]
b. Tom_1 persuaded Mary_2 [$\text{PRO}_{*1/2}$ to attend school]

The meaning of (13a) is that "Tom promised Mary that he, Tom, would attend school," while (13b) means that "Tom persuaded Mary that she, Mary, would attend school." PRO is used here to represent the invisible, understood subject α of the embedded clause [α *to attend school*].

3.2.3 Strong crossover configuration

"Wh-movement" is the configuration involving a wh-phrase, such as [*who*] or [*what person*], that appears displaced from its underlying argument position. A *trace* is used to mark the underlying positions of arguments. For example, in *Who*$_k$ *did Mary see* t_k, the underlying position of the wh-phrase *who*$_k$ as the direct object of the verb *see* has been marked with a trace t_k coindexed with it. This represents the fact that *who*$_k$ stands in the same relation to the verb *see* as it does in the related expression *Mary saw who*.

"Strong crossover" occurs when an anaphoric element α intervenes between a wh-phrase and its trace, and c-commands the trace. In such a configuration, α obviates the subject of the wh-phrase. This is shown in

(14a), where the pronoun *he* c-commands the trace t_k of the wh-phrase [*which person*], and for this reason must be understood as disjoint from the person who Mary kissed.

(14) a. [Which person]$_k$ did he$_{*k}$ say t_k kissed Mary.
b. [Which person]$_k$ t_k said he$_k$ kissed Mary.

In (14b), however, there is no strong crossover configuration, and no obviation. That is, (14b) has an interpretation of the form, "for which person x, did x say x kissed Mary." Although we need not postulate any traces for the purposes of the complexity proof below, such facts are difficult to explain without an explicit trace, because the wh-phrase [*which person*] and the pronoun *he* stand in the same structural relation in both sentences.

In the expression *The man who$_k$ Mary saw t_k*, we say that *who* heads the relative clause [*who Mary saw*], and that it predicates its subject, [*the man*]. When the relative clause contains a pronoun in a strong crossover configuration, then the pronoun obviates the subject of the relative clause, as in (15).

(15) [the man]$_1$ [who$_k$ he$_{*1}$ likes t_k].

3.2.4 Invisible obviation configuration

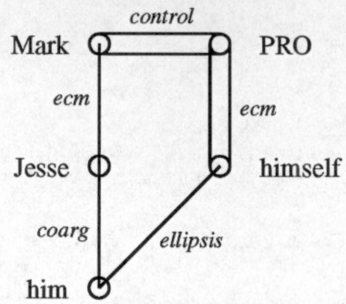
"Ellipsis" is the syntactic phenomenon where a phrase is understood but not expressed in words, as in *The men ate dinner and the women did too*, which can only be understood to mean that 'the women did <u>eat dinner</u> too'. For this example, we would say that the verb phrase [*eat dinner*] has been ellipsed in the second conjunct; this is called VP-ellipsis.

A configuration of "invisible obviation" arises between the subject of an ellipsed verb phrase and the direct and indirect objects of the overt verb phrase, because both subjects in effect locally c-command the other arguments of the verb. Observe that *him* can refer to *Mark* in (16a), but not in (16b):

(16) a. Mark$_1$ wanted Jesse to love him$_1$
b. *Mark$_1$ wanted Jesse to [love him$_1$] before PRO$_1$ allowing himself$_1$ to [*e*]

In example (16b), the reflexive *himself* is obligatorily linked to *PRO*, and *PRO* to the matrix subject *Mark*. The pronoun *him* invisibly obviates *himself*, the subject of the ellipsed VP, because they are in an invisible configuration of local c-command. The RDG for (16b) appears in (17):

(17)



Single lines depict relations of obviation, while double lines depict relations of coreference. Vertices are labeled with the corresponding noun phrase arguments, and the edges are labeled with the configurations to which they are attributable.

It must again be emphasized that the complexity classifications of this section in no way depend on the existence of traces, PRO, links, obviations, sentences, judgements of grammaticality, or on any other details of the linguistic analysis. The reduction relies only on empirical facts of interpetation, that an anaphoric element must corefer or be disjoint in reference from certain other elements in utterances such as (12b), (13), (14a), (15), and (16). The linguistic analysis is included for pedagogy, and to organize the reduction to unbounded utterances, when the facts of referential interpretation are more difficult to elicit.

This concludes our survey of the language user's knowledge of referential dependence, which has been studied extensively in recent years. The next step is a direct complexity proof.

3.3 From satisfiability to referential dependence

As seen above, the conceptually natural reduction is from graph coloring to the problem of computing referential dependencies. However, the transformation of arbitrary graphs into linguistic representations is cumbersome. To overcome this difficulty, we might reduce from 3SAT, by way of the graph 3-coloring problem. That is, on input a Boolean formula f in 3-CNF, we would first use the classic reduction of Lawler (1976) to construct a corresponding instance g of the graph 3-coloring problem. Next, from g we would construct an instance (a, e, o) of the anaphora problem, such that the anaphoric elements in e can find their antecedents in a iff g is 3-colorable (and f is satisfiable.) By restricting our attention to this class of "3SAT graph colorings," we would simplify the reduction into the task of transforming a simple class of "difficult" graphs into linguistic representations. Of course, the intermediate "graph 3-coloring" stage of the reduction won't really be used in the proof.

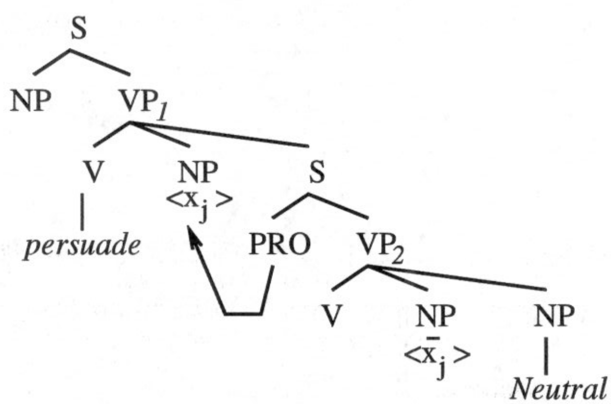
Lemma 3.2 *3SAT* $\propto$ *Referential dependencies.*

Proof. On input a Boolean formula f consisting of the clauses $C_1, C_2, \ldots, C_p$ in the variables $x_1, x_2, \ldots, x_n$, we construct a natural utterance whose linguistic representation contains the sets of available antecedents A, anaphoric elements E, and obviation relations O, such that every anaphoric element in E can find its antecedent in A without violating the obviation relations O iff f is satisfiable.

The set A will contain exactly three distinct antecedents: *True*, *False*, and *Neutral*. These noun phrases represent, respectively, the three possible truth values 'true', 'false', and 'unassigned.'

For every variable x_j in f, E will contain two pronouns, one to represent x_j and the other to represent its negation $\overline{x}_j$. In order to preserve the semantics of negation, these pronouns must obviate each other. Both will also obviate the proper noun *Neutral*, and therefore can only link to *True* or *False*. To be precise, for every Boolean variable we build an object control construction (18) that contains two possible targets for ellipsis, VP_1 and VP_2.

