No Vacuous Quantification Constraints in Syntax*

Christopher Potts

UC Santa Cruz

1 [$_S$ No vacuous quantification constraints in syntax!]

Much recent work appeals to a ban on vacuous quantification (NO VACUOUS QUANTIFICATION, henceforth NVQ) that operates not merely as a criterion of non-redundancy in an informal semantic sense but is actually assumed to constrain *syntactic* well-formedness. Enforcing NVQ is excessively expensive in computational terms, provably beyond the power of a context-free grammar (CFG) and probably requiring something more powerful than an indexed grammar for its statement (see section 3). There are syntactic phenomena that cannot be modeled in context-free terms, but, as Gazdar and Pullum (1985: §2.2.5) write, "no phenomena are known which would lead one to believe that the NLs fell outside their [the indexed grammars'] purview." Therefore, we should subject arguments for the syntactic use of NVQ to unusually close scrutiny. I undertake that task here, concluding that *NVQ has not been shown to be a necessary constraint on syntactic structures*.

The strategy is simple: reanalyze past appeals to NVQ using only a basic version of GPSG (as in Gazdar et al. 1985), which is known to define only context-free tree sets (Rogers 1996, 1997). In general, let P be a phenomenon that has been analyzed in terms of NVQ. Successful reanalysis of P using only statements in GPSG is a demonstration that P does not demand the formal complexity of NVQ, or any logic that could enforce it. It is in principle possible for reanalysis to fail for some choice of P; the Swiss-German crossing dependencies described by Shieber 1985 are provably beyond the power of context-free grammars, and so would not yield to this reanalysis. But, as we will see, no phenomena with this status have been found to motivate NVQ, and it seems unlikely that such are forthcoming.

The choice of GPSG is non-essential. Rogers (1998) provides an interpretation of GB theory in strictly context-free terms. And categorial grammars (CGs) with only directional application are context-free (Bar-Hillel et al. 1960). Hence, reanalysis within those theories would do equally well. But couching the discussion in GPSG terms is appropriate. Gazdar et al. 1985 is a paradigm case of the benefits of adopting a restrictive formalism and pushing it as far as possible. This generally leads to innovative analyses, and reveals areas in which additional power is required: a useful, abstract insight into language.

But the primary motivation for this project centers around the issue of decidability of the satisfaction question. All of GPSG can be defined using a decidable logic. But it is possible that the extension of

*I thank Geoff Pullum first and foremost for his invaluable guidance in writing this paper. Thanks also to Sandy Chung, Jim McCloskey, and Line Mikkelsen for comments and suggestions on earlier versions of this paper.

this logic required to state NVQ is undecidable. That is, the question "Is sentence X grammatical according to grammar G" might have no general answer if G contains NVQ. At issue is whether or not the grammar is a theory in the mathematical sense; in essence, the right to use predicates like "follows from" and "predicts" is in jeopardy. Adopting a restrictive formalism avoids these worries entirely, by assuring a positive answer to the decidability question.

More generally, complexity results are of interest in their own right. They provide content to adjectives like "constrained", a means for inter-framework comparison. Additionally, they are often essential for determining whether a proposal actually reduces the expressive power of the grammar.[1]

2 [$_{\text{NP}}$ No Vacuous Quantification (NVQ)]

Though often called upon, NVQ is rarely formulated. Kratzer (1995) is a notable exception; she offers (1).

(1) PROHIBITION AGAINST VACUOUS QUANTIFICATION (Kratzer 1995 (p. 131))
For every quantifier Q, there must be a variable x such that Q binds an occurrence of x in both its restrictive clause and its nuclear scope.

However, this is not broad enough to be the intended statement in most works. The null operator, *OP*, lacks a restriction, yet Kennedy (1997) calls upon NVQ to regulate this extractee; see section 6 below. Similarly, we want NVQ to block, e.g., (2b), which lacks a quantified element.

(2) a. the soup OP_1 Martha prepared t_1
b. *the soup OP_1 Martha prepared dinner

I propose, then, that we formulate NVQ as a constraint on abstraction in the meaning language, essentially adopting the λI calculus of Church 1941.

(3) NO VACUOUS QUANTIFICATION (ABSTRACTION) (see Bittner 1999 (p. 75))
If the semantic translation of a syntactic expression (subtree) α is α', then $\lambda x[\alpha']$ is well-formed if and only if x is a free variable in α'.

NVQ as stated in (3) suffices to block (2b) and also cases involving null operators, since both demand vacuous abstraction for the sake of function application; (4) illustrates.

(4) *Ford's new book, Chris dislikes the recent work.

a. $[\textbf{dislike}(the\text{-}work)(chris)](f's\text{-}recent\text{-}work)$ = failure of function application
b. *$\lambda x[\textbf{dislike}(the\text{-}work)(chris)](f's\text{-}recent\text{-}work)$ = abstraction blocked by (3)

[1] For instance, Rabin (1969) shows that any n-branching tree (every node has n daughters) can be embedded in a binary branching tree with dominance and precedence relations preserved (see Rogers 2000 (§6)). Not all constituency relations are preserved, but these can be restated in terms of precedence. Not all c-command relations are preserved, but all asymmetric c-command relations are preserved, and these are probably the important ones. So it is an issue whether the addition of a "binary branching only" axiom to the tree language reduces expressive power in linguistically relevant ways.

3 The Complexity of NVQ

3.1 A Proof and a Conjecture

Marsh and Partee (1984) prove that NVQ cannot be given a context-free grammar. Since this is an established result, I review it only briefly. We treat variables as strings, so that a variable x_i translates as a string of x's of length i, and say that the leftmost occurrence of a string of length i binds all occurrences of such strings to its right (with strings separated by brackets, say). The proof takes advantage of the fact that context-free languages are closed under image and erasure homomorphisms as well as intersection with regular languages. These closure properties permit reduction of the NVQ language, L_{NVQ}, to the language $\{a^i b^j c^i d^j \mid i \neq j\}$.

The pumping lemma works as follows for a language L. We set a string length n. Then we select a string $K = uvwxy$ of L that is at least as long as n and for which vx is non-empty. The pumping lemma says that L is not a context-free language if it is impossible to find a factorization of K into a string uv^iwx^iy where (i) the length of vwx is less than or equal to n; and (ii) K is in L for all positive, non-null values of i.

Assume the length of the relevant string K is set at n. Then we pick the string $a^n b^{n+1} c^n d^{n+1}$, which is in the above simplification of L_{NVQ}. It is clear that our only hope is to pump "binding pairs": either a's and c's, or b's and d's. Without loss of generality, suppose we pump a's and c's. Then the minimal length of vwx is $n + 2$, violating condition (i) above. The other possible choices for v and x also fail, either because they violate the conditions of the lemma, or because pumping produces strings that are not in the language. The pumping lemma claim thus fails for L_{NVQ}. Briefly, the fact that L_{NVQ} enforces two crossing dependencies entails its non-context-freeness.