(18)



This is the phrase structure that would be assigned to natural utterances such as *He persuaded him to introduce him to Hector.*

In this construction (18), PRO locally c-commands the pronoun for $\overline{x}_j$ in the lower clause, and so they are obviative (by condition B). Moreover, both PRO and the pronoun for $\overline{x}_j$ c-command the referring-expression *Neutral*, and therefore both must obviate it (by condition C). The object control verb *persuade* obligatorily links PRO to the pronoun for x_j, as depicted by the arrow. The subject position of *persuade* is filled with a dummy pronoun so as not to increase the number of available antecedents in the construction. The resulting RDG ensures that the pronoun for x_j must refer to either *True* or *False*; that the pronoun for $\overline{x}_j$ must refer to the other of the two antecedents; and that neither pronoun can refer to *Neutral*. In short, the construction (18) correctly ensures consistency of truth assignments, as well as correctly representing the semantics of Boolean negation.

Each pronoun that represents a literal of x_j in the construction (18) is in the direct object position of its own verb phrase. Both verb phrases in (18) are possible antecedents for VP ellipsis and hence each pronoun will invisibly obviate the overt subject of any such elliptical VP. We will take advantage of this below, by representing positive literals of x_j as ellipsis of the higher VP_1, and negative literals of x_j as ellipsis of the lower VP_2.

For each clause $C_i = (a_i \vee b_i \vee c_i)$, we construct the rather intricate syntactic structure shown in figure 1, whose graph of referential dependencies appears in figure 2. The effect of this obviation graph in combination with the limited number of available antecedents is to ensure that one of the three ellipsed verb phrases in figure 1 must contain a pronoun linked to the antecedent *True*. This corresponds exactly the requirement that each clause contain a true literal, concluding our reduction. □

It is of course not to be expected that the linguistic expressions constructed by the preceding reduction are easy to comprehend. Nor is this fact relevant here. Recall that our goal is to illuminate the computational structure inherent in innate linguistic knowledge, not to characterize the arbitrary limits of our finite abilities.

All that remains is to state the main theorem of this section.

Theorem 1 *The anaphora problem is NP-hard.*

Proof. By lemma 3.1, or by lemma 3.2. □

We now have two complexity lower-bounds for human language comprehension that rely only on the empirical facts of referential dependency, and on the uncontroverted assumption that these facts generalize in a reasonable manner. *It does not matter exactly how the conditions on coreference and disjoint reference are stated, only that there are such conditions, as there are in all known human languages.* Therefore, we may expect that this NP-hardness result will apply to all adequate linguistic theories, and that it is true of human language itself. Moreover, the directness of the reductions

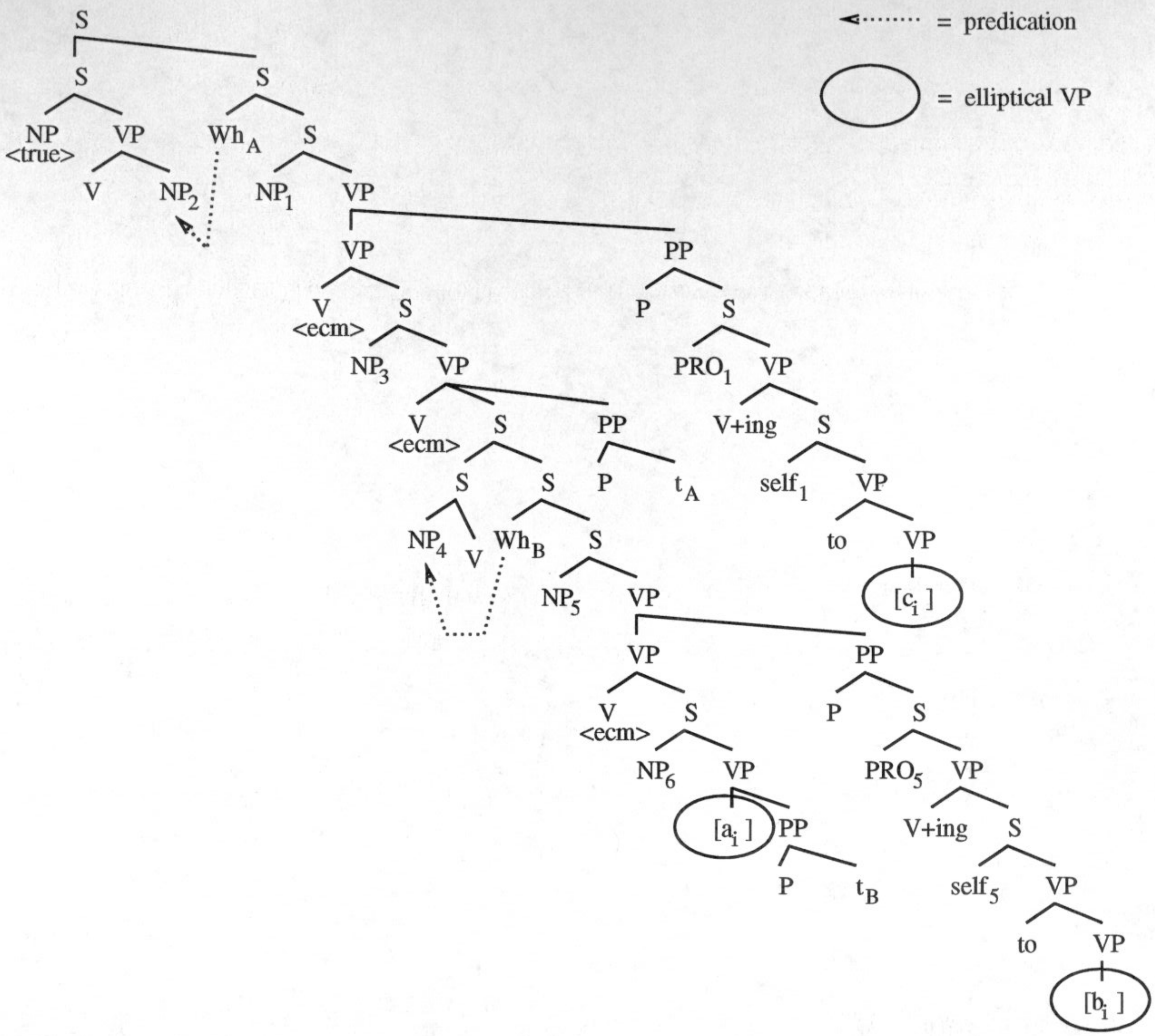


Figure 1: This phrase structure corresponds to the ith Boolean clause $C_i = (a_i \vee b_i \vee c_i)$, with irrelevant details suppressed. "S" is a clause, "NP" is a noun phrase, "VP" is a verb phrase, and "PP" a prepositional phrase. All NPs dominate pronouns. Dashed arrows depict the predication of a noun phrase by an extraposed relative clause. The structure contains configurations of local c-command, strong crossover, adjunct control, and invisible obviation, yielding the graph of referential dependencies shown in figure 2. Each of the ellipsed VPs ($[a_i]$, $[b_i]$, and $[c_i]$) refers to one of the overt VPs in the n literal constructions (18), as described above. Consequently, NP_6 and the pronoun that represents the literal a_i are in a relation of invisible obviation, as are NP_5 and the pronoun for b_i, and NP_1 and the pronoun for c_i. This is shown in figure 2. This phrase structure would be assigned to natural utterances such as *Hector met him$_2$, who$_A$ he$_1$ expected him$_3$ to want him$_4$ framed, who$_B$ he$_5$ believed he$_6$ did [e] with t_B after exposing himself$_5$ to [e] for t_A, before telling himself$_1$ to [e].* (Indices, traces of wh-movement, and elliptical VPs ("[e]") are included in the expression solely to help the reader align the expression with its phrase structure.)