Marsh and Partee (1984) conjecture that L_{NVQ} is not an indexed language. They offer compelling arguments, but no proof. There is a pumping lemma for indexed grammars (Hayashi 1973) and a shrinking lemma for them (Gilman 1996), but they are of little use when applied to complex languages like L_{NVQ}. To my knowledge, the exact complexity of L_{NVQ} remains unknown.

However, though I am unable to prove Marsh and Partee's conjecture, I can offer novel support for it. The argument runs as follows:

1. Rogers (1998: §9.3) shows that introducing into the decidable weak monadic second-order logic $L^2_{K,P}$ a predicate *CI* that holds between all and only coindexed nodes results in a theory for which the emptiness question ("does L generate any strings?") is undecidable.

2. It might be possible to restrict *CI* in such a way that it does not yield undecidability for the emptiness question, but, crucially, if there is no bound on the number of relations *CI* can identify, then *CI* yields a logic for which emptiness is undecidable.

3. Emptiness is decidable for the indexed languages.

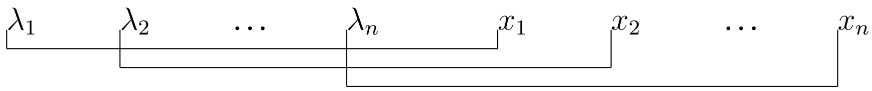
4. By 1-3, $L^2_{K,P}$+*CI* can define languages that are outside the class of indexed languages.

5. Therefore, *if* L_{NVQ} requires the predicate *CI* for its statement *and* L_{NVQ} is indexed, then L_{NVQ} can be stated using a bounded version of *CI*.

But it is easy to see that no such bound can exist. We proceed by *reductio*.

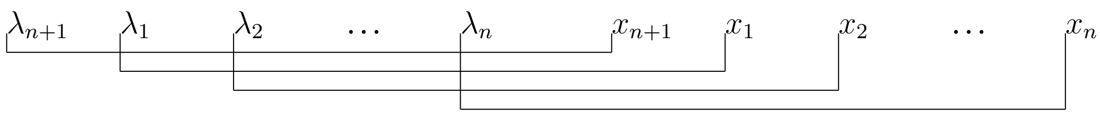
6. Set the bound on *CI* at n. Let S be a sentence containing n operator-variable pairs, and moreover let them all be crossing-dependencies, as in (5), so that we cannot reuse any indices.

(5)



Now we form a sentence S' by placing S in the scope of a lambda λ_{n+1} and insert a variable x_{n+1} with scope over all the other variables in S; this creates another crossing dependency, so the index $n+1$ cannot be among those in S:

(6)



S' is a sentence if S is. Hence we cannot place a numerical bound on *CI* if it is to capture L_{NVQ}. This entails undecidability of emptiness for L_{NVQ}, which entails non-indexed status for L_{NVQ}.

This is, I stress, not a proof that L_{NVQ} is not an indexed language. I have not established that *CI* is necessary to define the language. But can you think of another way?

3.2 Remarks on the Complexity of NVQ

Unfortunately, the pumping lemmas for both tree-adjoining grammars (Vijayshanker 1988) and head-grammars (Roach 1987) (two equivalent formalisms) fail for L_{NVQ}. In essence, the lemmas fail for this reason: they require that we pump a string in the language (one without vacuous quantifiers). Since tree-adjoining and head grammars can enforce crossing dependencies, we can always find a pumpable factorization—again thinking of variable binding in terms of matching strings, we just pick the matching strings of the longest length and pump them.

But crossing dependencies are at the heart of L_{NVQ}'s complexity. As suggested by the above conjecture, NVQ requires us to keep track of an unlimited number of such dependencies. While overlapping chains (of more than, say, two) are horribly ungrammatical for overt extraction in English, there is no limit to them in the area of variable binding; compare (7a) and (7b).

(7) a. Every director$_1$ wondered why every contract$_2$ said that every actress$_3$ could rightly claim that he$_1$ had to talk about it$_2$ with her$_3$ before they went into production.

b. *Which director$_1$ wondered which contract$_2$ they asked which actress$_3$ Francis said t$_1$ had to talk about t$_2$ with t$_3$?

This contrast is extremely important. Marsh and Partee (1984: p. 188) observe that NVQ can be enforced over certain limited domains. Context-free equivalent syntactic theories like GPSG, GB, and CG enforce a one-to-one correspondence between extractees and gaps by severely restricting the positions in which such items occur. As a result, NVQ holds, though it is nowhere stated in the grammar, nor can it be. I return to this issue in section 8.

4 Kratzer 1995 on Adverbs

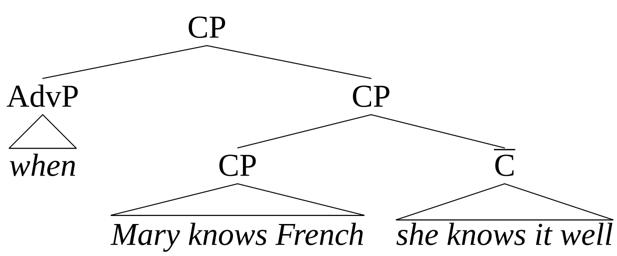
The appeals to NVQ by Kratzer (1995) prove useful in clarifying my limited claim about the status of NVQ. Kratzer explains the contrast in (8) (her (15a,b)) in terms of NVQ, which she defines as in (1) above.

(8) a. *When Mary knows French, she knows it well.
b. When a Moroccan knows French, she is knows it well.

Kratzer treats *when* as semantically equivalent to *always*. She assumes that it unselectively binds all free variables in its scope. In the case of (8a), there are no variables to bind, as indicated in the LF (9a).

(9) a. $\textbf{always}_x[(\textbf{know}(french)(mary) \rightarrow (\textbf{know-well}(french)(mary)]$
(Violates Kratzer's NVQ (1): $\textbf{always}_x$ binds no variables)
b. $\textbf{always}_x[\textbf{moroccan}(x) \wedge (\textbf{know}(french)(x) \rightarrow (\textbf{know-well}(french)(x)]$

There is an undeniable contrast between (9a) and (9b). But is (9a) *syntactically* ill-formed? That is, is (10) allowed by the grammar or not?

(10)


NVQ seem to be the only sensible way one could block this structure, which is isomorphic to that of (8b) up to terminal nodes.

But, importantly, (11) indicates that (10) *is* a model of the grammar.

(11) The medication Mary is taking causes her linguistic knowledge to flicker on and off like faulty neon sign. Some days she knows both of her languages (she's an English/French bilingual), other days she doesn't even know her own name. But one thing you can count on: when Mary knows French, she knows it well.

The example merely requires contextualization. Thus, it would be a mistake to say that vacuous quantification in the denotation of the lower CP node in (10) yields an ill-formed structure, as this would wrongly block (11).[2]

Fox (2000: 168) offers similar cases, one of which I repeat in (12a) along with his minimally contrasting example.