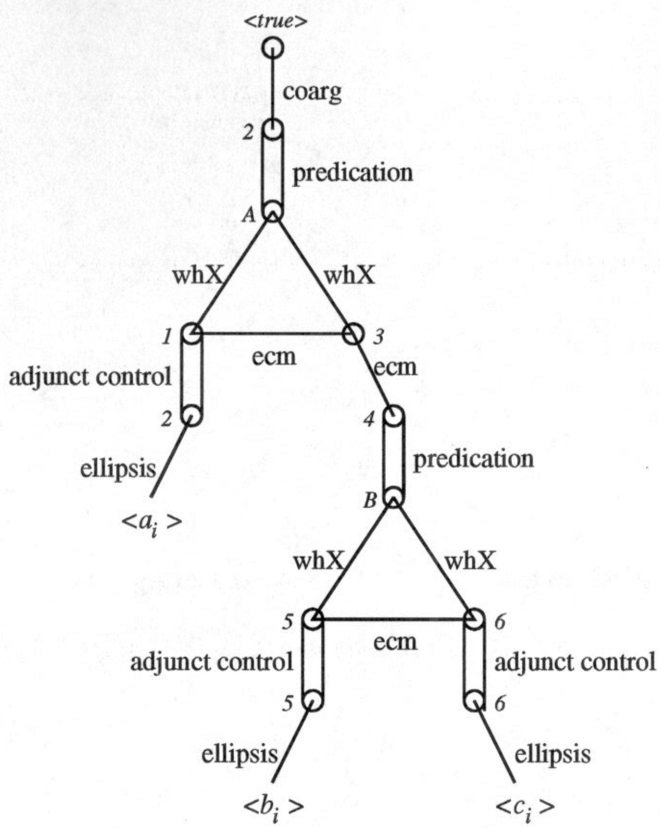


Figure 2: This is the graph of referential dependencies for the phrase structure of figure 1 (and the natural utterance to which it is associated). Single lines depict obviation relations; double lines depict relations of coreference, that is, links and predications. Vertices in the graph correspond to noun phrase arguments, and are labeled with the identifying indices from figure 1. The edges in the graphs are likewise labeled with the configurations to which they are attributable. Recall that "coarg" and "ecm" are configurations of local c-command; "whX" is for a strong crossover involving a wh-phrase; and "ellipsis" for the invisible obviation arising from the ellipsis of a verb phrase representing the relevant Boolean literal from (18). The obviation graph that results when coreferential vertices are coalesced, in combination with the three available antecedents, ensures that at least one of pronouns representing a_i, b_i, or c_i must link to the proper noun *True*. This corresponds exactly to the constraint that a true Boolean 3-clause contain at least one true literal.

suggests that the anaphora problem is one of the more difficult subproblems of language comprehension, because graph k-coloring is one of the most difficult NP-complete problems, with no known polynomial-time approximation scheme.

4 Evidence for an $\mathcal{NP}$ upper bound

Lacking a complete, scientifically plausible linguistic theory, it is not possible to prove an upper bound on the complexity of human language. It is however possible to provide empirical evidence for an upper bound, and that is the task of this section. The argument goes as follows. First, we examine the linguistic phenomenon of ellipsis, and present empirical arguments for a simple "copy theory" of the knowledge that language users have about such phenomenon. Next, we use this theory to prove that the anaphora problem is PSPACE-hard. Using the insights of the complexity proof, we reexamine the phenomenon of ellipsis, exhibit counterexamples to the copy theory, and suggest an descriptively superior predicate-sharing theory. Finally, we argue that the anaphora problem is in $\mathcal{NP}$ according to the predicate-sharing theory. By reducing the complexity of the anaphora problem in elliptical structures to inside $\mathcal{NP}$, while strictly improving the empirical adequacy of the theory of ellipsis, we provide evidence for an $\mathcal{NP}$ upper bound on the complexity of the anaphora problem.

In developing our simple linguistic theories, we will briefly introduce the relevant phenomenon, state the theory, and conclude with a concise enumeration of the empirical arguments in support of the theory.

4.1 Copy theory of ellipsis

A central goal of linguistics is to explicitly represent the knowledge that language users have about utterances. Let us *purely as a matter of convenience* distinguish the representation of how an utterance expressed in words and phrases, from a representation of the logical aspects of its meaning, such as referential dependencies and predication. Let us call the former representation the *surface form*, and the latter representation, the *logical form*. (The number of levels of representation does not in and of itself affect the computational complexity.)

Theory 1 *The logical form of ellipsis is constructed by (recursively) copying the overt structure into the position of the corresponding ellipsed structure;*

an anaphoric element α may link to its antecedent either before or after copying; when the antecedent of α is a quantified NP β, then α must link to β after copying.

Evidence. First, the elliptical structure is understood as though it were really there, as shown in (19).

(19) a. The men [ate dinner] and the women did [e] too.
  b. 'the women did *eat dinner* too'.

This fact about our linguistic knowledge must be represented somehow in the logical form, perhaps by copying the overt structure into the position of the null structure, as first suggested by Chomsky (1955).

Second, an elliptical structure may itself be understood as containing an elliptical structure, as in (20a), which is understood to mean (20b).

(20) a. Jack [[corrected his spelling mistakes]$_1$ before the teacher did [e]$_1$]$_2$ and Ted did [e]$_2$ too.
  b. Jack corrected his spelling mistakes before the teacher did *correct his spelling mistakes* and Ted did *correct his spelling mistakes before the teacher did correct his spelling mistakes.*

This suggests that copying is a recursive process. The depth of recursion does not appear to be constrained by the principles of grammar, as shown in (21):

(21) Harry [claims that Jack [[corrected his spelling mistakes]$_1$ before the teacher did [e]$_1$]$_2$ and that Ted did [e]$_2$ too]$_3$, but Bob doesn't [e]$_3$.

Third, the elliptical structure behaves as though it was really there. In particular, it can induce a violation of obviation, as in the discourse (22):

(22) Ann: Romeo$_1$ wants Rosaline to [love him$_i$] ($i = 1$)
Ben: Not any more—now Rosaline wants Romeo$_1$ to [e] ([*love him$_i$*], $i \neq 1$)

In this example, Ann's use of the pronoun *him* is most naturally understood as referring to *Romeo.* Yet when Ben replies, the coreferential interpretation ($i = 1$) is no longer possible in Ann's statement. These facts of invisible obviation are easily accounted for if the overt structure is in fact copied in the syntax, as illustrated in (23), where the obviation violation between *him*$_1$ and *Romeo*$_1$ has been made explicit by copying the overt VP *love him* into the position of the null VP.

(23) Rosaline wants [Romeo$_1$ to *love him*$_{*1}$]

The invisible structure is not merely an invisible VP-pronoun, simply because the obviation violation in (24a) vanishes when an overt pronoun is used instead in (24b).

(24) a. Juliet$_1$ thought the Friar$_2$ [poisoned her$_1$] without realizing that she$_{*1}$ did [*e*].
b. Juliet$_1$ thought the Friar$_2$ [poisoned her$_1$]$_3$ without realizing that she$_1$ did it$_3$.

Fourth, corresponding anaphoric elements in the overt and invisible structures may be understood as having different antecedents, as in (25), where the invisible pronoun *his* is ambiguous, referring either to *Felix* ('invariant' interpretation) or *Max* ('covariant' interpretation). (In each example, careful attention must be paid to the relevant construal of the null structure, indicated with brackets, and the intended reference of anaphoric elements, as indicated in the italicized parenthetical following the example.)

(25) Felix$_1$ [hates his$_1$ neighbors] and so does Max$_2$ [*e*].
([*hates his*$_{1/2}$ *neighbors*])

This suggests that an anaphoric element may be linked to its antecedent either <u>before</u> the overt structure is copied, resulting in the invariant interpretation (26), or <u>after</u>, resulting in the covariant interpretation (27).

(26) a. Felix$_1$ [hates his$_1$ neighbors] and so does Max$_2$ [*e*].
b. Felix$_1$ [hates his$_1$ neighbors] and so does Max$_2$ [hate his$_1$ neighbors].

(27) a. Felix$_1$ [hates his neighbors] and so does Max$_2$ [*e*].
b. Felix$_1$ [hates his$_1$ neighbors] and so does Max$_2$ [hate his$_2$ neighbors].

Fifth, the invisible pronoun must agree with its antecedent, which excludes the covariant interpretations in (28) that are possible in the minimally different examples in (29).

(28) a. Barbara$_1$ read her$_1$ book and Eric$_2$ did [*e*] too.
([*read her*$_{1/*2}$ *book*])
b. You$_1$ ate your$_1$ vegetables and so did Bob$_2$ [*e*].
([*eat your*$_{1/*2}$ *vegetables*])

(29) a. Barbara$_1$ read her$_1$ book and Kate$_2$ did [*e*] too.
([*read her*$_{1/2}$ *book*])
b. You$_1$ ate your$_1$ vegetables and so did you$_2$ [*e*].
([*eat your*$_{1/2}$ *vegetables*])

Sixth, the covariant interpretation is forced when the antecedent of the anaphoric element is a quantified noun phrase (QNP), as shown in (30).

(30) Every man$_1$ [ate his$_1$ dinner] and so did every boy$_2$ [*e*]
([*eat his*$_{*1/2}$ *dinner*])

That is, (30) must mean that every boy ate his own dinner; it cannot mean that every boy ate every man's dinner. Therefore, an anaphoric element must be linked to its antecedent β after copying when β is a quantified noun phrase.

To summarize, we have seen evidence that the overt structure must be copied to the position of null structure in the syntax, that copying is a recursive process, and that anaphoric elements may be linked to their antecedents either before or after the copying, and that they must be linked after copying when their antecedent is a quantified noun phrase. □

The first account of invariant and covariant interpretations in VP ellipsis, due to Ross (1967), is equivalent to the copy theory, because deletion in Ross's deep-structure to surface-structure derivation is identical to copying in the surface form to logical form mapping. This copy theory has also been proposed in recent linguistics literature. See for example, Pesetsky (1982), May (1985), Koster (1987), and Kitagawa (1989).

More generally, any linguistic theory that represents the meaning of an elliptical utterance using devices that can achieve the effect of copy and link operations will inherit the complexity of the copy theory. This is true regardless of how that linguistic theory is defined, how many levels of representation it has, or what they are called.

4.2 Complexity of anaphora in the copy theory

In this work, we only consider the problem of assigning linguistic representations to utterances, which is a trivial subproblem of the much more intractable and less well understood problem of determining the semantic 'truth value' of a given utterance. The following proof shows that assigning a complete linguistic representation to a given class of utterances can be as difficult as determining the truth of quantified Boolean formulas; the proof does *not* make the entirely unnourishing argument that determining the 'truth value' of human language utterances can be as difficult as determining the truth of quantified Boolean formulas.

Lemma 4.1 *The anaphora problem is PSPACE-hard in the copy theory.*

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 a surface form S and a set A of available antecedents, such that all the anaphoric elements in S have antecedents in A if and only if Ω is true. In order to verify that all anaphoric elements in S have antecedents, we must construct the logical form of S. The reduction uses one binary agreement feature to represent literal negation, and one n-valued agreement feature (or equivalently, $\log_2 n$ binary agreement features) to identify the n distinct Boolean variables.

The idea of the proof is to mimic the structure of Ω with linguistic constructions, by reducing the quantification of variables in Ω to the linking of pronouns in S. Each quantifier Qx in Ω will correspond to a pair of available antecedents in S, one to represent $x = 0$ and the other to represent $x = 1$. Boolean literals in Ω will correspond to pronouns in S. As shown in figure 3, the surface form S is built from three distinct components: universal quantifiers, existential quantifiers, and Boolean clauses. We will now motivate each of these parts in turn, using intricate yet still natural English sentences.

The first step is to simulate a universal quantifier. Recall that a universally quantified predicate $[\forall x_i P(x_i)]$ is true if and only if $[P(x_i = 0) \wedge P(x_i = 1)]$. The latter Boolean formula can be expressed in a VP-ellipsis construction whose surface form is abstracted in (31).

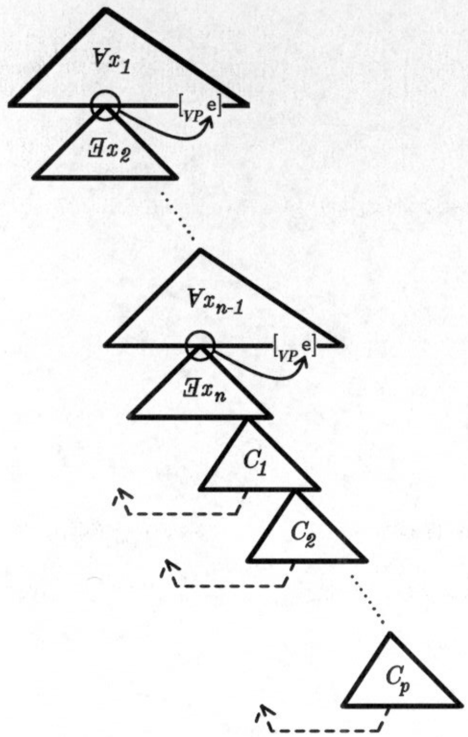
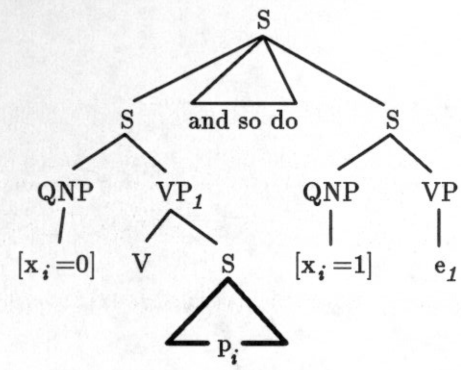