[2] It is common these days to find variables and lambdas adjoined to the trees. These translate directly into semantic objects. On such a view, the structures of (12a,b) might not be the same, since (12b) would involve additional nodes. Two things about this are worth emphasizing: first, these nodes are dubious for independent reasons, since they are not syntactically motivated, nor do they in any sense reduce the complexity of the semantics; rather, their only motivation seems to be pedagogical: they make trees easier to decipher. And second, this view does not derive the contrast in (12a,b), since nothing would block such variables from appearing in the structures of both. I thank Norvin Richards and Yael Sharvit for challenging comments on this point.

(12) a. * [Which languages spoken in a country a linguist_1 comes from] does he_1 usually know t?

b. [Which languages spoken in a country he_1 comes from] does a linguist_1 usually know t?

Fox says of (12a), "...the principle that bans vacuous quantification forces scope reconstruction. Scope reconstruction in turns yields a Condition C effect, and the sentence is ruled out" (p. 168).

Importantly, though, the Condition C violation above does not involve coreference but rather a failure of semantic binding. Since the quantifier, *a linguist*, cannot c-command he_1, even under reconstruction, we expect binding to fail; coindexation is not ungrammatical so much as meaningless (interpretation of *a linguist* is not assignment-dependent). This means that he_1 is a free pronoun. The awkwardness of (12a) then has the same explanation as the awkwardness of Kratzer's examples: the individual denoted by $g(1)$ (g an assignment) is asserted to *usually* know some fact, which is normally odd, but is natural in certain contexts. The anomaly of both this example and Kratzer's is probably the same one that arises from universal quantification over a singleton set (#*Every member of the Potts family that spoke at NELS-32 is tall*); see Partee 1987 (p. 127). This is not a fact about syntax.

5 Chomsky 1982, 1995

Chomsky appeals to NVQ for data such as (13); similar examples appear in Heim and Kratzer 1998 (p. 127).

(13) a. *the man who John saw Bill

b. *the man who_1 John saw him_2.

c. *the man on whom John depends on Bill.

Here again one must ask whether these examples are syntactically ill-formed, as Chomsky assumes, or simply semantically anomalous, the position of Akmajian and Kitagawa (1976) and McCloskey (1979: §4). Although (13a-c) are quite deviant, similar examples are attested; see (14).

(14) a. "There is great enthusiasm for this photo around Asymmetrical Productions, and they're looking into the legalities of using it in Lost Highway ads and posters, which if I was the guy in the photo I'd want a truly astronomical permission fee."
—David Foster Wallace. David Lynch Keeps his Head. *A Supposedly Fun Thing I'll Never Do Again* (p. 149, nt. 4).

b. "You strike me as the kind of guy that, one big screwup, we're living over a candy store on Flatbush Avenue."
Saturday Night Live (thanks to Andrew Dowd for this example (p.c., 7/1/01)).

c. " 'One of those things where you let yourself fall backwards and everyone catches you?' "
—Walter Kirn. *Up in the Air* (p. 183).

d. "Conan O'Brian: There are funny things like "If They Mated," which, we're not kidding ourselves: We know that it's just funny pictures."
—Interview in *The Onion*, May 23, 2001 (interviewer Keith Phipps).

If these examples involve *Wh*-operators (existentials) that do not bind traces, then NVQ is not even a descriptively accurate principle, and is thus to be abandoned on empirical grounds alone. But this is probably too hasty; (13) and (14) obviously contrast. So let's suppose the examples should be treated as syntactically ill-formed. I pursue two analyses that begin with this supposition, both of which deny that NVQ is a principle: (i) Chomsky's attempt to derive NVQ from more general considerations; and (ii) a simple GPSG account that might also entail NVQ.

5.1 Full Interpretation (FI) Implies NVQ

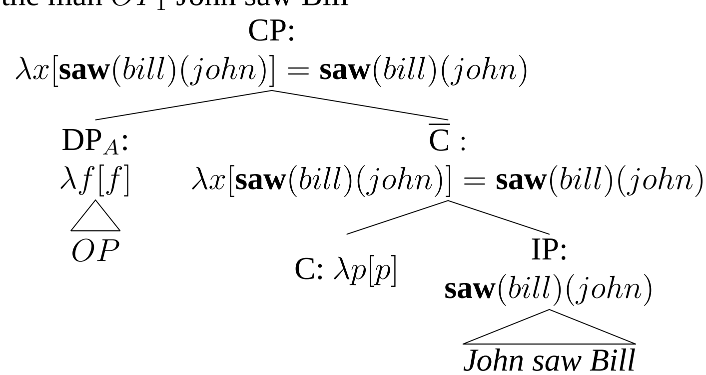
Chomsky's (1995) proposal is that NVQ is not a principle of grammar, but rather a consequence of FULL INTERPRETATION, which is not given explicit definition, but which Chomsky comments on as follows (see also Chomsky 1982: p. 10-11):[3]

(15) a. "Thus, the notion of "Full Interpretation" (FI) requires that representations be minimal in a certain sense" (p. 130).

b. "…just as there can be no superfluous steps in derivations, so there can be no superfluous symbols in representations. This is the intuitive content of the notion of Full Interpretation (FI).…" (p. 151)

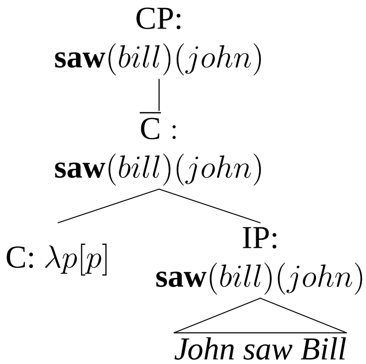
An interpretation of this proposal that seems to succeed in attaining NVQ as a theorem runs as follows. We first treat both (16) and (17) as possible structures.

(16) *the man OP_1 John saw Bill



[3]Chomsky (1982: p. 11) also appeals to NVQ to explain why **all some men* is ill-formed. But this follows from the lexical denotations of the morphemes involves. *some men* denotes a generalized quantifier, of type $\langle\langle e,t\rangle,t\rangle$; *all* probably denotes in more than one domain, but it is never a function in $\langle\langle\langle e,t\rangle,t\rangle,\tau\rangle$. So this is simply a type-mismatch. If *some men* denotes a set, then *all* binds a variable in this set, and so the example satisfies NVQ.

(17) John saw Bill



Then we apply the premise in (15) as follows: the trees in (16) and (17) have identical model-theoretic denotations, as indicated by the translations on their root nodes. But the set of nodes (constituents) in (16) is equivalent to the union of the set of nodes in (17) with the nodes of the subtree DP_A. By (15a), FI requires representations to be "minimal". Hence, (17) blocks (16), since it is the smaller of the two.[4]

Though we obtain the desired result, this calculation is worrisome in its complexity. In particular, it sets well-formedness relative to sets of independent structures and is thus intrinsically transderivational. Linguistic theory as it stands now cannot even accommodate such constraints, since models are generally taken to be individual trees (with nodes as individuals), not sets of trees. Moreover, it is unclear that we should make the necessary widescale adjustments to the theory that are required to state these principles, since even the simplest logic that results will have conceptual and computational drawbacks. For further discussion of these issues, see Johnson and Lappin 1999 and Potts 2001.