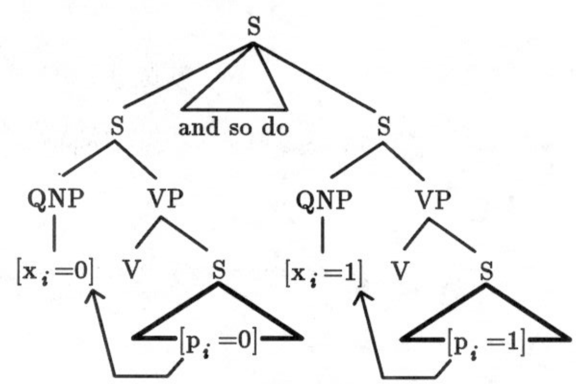
Figure 3: The surface form S that corresponds to the input instance $\Omega = \forall x_1 \exists x_2 \ldots \forall x_{n-1} \exists x_n [C_1, C_2, \ldots C_p]$. The quantifier constructions contain two antecedents to represent the two possible truth assignments to the quantified variable. Each universal quantifier $\forall x_i$ is represented by a VP-ellipsis template. In the logical form that corresponds to S, each of the $n/2$ circled overt VPs is copied to its corresponding ellipsed VP position $[_{\mathrm{VP}} e]$, according to the copy theory. Each existential quantifier $\exists x_{i+1}$ is represented by an extraposed strong crossover template, as discussed in the text. Each clause C_j is represented by a pigeonhole construction that contains three pronouns, one for each literal in C_j; one of these pronouns (the *selected* pronoun) must link to an antecedent outside that construction, in some dominating quantifier construction. These obligatory long distance links are drawn with dashed arrows. The selected pronouns represent the literals that satisfy the clauses.

(31)



According to the copy theory, the language user's knowledge of the construction (31) is represented in the abstracted logical form (32). First, the overt VP is copied to the position of the null VP. Next, pronouns inside the original and copied VPs link to their antecedents independently.

(32)

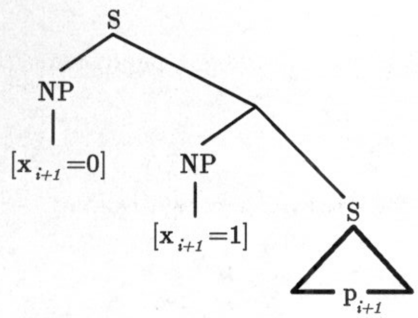


The VP is used in the reduction to represent the Boolean predicate $P(x_i)$; the embedded pronoun p_i represents a true literal of x_i inside P; the two QNP subjects represent the truth values $x_i = 0$ and $x_i = 1$, respectively. Each p_i must link to the subject of its own conjunct in the logical form, because the subjects are quantified noun phrases. Therefore the pronoun p_i in the first VP may only link to the first subject $[_{\text{QNP}}\ x_i = 0]$, which represents the conjunct $P(x_i = 0)$, and the pronoun p_i in the second (copied) VP may only link to the second subject $[_{\text{QNP}}\ x_i = 1]$, which represents the conjunct $P(x_i = 1)$. As shown in figure 3 above, the verb phrase will also contain the construction (33) that represents the next quantifier $\exists x_{i+1}$.

The second step is to simulate an existential quantifier. An existentially quantified predicate $[\exists x_{i+1} P(x_{i+1})]$ is true if and only if $[P(x_{i+1} = 0) \vee$

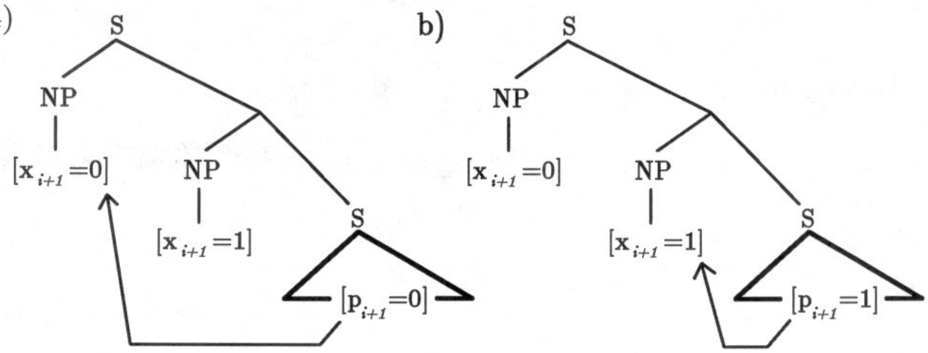
$P(x_{i+1} = 1)]$. The latter Boolean formula can be expressed in a construction whose surface form is (33).

(33)



This structure will have two possible meanings, as represented by the two logical forms in (34):

(34) a)



The embedded sentence represents the predicate $P(x_{i+1})$; the embedded pronoun p_{i+1} represent a true literal of x_{i+1} inside the predicate P; the two NPs represent the truth values $x_{i+1} = 0$ and $x_{i+1} = 1$, respectively. Linguistic constraints to be disclosed below ensure that p_{i+1} can only be linked to one of the two noun phrases, and that p_{i+1} can be linked to the first NP $[_{\mathrm{NP}}\ x_{i+1} = 0]$ if and only if $P(x_{i+1} = 0)$ is true, as shown in (34a); and that p_{i+1} can be linked to the second NP $[_{\mathrm{NP}}\ x_{i+1} = 1]$ if and only if $P(x_{i+1} = 1)$ is true, as shown in (34b). The embedded clause will also contain the construction (31) that represents the next quantifier $\forall x_{i+2}$, as shown in figure 3.

In order to ensure consistency of truth assignments, all embedded pronouns representing true literals of x_{i+1} must link to the same antecedent. This constraint may be enforced using the powerful strong crossover configuration introduced in the previous section. The details of how this might be done arose from discussion with Alec Marantz, who suggested all the examples.

Recall that strong crossover is the configuration where an anaphoric element α intervenes between a displaced wh-phrase and its trace, and c-commands the trace. In such a configuration, α obviates the subject of the wh-phrase.

(35) a. Who$_k$ did he$_{*k}$ say Mary kissed t_k.
b. [the man]$_1$ [who$_k$ he$_{*1}$ likes t_k].

The noun phrase in (35b) contains a relative clause [*who he likes t*] that predicates [*the man*]; the pronoun *he* is in a strong crossover configuration, and therefore cannot refer to [*the man*], which is the subject of the relative clause.

Now consider the effect of extraposing a relative clause containing a strong crossover configuration in (36).

(36) a. At the airport, a man$_1$ met Jane$_2$, who$_{k=1/*2}$ she$_2$ likes t_k.
b. At the airport, a man$_1$ met Jane$_2$, who$_{k=*1/2}$ he$_1$ likes t_k.

In (36a), if we understand *she* as referring to *Jane*, then we must understand *who* as predicating *a man*. Conversely, if we understand *he* as referring to *a man* in (36b), then *who* must predicate *Jane*. This simple example establishes the ambiguity of predication when the predicate is an extraposed relative clause containing a strong crossover configuration.

When the extraposed relative clause contains two obviative pronouns, as in (37), then the sentences cannot have the intended interpretation because the relative clause must predicate some subject, yet cannot without violating strong crossover.

(37) a. *At the airport, [a man]$_1$ met Jane$_2$, who$_k$ she$_2$ thinks he$_1$ likes t_k.
b. *At the airport, [a man]$_1$ met Jane$_2$, who$_k$ he$_1$ thinks she$_2$ likes t_k.

This example establishes that the strong crossover configuration gives rise to inviolable obviation between the wh-phrase and all embedded pronouns that c-command its trace.

Now we have our construction:

(38) At the airport, NP$_0$ met NP$_1$, [who$_k$ $\ldots \alpha_{*k} \ldots$ t_k].