But the most serious objection to the FI account is that it is descriptively inaccurate. For instance, it fails if the putative operator carries an entailment. Thus, although (18) is structurally parallel to (13a), the FI account allows it, since the meaning of *who* existentially quantifies over people. That is (18b) entails (18a), but not vice versa.

(18) *the man [$_{\text{CP}}$ who$_1$ [$_{\text{IP}}$ that rhino is dangerous]]

a. [$_{\text{IP}}$ that rhino is dangerous] = $\textbf{dangerous}(that\text{-}rhino)$

b. [$_{\text{CP}}$ who$_1$ [$_{\text{IP}}$ that rhino is dangerous]] =
$\exists x : \textbf{person}(x) \wedge \textbf{dangerous}(that\text{-}rhino)$

Most seriously of all, FI presupposes that every licensed syntactic element makes a semantic contribution. But natural language syntax is replete with semantically empty elements; see (19), where FI certainly puts the star on the wrong example.

(19) a. What did Martha munch upon?
$\lambda p[\exists x : \textbf{munch-upon}(x)(martha)]$

b. *What Martha munch upon?
$\lambda p[\exists x : \textbf{munch-upon}(x)(martha)]$

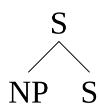
[4]Johnson and Lappin (1999: §2.6) criticize this mode of reasoning in their study of the HAVE AN EFFECT ON OUTPUT CONDITION (HEOC) of Chomsky (1995: p. 294). That principle also favors (17) over (16), indicating redundancy. The criticisms leveled against FI in this section apply equally to the HEOC, and FI is subject to the same conceptual and computational problems observed by Johnson and Lappin for the HEOC.

Auxiliary-*did* contributes nothing to the model-theoretic denotation of the question; an arbitrary restriction suppresses it in embedded contexts. This is one of many cases in which the syntax countenances meaningless elements. Others include the optional complementizer *that*, expletive objects (*I take* *(*it*) *that you're unhappy*), and a host of prepositions (e.g., *I am certain that we'll succeed* versus *That we'll succeed, I am certain* *(*of*)). So appeals to FI to govern syntax are bound to require lots of *ad hoc* explanations for these elements' presence in the grammar.

5.2 A Simple GPSG Account

The complexity of the FI account is not warranted. Within GPSG, we can obtain the desired result by appeal to two quite central tenets of the framework. The first is that a constituent containing a gap is marked with a slash feature, so that, e.g., 'S|NP' labels a tree rooted at S with an NP gap. We then simply stipulate that no rule licenses a tree of the form in (20).

(20) *



Unfortunately, space precludes a complete discussion, but it is worth noting that this restriction can easily be stated in the logic $L^2_{K,P}$ of Rogers (1996, 1997), which is expressively equivalent to a context-free formalism; see (21), in which '$\triangleleft$' is immediate domination.

(21) $\forall x[\mathrm{S}(x) \rightarrow \neg\exists y[\mathrm{S}(y) \wedge x \triangleleft y]]$
—if a node is labeled S, then it does not have a daughter labeled S

Most other frameworks can provide accounts at this level of simplicity. For instance, in CG the result follows from the fact that strings like those in (13) simply lack proofs—their elements cannot be combined to yield something of category S denoting in $\langle t \rangle$. Thus, it is clear that these facts do not demand the radical complications inherent in either NVQ or Chomsky's FI.

For an indication of how a GPSG-style framework might have NVQ as a theorem, see section 8.

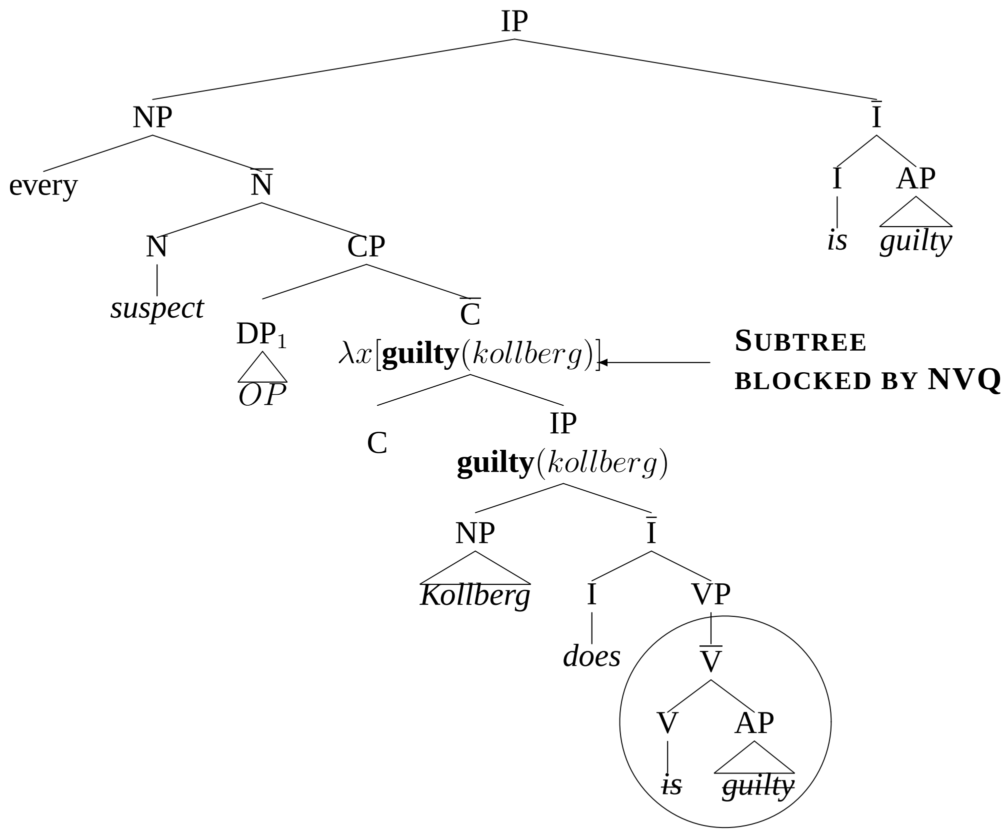
6 Kennedy 1997 on ACD

The appeal to NVQ in Kennedy 1997 offers a slightly different challenge. Kennedy says that antecedent contained deletion (ACD) in (22) is blocked "because it would involve copying the lower VP, generating an LF representation that would be ill formed because the relative operator would not bind a variable" (p. 667, nt. 8).[5]

[5] Wide scope resolution is blocked because it requires the object quantifier to raise out of the embedded tensed clause, which is disallowed by the finite clause-boundedness of non-existential quantifiers.

(22) a. *Beck believes that every suspect Kollberg does is guilty.

b.



If VP-ellipsis involves copying of syntactic material, then (ACD) is blocked for the same reason as (13). But if VP-ellipsis is semantic, it is initially unclear why the semantically $\langle e, t\rangle$-type expression (*is*) *guilty*($\lambda x[\textbf{guilty}(x)]$) cannot fill the ellipsis site. NVQ supplies one answer, but in fact an even simpler principle suffices:

(23) An ACD gap is licit only if its meaning can be supplied by a *two-place relation* somewhere in the derivation (at some node, say) (see Jacobson 1992a,b).

Since **guilty** is a one-place relation, it is not a candidate for ACD resolution. Even (23) probably need not appear in the grammar. Its effects might be obtainable by a strategic assignment of types to the elements that can appear adjacent to ellipsis sites.

7 Fox (2000) on Binding in Coordinate Structures

The most intricate appeals to NVQ in the area of syntax are those of Fox (2000: §2.3), who relies on this principle in his account of the Coordinate Structure Constraint (CSC). Fox's proposal is that CSC violations like (24) follow from a multi-dimensional analysis of coordinate structures plus NVQ.

(24) *Who do you think Mary likes and Bill hates Sue.

a. Who_1 do you think Mary likes t_1?

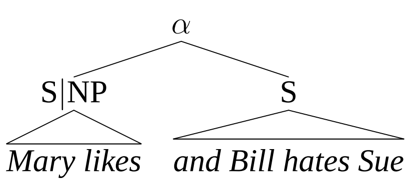
b. Who do you think Bill hates Sue?
(*who* ($\rightsquigarrow \lambda f[\exists x : \textbf{person}(x) \wedge f(x)]$) binds no variable, violating NVQ)

The sentence in (24) is held to have the component LFs in (24a,b). (24a) is well-formed. But if function application is to proceed in (24b), we require vacuous abstraction over the saturated expression [**think**(**bill-hates-sue**)(*you*)],[6] so that it forms a set-denoting argument for **who**.

One might challenge this account of coordination on the grounds that it involves many ancillary statements that are unnecessary in other theories. For example, one must prevent *Who sang and danced?*, decomposed into *Who sang?* and *Who danced?*, from being answered with, e.g., *Eddie sang and Ali danced*. The account is also of unknown complexity. It requires a tree isomorphism, a mapping that is not context-free definable (Rogers 1998: §5.3.5, §9.1).[7] Given the considerations reviewed in section 1, it would be surprising if basic coordinations required this level of computation power.

What's more, it is easy to provide a context-free account of (24) and its ilk. Within GPSG, we can again block these cases by restricting the class of local trees. Simplifying somewhat, the GPSG head feature condition (HFC) says, *inter alia*, that a slash feature S that is not specified in any lexical entry can be instantiated on a head daughter D if and only if S is instantiated on the mother of D. GPSG also assumes that the daughters of a coordinate node are all heads. It follows from these two premises that the tree in (25) is illicit. If α is of category S|NP, then its right daughter violates the HFC, because it lacks the slash feature of its mother. If α is of category S, then its left daughter violates the HFC, because it has a slash feature that its mother lacks. And if α's category is neither S nor S|NP, then both daughters violate the HFC.

(25)



The HPSG account (stated in terms of intersection) is identical in relevant respects. Similarly, the CG assumption that *and* denotes a family of categories and is semantically intersective derives this restriction.[8] So (24) cannot motivate this complex account. But Fox's argument is actually much more subtle.

[6] More precisely, the expression *you think Bill hates Sue* denotes a proposition, and is thus a function from worlds to truth values (set of worlds). But the only variables it contains are world variables, and the abstraction necessary for application is blocked by NVQ.

[7] This might not be too great a worry. Lindell (1992) shows that the graph isomorphism problem for trees of the sort used in linguistics is computable in logspace, and is thus a fairly reasonable computation to attribute to language users. I thank Phokion Kolaitis for providing this reference.

[8] It is probably more accurate to say that coordination requires simply semantic type identity, which in most cases entails syntactic identity but allows for cases like (i) - (ii).

(i) "Virtually all those with whom I talked who knew him well in those years saw him as studious and a member of the lowest ranking high school clique..."
—Alston Chase. Harvard and the Making of the Unabomber. *The Atlantic Monthly*, June 2000 (p. 49).

(ii) "Bill recalled a pretty girl with rosy cheeks and curly blond hair, quiet but walked down the hall like she was hoping somebody would kiss her." —Alan Lightman. *The Diagnosis* (p. 306).

This need not affect the argument in the text. In CG, for instance, *Mary likes* extensionally denotes the set of objects Mary likes, whereas *Bill hates Sue* denotes a truth value.

7.1 NVQ and Scope in Coordinations

Fox's primary support for NVQ as essential to coordination is based initially on (25).

(26) A (#different) student [[likes every professor] and [hates the dean]]

a. [every professor]$_1$ A (#different) student likes t_1

b. [every professor]$_1$ A (#different) student hates the dean

The component (25a) is well-formed. But component (25b) requires vacuous abstraction. Thus, we have an account of the infelicity of *different*, which requires the students to vary with the professors—i.e., a wide scope universal. But this violates NVQ in one conjunct (in component (25b)).

Fox proceeds to offer (26) as decisive evidence that NVQ is the main principle operative in this area. He claims that wide-scope for the universal is exceptionally licensed here because the pronoun in the second conjunct functions as a bindable variable.

(27) A (different) student [[likes every professor] and [wants him$_1$ to be on the committee]]

a. [every professor]$_1$ a (different) student likes t_1

b. [every professor]$_1$ a (different) student wants him$_1$ to be on the committee

In (26b), in contrast to (25b), abstraction need not be vacuous, since *him* can translate as a variable bound by the wide-scope universal. Is *him* a true bound variable? If it were, then the above would be a compelling argument for NVQ as a principle regulating LF structures. But closer inspection reveals that these cases do not involve variable binding.

7.2 Objections to the Variable Binding Analysis

7.2.1 Linear Order Restrictions

A first reason to be skeptical of the claim that (26) involves binding into both conjuncts is that the linear order of the conjuncts is crucial for grammaticality:

(28) a. *A student wants him$_1$ to be on his committee and likes every professor$_1$.

b. A student wants every professor$_1$ to be on his committee and likes him$_1$.

This is a surprise under Fox's analysis, which evaluates the component structures separately. That is, (28) is the analysis of both examples in (27).

(29) a. [every professor]$_1$ A student likes t_1

b. [every professor]$_1$ A student wants him$_1$ to be on his committee

An appeal to linear ordering constraints on variable-binding is not sufficient, given, e.g., (29).

(30) That he$_1$ is guilty, no prisoner$_1$ will admit.