As before, two possible antecedents NP_0 and NP_1 represent the truth assignments $x_{i+1} = 0$ and $x_{i+1} = 1$, respectively. Pronouns in the embedded clause that represent true negative literals of x_{i+1} can only link to the 'false' noun phrase NP_0; pronouns that represent true positive literals of x_{i+1} can only link to the 'true' noun phrase NP_1. Observe that the relative pronoun *who*$_k$ may predicate either NP_0 or NP_1 in the example (38). The strong crossover configuration ensures that an anaphoric element α in the extraposed relative clause must obviate the subject of the wh-phrase *who*$_k$. Therefore, once the ambiguous predication relation is determined, pronouns representing literals of x_{i+1} must all be linked to the same antecedent because (i) the pronouns must all obviate the predicated noun phrase by strong crossover and (ii) there is only one other permissible antecedent by construction. This exactly corresponds to assigning a consistent truth value to x_{i+1} everywhere.

The third and final step of the reduction is to simulate a Boolean 3-clause $C_j = (a_j \vee b_j \vee c_j)$ using the pigeonhole principle. A Boolean clause C_j is true if and only if one of its literals it true: let us call the literal that satisfies the clause the *selected literal.* Only selected literals need be assigned consistent truth values: nonselected literals simply don't matter, and can receive any arbitrary inconsistent value, or none at all. We have been reducing the quantification of variables to the binding of pronouns, and so must now represent each literal in C_j with a pronoun. For each 3-clause, the reduction builds a sentence that contains three disjoint pronouns and only two possible antecedents. At least one of the pronouns must link to an antecedent outside the sentence—this pronoun represents the selected literal. The following English sentence shows how this works:

(39) [$_S$ [the student] thought [the teacher] said that
[he$_a$ introduced her$_b$ to him$_c$]]

Only two neutral antecedents [*the student*] and [*the teacher*] are locally available to the three obviative pronouns *he*$_a$, *her*$_b$, and *him*$_c$ in this construction. Therefore at least one of these three pronouns must link to an available antecedent outside the clause, to one of the noun phrases in some dominating quantifier construction (either (31) or (33)). This selected pronoun corresponds to a true literal that satisfies the clause C_j. Agreement features on pronouns and their antecedents ensure that a pronoun representing a literal of x_i can only link to an antecedent representing the quantifier of x_i.

Note that this construction is contained inside $n/2$ VP-deletion constructions in the surface form of the entire sentence S, and that therefore the corresponding logical form will contain $2^{n/2}$ copies of each such construction, each copy with its own selected pronoun. (This corresponds to the fact that different literals may satisfy a given quantified clause, under different quantifier-determined truth assignments.) The verb phrase that appears in our English example (39) as [*he introduced her to him*] will immediately contain the construction representing the next Boolean clause C_{j+1}, as shown in figure 3.

The pigeonhole construction representing C_j is permissible iff all of its logical form copies are all 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 quantifier constructions (either (31) or (33)). Therefore, the logical form for the complete surface form S is permissible iff the quantified Boolean formula Ω is true. $\square$

Note that the constructions used in this proof to represent existential quantifiers (33) and Boolean clauses (39) can be combined to give a third direct NP-hardness proof for the anaphora problem, where each pronoun is no more than four-ways ambiguous and no elliptical contexts are used. Such a proof requires significantly fewer agreement features than used in the proof of lemma 3.1.

The epilogue to this proof is a demonstration of how the preceding reduction might concretely represent the QBF formula (40) in an actual English sentence:

(40) $\forall x \exists y[(\overline{y} \vee \overline{x} \vee y), (x \vee \overline{x} \vee y)]$

There are two minor difficulties, that are entirely coincidental to the English language: the English plural pronoun *they* is unspecified for gender; there are no entirely neutral arguments in English, that can be the antecedent of any pronoun. Rather than construct our example in a different language, say Italian, let us make the following allowances.

To overcome the first difficulty, let *they*$_0$ be the masculine plural pronoun, and *they*$_1$ the feminine plural pronoun.

To overcome the second difficulty, we observe that a plural pronoun can always have a split antecedent, as in example (41), and that the condition of local obviation holds between *they* and *him*. That is, *they* and *him* cannot share an antecedent when they are locally obviative.

(41) John$_1$ suggested to Tom$_2$ that they$_{\{1,2\}}$ tell him$_{*1/*2}$ to leave.

We will use split antecedents below.

The given formula (40) has two variables, x and y, which we will identify via the plural/singular number distinction: plural pronouns represent literals of x, while singular pronouns represent literals of y. Negation will be represented via the the masculine/feminine gender distinction: masculine pronouns for negative literals, feminine pronouns for positive literals. These correspondences are summarized in the table:

$$\begin{array}{r|l} \overline{x} \mapsto \textit{they}_0 & \overline{y} \mapsto \textit{he}_0 \\ x \mapsto \textit{they}_1 & y \mapsto \textit{she}_1 \end{array}$$

The constructed sentence consists of four components:

- The VP-ellipsis construction (31) to represent $\forall x$:

  (42) [[$_{\mathrm{NP}_0}$ some stewards] [$_{\mathrm{VP}}$ say [$_{\mathrm{S}}$ ...]]]
  and so do [[$_{\mathrm{NP}_1}$ some stewardesses] [$_{\mathrm{VP}}$ e]]

- The extraposed strong crossover configuration (38) to represent $\exists y$:

  (43) [$_{\mathrm{S}}$ at the airport [$_0$ a KGB man] met [$_1$ Jane], [$_{\mathrm{S}}$ who$_k$ [... t_k] and [... t_k]]

- The pigeonhole construction (39) to represent $(\overline{y} \vee \overline{x} \vee y)$ using split antecedents.

  (44) [$_{\mathrm{S}}$ the officer, the agent, and the mechanic suspected
  [he$_0$ expected them$_0$ to talk to her$_1$]]

  There are three locally available antecedents, all singular and unspecified for gender. The three pronouns in the embedded clause are obviative, and require at least four singular antecedents. Therefore, at least one of the pronouns must be linked to an argument outside the construction (44).

- A second pigeonhole construction to represent $(x \vee \overline{x} \vee y)$, again using split antecedents.

  (45) [$_{\mathrm{S}}$ the crew, the pilot, and the co-pilot knew [they$_1$ traded them$_0$ to her$_1$]]

There are three locally available antecedents: one is plural neuter (*the crew*), and the remaining two are singular neuter. The three pronouns in the embedded clause are obviative, and require at least one plural antecedent and three singular antecedents. Therefore, at least one of the pronouns must be linked to an argument outside the construction (45).

The resulting sentence, in all its glory, is:

(46) [[$_{\text{NP}_0}$ some stewards] [$_{\text{VP}}$ say
[$_{\text{S}}$ at the airport [$_0$ a KGB man] met [$_1$ Jane], [$_{\text{S}'}$ who$_k$
[$_{\text{S}}$ the officer, the agent, and the mechanic suspected
[he$_0$ expected them$_0$ to talk to her$_1$ about t_k]]
and
[$_{\text{S}}$ the crew, the pilot, and the co-pilot knew
[they$_1$ traded them$_0$ to her$_1$ for t_k]]]]]]
and so do
[[$_{\text{NP}_1}$ some stewardesses] [$_{\text{VP}}$ e]]

This concludes the presentation of the lemma 4.1.