Examples like (29) are often analyzed as being licensed by syntactic reconstruction of the topicalized clause into a position within the c-command domain of *no prisoner*. One could then maintain that a linear ordering constraint holds on variable-binding at LF even in the face of this example. But it is doubtful that the notion of linear ordering even makes sense for LF representations. As Chomsky (1995: §4.8) says, "there is no clear evidence that order plays a role at LF"(p. 334).

7.2.2 Weak Crossover Considerations

In (26), the quantifier does not c-command the pronoun at surface structure. Such configurations generally yield weak-crossover violations; see (30) for typical examples.

(31) a. *His$_1$ student disparaged every professor$_1$ behind his back.
b. *No professor$_1$ showed up late, and his$_1$ students were upset by this.

Although there are exceptions, for instance (31)—

(32) a. No professor$_1$'s student disparaged him$_1$.
b. Everybody in some city$_1$ hates its climate. (Büring 2001: (9a))

—these too fail to involve straightforward variable-binding. Büring (2001) offers a promising analysis involving PAYCHECK PRONOUNS. That is, the apparently bound pronoun is actually functional—essentially giving rise to an interpretation equivalent to a definite description. As I show in the next section, Fox's cases are accurately described only terms such as these.

7.2.3 Downward-Entailing Operators

The interpretation of examples identical in form to Fox's (26) reveal that we do not have bound variables in the second conjunct. This line of reasoning derives from Fox's own insights into the way variable binding works. He observes that certain cases of apparent binding actually involve what he calls ILLUSORY VARIABLE BINDING or TELESCOPING. Downward entailing operators like *only one* yield interpretive differences when binding is illusory:[9]

(33) John loves three of the women he knows.
#However, he loves only one of them and expects her to love him back.

= *Contradictory*. John loves just one woman, call her Ali. He expects Ali to love him back.

(34) John loves three of the women he knows.
However, there is only one of them that he loves and expects to love him back.

= *Unexceptional*, because *only one of them* scopes over both conjuncts.

Only (33) is interpreted as a case of variable binding; this is in accord with the conclusion in section 7.2.2 that surface c-command is required for such binding. In (32), *only one of them* does not c-command *her*, and hence *her* is interpreted as, roughly, *the woman John loves*. To be concrete, we can assume that the pronoun *her*$_1$ is a referential pronoun—i.e., the assignment maps the index 1 to the

[9]Is *only* downward entailing? There are two schools of thought:

(i) "only is not downward entailing": Gamut 1991 (p. 239); Szabolsci and Zwarts 1997 (p. 225), Atlas 1993, Horn 2001

(ii) "only is downward entailing": Hoeksema 1986 (p. 38); Horn 1997 contextually salient entity that John loves. Since this entity is made salient by the material in the initial conjunct, we expect linear order restrictions between the two conjuncts.

If we heed Fox's warning and use downward entailing operators to test for binding in coordinate cases like (26), it becomes apparent that they involve illusory variable binding. A case of the required type is (34).

(35) a. $\neq$ There is exactly one professor such that every student both likes him and wants him to be on the committee. (There might be lots of professors that every student simply likes, and lots that every student wants on committee. But just one possesses both these properties.)

b. = There is at most one professor that every student likes. This lucky professor is such that every student wants him on the committee.

(34a) is the true variable binding interpretation; it is equivalent to *Every student likes only one professor that he wants to be on his committee*. But this reading is not available for (34), which is interpreted as in (34b), in which *him* in the second conjunct is interpreted as *the professor that every student likes*. Since the first conjunct introduces this entity, we derive the linear order restrictions of Section 7.2.1 from the fact that *him* is a definite pronoun.

A precise semantic representation of (34) is given in (35), which maintains Fox's multi-dimensional analysis

(36) a. Every student likes only one professor $\rightsquigarrow$
$\exists y[\textbf{professor}(y) \wedge \forall x[\textbf{student}(x) \wedge \exists z[\textbf{professor}(z) \wedge \textbf{like}(z)(x)]] \rightarrow z = y]$

b. Every student wants him1 to be on the committee $\rightsquigarrow$
$\forall x[\textbf{student}(x) \rightarrow \textbf{want}(\textbf{on-the-committee}(\textit{him}_2))(x)]$
, where the assignment maps 2 to the contextually salient individual that every student likes

Another reason to be skeptical of the binding analysis is that the pronoun in the second conjunct can receive a functional (telescoping) interpretation even when *only one of them* takes narrow scope with respect to *every student*. To be concrete, (34) is felicitous in the situation in (36), in which the students like different professors.

(37)

student	=	$\{$*ali, art, brad*$\}$
professor	=	$\{$*mrs.-tenney, mr.-chapin, mr.-larson*$\}$
like	=	$\{\langle$*ali, mrs.-tenney*$\rangle, \langle$*art, mr.-chapin*$\rangle, \langle$*brad, mr.-larson*$\rangle\}$
$\textbf{him}_2$	=	$\lambda x : \textbf{student}(x)[\iota y : \textbf{professor}(y) \wedge \textbf{like}(y)(x)]$

In sum, *him* can appear to be bound even when *only one of them* remains in the scope of *every professor*, and hence outside its binding domain.

These interpretive facts, when considered alongside the linear order restrictions and the weak crossover data, seem decisive. This is not to say, though, that Fox's cases are uninteresting. On the contrary, they present an important, unsolved puzzle concerning quantifier scope and E-type anaphora; the contrast between (25) and (26) still lacks an explanation. In the next section, I deepen the mystery, and take steps towards its resolution.[10]

[10]The data reviewed in this section are quite complex. I have yet to encounter speakers who allow variable-binding in these cases, but testing is of course difficult. Suppose such speakers exist (or that we can find examples that reveal such a

7.3 Overt Extraction and the CSC

The exceptional cases that Fox investigates are not limited to covert movement. Some cases of overt extraction violate the CSC, as seen in (37).

(38) a. "Let's get to our first guest, who I asked for and was so delighted that he could make it."
(from Prince 1990: (4), citing Orson Welles on *The Tonight Show*)

b. ...who$_1$ I [$_{VP}$ asked for t$_1$ and [$_{VP}$ was so delighted that he$_1$ could make it...

(39) a. *...our first guest, who I asked for and was so delighted that the show started on time...

b. ?...who I asked for and was so delighted we could have such an illustrious line-up...

However, a pronoun in one conjunct is not a necessary condition on extraction from a single coordinate phrase; see (39), which I have annotated.

(40) a. "This week also features an unbelievably large and involved jigsaw puzzle that sits about half-done on an oak table in the corner, which$_1$ all sorts of different old people [$_{VP}$ come in and work on t$_1$ in shifts]."
—David Foster Wallace. A supposedly fun thing I'll never do again. *A Supposedly Fun Thing I'll Never Do Again* (p. 325)

b. "But now, just as I need to leave in order not to be late for 1500h.'s much-anticipated skeetshooting, Scott Peterson starts to relate an anecdote OP_1 that engages my various on-board dreads and fascinations enough for me to [$_{VP}$ stay and try to write down t$_1$]."
—David Foster Wallace. A supposedly fun thing I'll never do again. *A Supposedly Fun Thing I'll Never Do Again* (p. 340)

c. It was **de Gaulle** who$_i$ Frank [criticized t$_1$ and in doing so criticized a Frenchman]. (Levine 2001: p. 156-7, nt. 12)

Can we make sense of these counterexamples to the CSC (in some cases, counterexamples to both the GPSG analysis offered above and Fox's multi-dimensional analysis plus NVQ)? I propose that the source of asymmetries like those above stems from the predication formed by the coordinate phrase. It seems that wide-scope for (or extraction of) an operator in a single conjunct is allowed only if the other conjunct(s) can be interpreted as relevant to the operator's domain. An E-type pronoun in one conjunct fits the bill, but this is not the only way such a predication can be obtained. In (39c), for instance, the

reading). Does this suffice to motivate NVQ? I think not; an alternative is available. For instance, such (presently hypothetical) cases might exceptionally involve the ASYMMETRIC COORDINATION studied by Büring and Hartmann (1998), in which a quantifier embedded in an initial conjunct binds a variable in a second conjunct, as in (i) (their (24c)).

(i) Im Zirkus Krone serviert der Dompteuse jedem Löwen eine Antilope$_1$ und würzt sie$_1$ mit Löwensenf.
in circus Krone serves the trainer every lion an antelope$_1$ and seasons it$_1$ with spicy-mustard
'In circus Krone, a trainer serves every lion an antelope that is seasons with spicy mustard.'

Büring and Hartmann support an analysis in which *und würzt sie1 mit Löwensenf* is a VP-adjunct, hence in the scope of the scrambled quantifier *eine Antilope*. These examples also display the linear order restrictions of section 7.2.1. See Johnson 2000 for evidence from certain coordinate structures (gapping) that this adjunction is licit in English as well, at least under special conditions.

adverbial *in doing so* explicitly links the two conjuncts, and *a Frenchman* is semantically related to the ultimate interpretation of the extractee *who*. This reasoning also extends to the subtle contrast in (38).

This proposal has two significant consequences: (i) it means that the CSC is not a narrowly grammatical restriction; the account Fox offers, and the GPSG account based on the ill-formedness of a class of local subtrees, are too blunt (see Levine 2001 for arguments for this conclusion and additional references); and (ii) it means that the CSC cannot be used to motivate NVQ as a principle of syntactic structure, since, by (i) it is not a factor in structural well-formedness.

8 Closing Remarks

NVQ is of considerable complexity and thus should be adopted only as a last resort; the theory is best served by resistance to it. What's more, at present, motivation for it is lacking; it is easy to reanalyze the variable-binding data that researchers have marshaled in its support using only context-free techniques. Thus, the facts in question do not even motivate a more powerful syntax than that provided by GPSG.

It is possible, though, that when one limits attention to the syntax, NVQ holds (this assumes that we can find alternative explanations for the counterexamples like (14)). This does not, however, mean that it is a statement in the grammar. Rather, it might be a *theorem* of the grammar, and thus have the meta-level status of, e.g., a completeness proof in logic. To see that this is a real possibility, consider the grammar fragment in (40).

(41) a.

$$\begin{array}{ll} S & \rightarrow NP\ S|NP \\ S|NP & \rightarrow NP\ VP|NP \\ VP|NP & \rightarrow V \end{array}$$

b.

$$\begin{array}{ll} S & \rightarrow NP\ VP \\ VP & \rightarrow V\ NP \end{array}$$

This grammar enforces NVQ in the sense that an NP that is sister to a slashed category S|NP must be part of a derivation that contains an object gap. There is no way to move from an initial expansion S → NP S|NP (an expansion with an extracted operator) to a constituent without a gap (an expansion VP → V NP); the rules in (40a) do not interact with those in (40b). It would be a confusion to proceed to state NVQ as a principle of this grammar itself. This is in the spirit of Chomsky's view of NVQ as I interpret it in section 5.1, though that derivation evidently introduces more complexity than NVQ itself. Results like this are of great value, since they exploit the deductive structure of the theory. For steps in this direction regarding the economy condition Shortest Move, see Kracht 2001 (§10).

I close on a speculative note. It is clear that the semantics of natural languages requires a very complex logic. The calculation of presuppositions, for instance, seems based on a sophisticated, implicit knowledge of set theory. But I suggest that the syntax is something very close to context-free; at present we lack evidence that it is more powerful than a tree-adjoining grammar. Why the asymmetry?

Katz (1981) argues that the defining feature of natural languages is their ability to express *every* proposition. This is his Principle of Effability (p. 226). This would seem to require a maximally expressive system. But natural languages are used in real time, hence must be efficiently parsable. This requirement can be met only by very restrictive apparatus—perhaps something very nearly context-free.

References

Akmajian, Adrian and Chisato Kitagawa. 1976. Deep-Structure Binding of Pronouns and Anaphoric Bleeding. *Language* 52 (61-77).

Atlas, Jay David. 1993. The Importance of Being *only*: Testing the Neo-Gricean versus Neo-Entailment Paradigms. *Journal of Semantics* 10 (301-18).

Bar-Hillel, Y., H. Gaifman, and E. Shamir. 1960. On Categorial and Phrase Structure Grammars. *Bulletin Research Council of Israel* 9: 1-16.

Bittner, Maria. 1999. Concealed Causatives. *Natural Language Semantics* 7 (1-78).

Büring, Daniel. 2001. A Situation Semantics for Binding out of DP. *Proceedings of SALT 11*.

Büring, Daniel and Katharina Hartmann. 1998. Asymmetrische Koordination. *Linguistische Berichte* 174: 172-201.

Chomsky, Noam. 1982. *Some Concept and Consequences of the Theory of Government and Binding*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Church, Alonzo. 1941. *The Calculi of Lambda-Conversion*. Number 6 in *Annals of Mathematical Studies*. Princeton, NJ: Princeton University Press.

Fox, Danny. 2000. *Economy and Semantic Interpretation*. Cambridge, MA: MIT Press.

Gamut, L.T.F. 1991. *Logic, Language, and Meaning, Volume 2: intensional logic and logical grammar*. Chicago, IL: University of Chicago Press.

Gazdar, Gerold, Ewan Klein, Geoffrey K. Pullum, and Ivan Sag. 1985. *Generalized Phrase Structure Grammar*. Cambridge, MA: Harvard University Press.

Gazdar, Gerold and Geoffrey K. Pullum. 1985. Computationally Relevant Properties of Natural Languages and their Grammars. *New Generation Computing* 3 (273-306).

Gilman, Robert H. 1996. A Shrinking Lemma for Indexed Languages. *Theoretical Computer Science* 163: 277-281.

Hayashi, T. 1973. On Derivation Trees of Indexed Grammars: An Extension of the $uvwxy$-Theorem. *Publication of RIMS Kyoto University* 9 (61-92).

Heim, Irene and Angelika Kratzer. 1998. *Semantics in Generative Grammar*. Malden, MA: Blackwell.