4.3 Ellipsis reconsidered

In the previous section, we proved that the anaphora problem is PSPACE-hard according to the copy theory. The thesis we are defending states that language comprehension is NP-complete. Therefore, the thesis predicts that there is a empirical defect in the linguistic analysis that led to the PSPACE-hardness result. The thesis also tells us exactly where to look for the defect: we must reexamine that part of the analysis that allowed us to simulate a computation outside of $\mathcal{NP}$. In the case of a reduction from QBF, the defect must be in that part of the analysis used to simulate the unnaturally powerful universal quantifier. Therefore, let us reexamine the copy theory of ellipsis.

A copy operation naturally makes two predictions; neither holds.

The first prediction is that the original (overt) structure and its copy will obey the same post-copying linguistic constraints, including agreement and the linking conditions. (If agreement and the linking conditions did not apply after copying, then it would always be possible to vacuously satisfy those constraints, simply by postponing all linking until after copying had applied. Therefore, agreement and the linking conditions must apply both before and after copying.) This expected post-copying equivalence is violated. Although overt pronouns must agree with their antecedent on gender and number (47a), copied pronouns can disagree with their antecedents, as in (47b):[7]

(47) a. Tom_1 read $\text{his}_{1/*2}$ book and Barbara_2 read $\text{his}_{1/*2}$ book (too).

b. Tom_1 [read his_1 book] and Barbara_2 did [*e*] too.
([*read* $\textit{his}_{1/2}$ *book*])

Moreover, although overt reflexives must have local antecedents in (48a), copied reflexives need not, as shown in (48b):

(48) The prisoner_1 hung himself_1 before [the executioner_2 could hang $\text{himself}_{*1/2}$].

a. The prisoner_1 [hung himself_1] before the executioner_2 could [*e*].
([*hang* $\textit{himself}_{1/2}$])

The second prediction is that processes that apply after copying, such as linking, will apply independently in both the original (overt) structure and its copy. This expected post-copying independence is also violated. In particular, linking is not independent in both the original structure and its copy, as shown by example (49), which is incorrectly predicted to have five readings that introduce no new information (two when linking precedes copying, four when linking follows copying, and one overlap).

(49) Bob [introduced Felix to his neighbors] and so did Max [*e*].

In particular, there should be a reading where the overt *his* refers to *Felix* and the null/copied *his* refers to *Max*. However, this reading is not available. In fact, only three interpretations of (49) are attested (two invariant, one covariant), as shown in (50):

(50) a. Bob_1 [introduced Felix_2 to his_2 neighbors] and so did Max_3 [*e*].
([*introduced* $\textit{Felix}_2$ *to* $\textit{his}_{*1/2/*3}$ *neighbors*])

b. Bob_1 [introduced Felix_2 to his_1 neighbors] and so did Max_3 [*e*].
([*introduced* $\textit{Felix}_2$ *to* $\textit{his}_{1/*2/3}$ *neighbors*])

[7] The difficulty some speakers have in obtaining the covariant interpretation for *$\textit{Barbara}_1$ read $\textit{her}_1$ book and Eric did too*, or for *$\textit{You}_1$ ate $\textit{your}_1$ vegetables and so did Bob*, does not weaken my criticism of the copy theory. My criticism is based on the necessity of discriminating (47a) and (47b), which the copy theory is unable to do.

In other words, a pronoun must link to the same position in both the visible verb phrase and its understood copy. This is not real copying, but a kind of logical predicate-sharing that can always be represented without copying.

That is, we should think of a verb phrase as a kind of logical predicate, that is applied to a logical argument (the grammatical subject), resulting in a logical proposition (the sentence). In VP ellipsis then, the overt VP predicate is *shared* between two propositions. That is, it is applied to the subject of the invisible VP, as well as to the subject of the overt VP.

Continuing in this vein, let us think of a verb as a higher-order logical function that contains a sequence of thematically-typed argument positions. For example, the English verb *see* is a logical function, from an argument of thematic type "patient" (the object seen), to a predicate of an argument of thematic type "agent" (the entity seeing). So the indirect object, the direct object, and the subject of a sentence are logical arguments to the main verb of that sentence; each argument is assigned a thematic type by the argument position that it saturates in the verb's logical function.

This line of thinking leads to the following predicate-sharing theory of ellipsis:

Theory 2 *The logical form of ellipsis is constructed by sharing the same logical predicate between the overt and ellipsed structures; obviation is a relation between argument positions in a logical predicate; an anaphoric element may link to an argument or to an argument position.*

Evidence. First, verbs are logical functions from their direct and indirect objects to a predicate; the predicate corresponding to a verb phrase is function from the subject to a logical proposition, that is realized syntactically as a clause. For example, the expression *Felix hates vegetables* would be assigned the logical form (51), after lambda-abstracting the predicate ($\lambda x.$[x *hates vegetables*]).

(51) ($\lambda x.$[x hates vegetables]) ($\lambda P.$[(Felix P)])

Second, VP ellipsis is the sharing of one VP predicate between two clauses. One way to represent the logical form of an elliptical structure to lambda-abstract the VP predicate. For example, the surface form (52a) would be assigned the logical form representation (52b), with the VP predicate ($\lambda x.$[x *ate dinner*]) applied "in parallel" to both subjects.

(52) a. [Felix [ate dinner]] and so did [Tom [e]]
b. ($\lambda x.$[x ate dinner]) ($\lambda P.$[(P Felix) and so did (P Tom)])

Third, obviation is a relation between the argument positions in the VP predicate, as illustrated in (53b) for the surface form (53a).

(53) a. Romeo$_1$ wants Rosaline to [love him$_{*1}$] before wanting himself$_1$ to [e].

b. $(\lambda x_i.[x_i$ to love him$_{*i}])$ $(\lambda P.[$Romeo$_j$ wants $[($Rosaline $P)]$ before wanting $[($himself$_j$ $P)]]) \Rightarrow [i \neq j]$

This logical form representation accounts for all the cases of invisible obviation, without an unnaturally powerful copying operation.

Fourth, an anaphoric element may link to an argument directly (54b), resulting in the invariant interpretation, or indirectly, to an argument position in the VP predicate (54c), resulting in the covariant interpretation.

(54) a. Felix$_1$ [hates his$_1$ neighbors] and so does Max [e]

b. $(\lambda x.[x$ hates his$_i$ neighbors$])$ $(\lambda P.[($Felix$_i$ $P)$ and $($Max $P)])$

c. $(\lambda x_i.[x_i$ hates his$_i$ dinner$])$ $(\lambda P.[($Felix $P)$ and $($Max $P)])$

□

This predicate-sharing theory correctly predicts that the example (49) has exactly three interpretations, one for each of the three possible verbal predicates shown in (55).

(55) a. $(\lambda x.[x$ introduced Felix to Felix's neighbors$])$ $(\lambda P.[(P$ Bob$)$ and $(P$ Max$)])$

b. $(\lambda x.[x$ introduced Felix to x's neighbors$])$ $(\lambda P.[(P$ Bob$)$ and $(P$ Max$)])$

c. $(\lambda x.[x$ introduced Felix to Bob's neighbors$])$ $(\lambda P.[(P$ Bob$)$ and $(P$ Max$)])$

While predicate-sharing is conceptually simple, a detailed empirical investigation is needed to confirm such a linguistic theory. This is the task of appendix B in Ristad (1990b).