Hoeksema, Jack. 1986. Monotonicity Phenomena in Natural Language. *Linguistic Analysis* 16 (25-40).

Horn, Laurence R. 1997. Negative Polarity and the Dynamics of Vertical Inference. In Danielle Forget, Paul Hirschbühler, France Martineau, and Martia-Luisa Rivero (eds.), *Negation and Polarity, Current Issues in Linguistic Theory: 155*. Amsterdam, Holland: John Benjamins Publishing Company.

Jacobson, Pauline. 1992. Flexible Categorial Grammars: Questions and Prospects. In Robert Levine (ed.), *Formal Grammar: Theory and Implementation*, 129-167. New York, NY: Oxford University Press.

Jacobson, Pauline. 1992. Antecedent Contained Deletion in a Variable Free Semantics. In Chris Barker and David Dowty (eds.), *Proceedings of SALT II*, 193-213. Columbus, OH: OSU Working Papers in Linguistics, Ohio State University.

Johnson, David E. and Shalom Lappin. 1999. *Local Constraints vs. Economy*. Stanford, CA: CSLI Publications.

Johnson, Kyle. 2000. Few Dogs Eat Whiskers or Cats Alpo. In Kiyomi Kusumoto and Elisabeth Villalta (eds.) *University of Massachusetts Occasional Papers 23: Issues in Semantics*, 59-2. University of Massachusetts, Amherst, MA; GLSA Publications.

Katz, Jerrold. *Language and Other Abstract Objects*. Totowa, NJ: Rowman and Littlefield.

Kennedy, Christopher. 1997. Antecedent Contained Deletion and the Syntax of Quantification. *Linguistic Inquiry* 28 (662-88).

Kracht, Marcus. 2001a. Logic and Syntax—A Personal Perspective. In Michael Zakharyaschev, Krister Segerberg, Maarten de Rijke, and Heinrich Wansing, *Advances in Modal Logic, Volume 2*, 337-366. Stanford, CA: CSLI Publications.

Kratzer, Angelika. 1995. Stage-Level and Individual-Level Predicates as Inherent Generics. In Gregory N. Carlson and Francis Jeffry Pelletier (eds.), *The Generic Book*, 125-75. Chicago, IL: University of Chicago Press.

Levine, Robert D. (2001). The Extraction Riddle: Just What are We Missing? *The Journal of Linguistics* 37 (145-174).

Lindell, S. 1992 A Logspace Algorithm for Tree Canonization. In *Proceedings of the Twenty-Fourth Annual ACM Symposium on the Theory of Computing*, 400-404. New York, NY: ACM.

Marsh, William and Barbara Partee. 1984. How Non-Context Free is Variable Binding? In Mark Cobler, Suannah MacKaye and Michael Wescoat (eds.), *Proceedings of the West Coast Conference on Formal Linguistics, Volume 3*, 179-90. Stanford University: The Stanford Linguistics Association. [Reprinted in Savitch, Walter J. 1987. *The Formal Complexity of Natural Language*, 369-386. Dordrecht: D. Reidel Publishing Company.]

McCloskey, James. 1979. *Transformational Syntax and Model Theoretic Semantics*. Dordrecht: D. Reidel.

Partee, B. 1987. Noun Phrase Interpretation and Type-Shifting Principles. In J. Groenendijk, D. de Jongh, and M. Stokhof (eds.), *Studies in Discourse Representation Theory and the Theory of Generalized Quantifiers*, 115-43. Dordrecht: Foris Publications.

Partee, Barbara, Alice ter Meulen, and Robert E. Wall. 1993. *Mathematical Methods in Linguistics: corrected 1st edition (student edition)*. Dordrecht: Kluwer Academic Publishers.

Potts, Christopher. 2001. Three Kinds of Transderivational Constraint. In *Syntax at Santa Cruz, Volume 3*. University of California, Santa Cruz: Slug Pubs

Prince, Ellen F. 1990. Syntax and Discourse: A Look at Resumptive Pronouns. In S. Reinman, et al. (eds.), *Proceedings of the Sixteenth Annual Meeting of the Berkeley Linguistics Society. Parasession on the Legacy of Grice*, 482-497.

Rabin, Michael O. Decidability of second-order theories and automata on infinite trees. *Transactions of the American Mathematical Society* 141 (1-35).

Roach, Kelly. 1987. Formal Properties of Head Grammars. In Alexis Manaster-Ramer (ed.), *Mathematics of Language*, 293-347. Amsteram, Holland: John Benjamins Publishing Company.

Rogers, James. 1996b. What does a Grammar Formalism Say about a Language? IRCS Technical Report Number 96-10. University of Pennsylvania. [Available at `http//www.cs.earlham.edu/˜jrogers`]

Rogers, James. 1997. “Grammarless” Phrase Structure Grammar. *Linguistics and Philosophy* 20 (721-46).

Rogers, James. 1998. *A Descriptive Approach to Language-Theoretic Complexity*. Stanford, CA: CSLI/FoLLI.

Rogers, James. 1999. The Descriptive Complexity of Generalized Local Sets. In Hans-Peter Kolb and Uwe Mönnich (eds.), *The Mathematics of Syntactic Structure: Trees and Their Logics*, Studies in Generative Grammar 44, 21-40. Mouton de Gruyter, Berlin.

Rogers, James. 2000. Syntactic Structures as Multi-Dimensional Trees. *Language and Computation*. [Available at `http//www.cs.earlham.edu/˜jrogers`]

Ross, John Robert. 1967/1984. *Constraints on Variables in Syntax*. Doctoral dissertation, MIT, Cambridge, MA. [Published 1984 as *Infinite Syntax*! Norwood, NJ: Ablex Publishing.]

Shieber, Stuart M. 1985. Evidence Against the Context-Freeness of Natural Language. *Linguistics and Philosophy* 8 (333-343). [Reprinted in Savitch, Walter J. 1987. *The Formal Complexity of Natural Language*, 320-334. Dordrecht: D. Reidel Publishing Company.]

Szabolsci, Anna and Frans Zwarts. 1997. Weak Islands and an Algebraic Semantics for Scope Taking. In Anna Szabolsci (ed.), *Ways of Scope Taking*, 217-62. The Netherlands: Kluwer Academic Publishers. [Reprinted, with slight alterations, from *Natural Language Semantics* 1 (1993 (235-84)).]

Vijayshanker, K. 1988. *A Study of Tree Adjoining Grammars*. Doctoral Dissertation. University of Pennsylvania: Department of Computer and Information Sciences.

Weir, D.J. and Aravind K. Joshi. 1988. Combinatory Categorial Grammars: Generative Power and Relationship to Linear and Context-Free Rewriting Systems. *Proceedings of the 26th Meeting of the Association for Computational Linguistics*, 278-285.

Department of Linguistics
UC Santa Cruz
1156 High Street
Santa Cruz, CA 95064-1077

potts@ling.ucsc.edu