Recall that it is not possible to *prove* an upper bound on the complexity of the anaphora problem without a comprehensive empirically-plausible formal model of human language. We can, however, accumulate evidence for our conjectured $\mathcal{NP}$ upper bound. The first piece of evidence is that the proof of lemma 4.1 is not possible with the predicate-sharing theory of ellipsis. Moreover, the following argument, somewhere between a proof and a conjecture, also gives us the $\mathcal{NP}$ upper bound predicted by the complexity thesis:

Conjecture 2 *The anaphora problem is in $\mathcal{NP}$ for nonelliptical structures, and for elliptical structures with predicate-sharing.*

Evidence. Covert arguments in a linguistic representation are either coreferential with an overt argument in the representation (for example, control PRO or wh-trace), in which case their corresponding vertices in the graph of referential dependencies may be coalesced, or they are assigned an arbitrary interpretation (PRO_{arb}), in which case they do not participate in the graph of referential dependencies and may be ignored entirely. Therefore, the number of obviation relations in a linguistic representation is at most quadratic in the number of overt arguments, an upper bound that is obtained in the case of a complete obviation graph.

Moreover, each obviation relation may be easily computed from an utterance's phrase structure representation. This is trivially so for the configurations considered above in section 3.2, as well as for all other configurations that the author is aware of.

Next, the logical forms licensed by the predicate-sharing theory are nearly the same size as their corresponding surface forms, because we can always lambda-abstract the shared predicate, if the structure is elliptical. Otherwise, the structure is nonelliptical and logical and surface forms are the same size, because operators that map surface forms to logical forms, such as quantifier scope assignment, do not increase the number of arguments and therefore cannot increase the size of the graph of referential dependencies beyond quadratic.

Each anaphoric element may be nondeterministically linked either to an argument in the set A of available antecedents, or to an open argument position. Clearly this may be done in nondeterministic polynomial time. Finally, we check that the linking conditions are satisfied, including invisible obviation, in deterministic time proportional to the number of links, verifying the semantics of obviation relations by propagated "referential value" markers along the links, checking for cyclic dependencies, and so forth. □

Conjecture 3 *The anaphora problem is NP-complete in the predicate-sharing theory.*

Evidence. By conjecture 2 and theorem 1. □

As can be seen from section 4.3, and in greater detail from Ristad (1990b), the predicate-sharing theory is strictly superior to the copy theory.

That is, the predicate-sharing theory assigns empirically superior representations to the class of elliptical utterances than the copy theory does, and no utterances are assigned better representations by the copy theory. However, the significance of the predicate-sharing theory goes beyond merely the number of additional linguistic examples correctly represented.

Recall that our central scientific goal is to understand the comprehension, production, and acquisition of human language; modern linguistic theory is interesting only in so far as it advances this goal. The solution to a PSPACE-hard problem may be exponentially large in the size of the problem statement. If anaphora comprehension were PSPACE-hard, as it is according to the copy theory of ellipsis, then the mental representations required to produce and comprehend elliptical anaphora could be infeasibly large. Language users could not even comprehend the utterances that they themselves produced. And current linguistic theories of anaphora would not yield a plausible account of language comprehension and production.

But by reducing the complexity of the anaphora problem from PSPACE to $\mathcal{NP}$, we argue that the anaphora problem has efficient witnesses, and in turn show that current generative theory is the basis of a plausible account of anaphora comprehension and production.

To summarize, in this essay we proved that two distinct models of the anaphora problem are NP-hard. Next, we argued that including the computationally complex phenomenon of syntactic ellipsis in the referential dependence model does not increase the complexity of the anaphora problem outside of $\mathcal{NP}$. This is all by way of demonstrating a novel, largely theory-neutral approach to the mathematical analysis of human knowledge.

5 References

Barton, E., R. Berwick, and E. Ristad. (1987), "Computational complexity and natural language," Cambridge: MIT Press.

Canny, J. (1988), "The Complexity of Robot Motion Planning," MIT Press, Cambridge.

Chomsky, N. (1955), "The Logical Structure of Linguistic Theory," mimeographed, Harvard University. (Published in part by Plenum Press, 1975).

Chomsky, N. (1956), Three models for the description of language, *I.R.E. Transations on Information Theory*, Vol. IT-2, pp. 113–124. Reprinted, with corrections, in R.D. Luce, R. Bush, and E. Galanter (eds.), "Readings in Mathematical Psychology," Vol.II. New York: Wiley, 1965.

Chomsky, N. (1986), "Knowledge of Language: Its Origins, Nature, and Use," Praeger Publishers, New York.

Hopcroft, John and Jeffrey Ullman. 1979. "Introduction to Automata Theory, Languages, and Computation," Reading, MA: Addison-Wesley.

Higginbotham, J. (1983), Logical form, binding, and nominals, *Linguistic Inquiry* 14:395–419.

Higginbotham, J. (1985), On semantics, *Linguistic Inquiry* 16:547–593.

Kitagawa, Y. (1989), Copying identity, ms., Department of Linguistics, University of Massachusetts at Amherst.

Koster, J. (1987), "Domains and Dynasties: the Radical Autonomy of Syntax," Foris Publications, Dordrecht.

Lawler, E.L. (1976), A note on the complexity of the chromatic number problem, *Information Processing Letters* 5:66–67.

May, R. (1985), "Logical Form: Its Structure and Derivation," MIT Press, Cambridge.

Pesetsky, D. (1982), "Paths and Categories," Unpublished Ph.D dissertation, MIT Department of Linguistics and Philosophy.

Peters, S., and Ritchie, R.W. (1973), On the generative power of transformational grammars, *Information Science* 6:49–83.

Pullum, G. (1983), How many possible human languages are there?, *Linguistic Inquiry* 14(3):447–468.

Reif, J. (1979), Complexity of the mover's problem and generalizations, *in* "Proceedings, 20th Annual Symposium on Foundations of Computer Science," pp.421–427, IEEE Computer Society, New York.

Reif, J. and Sharir, M. (1985), Motion planning in the presence of moving obstacles, *in* "Proceedings, 26th Annual Symposium on Foundations of Computer Science," pp.144–154, IEEE Computer Society, New York.

Ristad, E. (1990a), Computational structure of GPSG models, *Linguistics and Philosophy*, 13:523–590.

Ristad, E. (1990b), "Computational Structure of Human Language," Ph.D dissertation, MIT Department of Electrical Engineering and Computer Science. (Revised version available as MIT Artificial Intellegence Laboratory technical report 1260; to appear with additional material as MIT Press monograph)

Ross, J. R. (1967), "Constraints on Variables in Syntax," Ph.D dissertation, MIT Department of Linguistics and Philosophy. (Published in 1986 as "Infinite Syntax!," Ablex, Norwood, NJ.)

Rounds, W. (1975), A grammatical characterization of the exponential time languages, *in* "Proceedings, 16th Annual Symposium on Switching Theory and Automata," pp.135–143, IEEE Computer Society, New York.

Sells, P. (1987), Aspects of logophoricity, *Linguistic Inquiry* 18:445–479.

Wiesemann, U., ed. (1986), "Pronominal Systems," Gunter Narr Verlag, Tübingen.

Winston, Patrick. (1984), "Artificial Intelligence," Reading, MA: Addison-Wesley.