Expressive power of the syntactic theory implicit in *The Cambridge Grammar of the English Language*[*]

Geoffrey K. Pullum
Linguistics and English Language
University of Edinburgh

James Rogers
Department of Computer Science
Earlham College

1 Introduction and motivation

This paper has three aims. First, to illustrate a particular model-theoretic technique for obtaining results about formalized theories of syntax. Second, to show how large-scale informal grammatical description can be drawn somewhat closer to formal syntactic theory. And third, to demonstrate a way of obtaining an expressive power result by a rather unexpected route, achieving a (conditional) result on generative capacity in a way that entirely avoids reference to generative grammars.

The topic is the expressive power of the syntactic theory implicit in *The Cambridge Grammar of the English Language* (Huddleston et al. 2002), which we refer to henceforth as *CGEL*. *CGEL* is an informal survey of the syntax and morphology of contemporary Standard English on a consistent basis of assumptions and terminology. It focuses on description rather than general linguistic theory, yet says enough to permit inferences to be drawn concerning many aspects of the grammatical framework it assumes. We undertake here the exercise of formalizing some central aspects of that tacitly assumed framework, with a view to determining its expressive power.

The analyses of *CGEL* are not, of course, assumed here to be definitive. They may well be mistaken, or may need elaboration or abandonment in some cases. But our work is at a fairly high level of generality, and should apply to modified variants of the *CGEL* framework just as well as to the one presented in *CGEL* itself.

2 *CGEL* and the structures it tacitly assumes

Table 1 gives a list of the 16 major grammatical categories employed in *CGEL*. None are particularly controversial — though note that **DP** (Determinative Phrase) is the category of phrases like *hardly any*, not that of phrases like *hardly any banks*, which would be an **NP** (Noun Phrase) for *CGEL*, as for the traditional grammars that preceded it, and generative grammars before 1987.

Table 2 shows the twenty grammatical functions *CGEL* employs. Four of the most common of those are illustrated in Figure fig.egstruc. Note that nearly all are special cases of others: every IndObj is an Object, every Object is a Complement, every Attributive is a Modifier, every Modifier is an Adjunct, and all Complements and Adjuncts are Dependents. The fundamental distinction is between Head and Dependent.

In Figure 1 we provide an example of a *CGEL* analysis, represented in a rather cluttered diagrammatic form that we shall use only temporarily. The names next to the lines are names of what *CGEL* calls grammatical functions.

The labels on the lines in sanserif font like Head, Det (for 'Determiner of'), Comp (for 'Complement of'), and so on, correspond to grammatical functions. The labels in boldface at line junctions are categories.

[*]This work was supported by the Radcliffe Institute for Advanced Study at Harvard University. We are grateful to Alexander Koller for noticing an error in an earlier version.

LABEL	EXPLANATION	EXAMPLE
Adj	adjective	*insincere*
AdjP	adjective phrase	*quite obviously insincere*
Adv	adverb	*obviously*
AdvP	adverb phrase	*quite obviously*
Clause	clause	*every word of the statement seemed quite obviously insincere*
Crd	coordinator	*and*
D	determinative	*every*
DP	determinative phrase	*almost every*
N	noun	*claim*
Nom	nominal group	*word of the statement*
NP	full noun phrase	*every word of the statement*
P	preposition	*of*
PP	preposition phrase	*of his and her friends*
Sbr	subordinator	*that*
V	verb	*seemed*
VP	verb phrase	*seemed quite obviously insincere*

Table 1: Grammatical category labels employed in *The Cambridge Grammar*

Notice (for it will be a point of central importance in what follows) that although most of the structure in Figure 1 is treelike, the Det of the **NP** and Head of the **Nom** are fused in the subject noun phrase constituent *some of her friends*.

We now consider how to formalize structures like the one in Figure fig.egstruc. Like just about any imaginable kind of syntactic representation, *CGEL*'s syntactic representations can be formalized as graphs (that is, sets of nodes with a relation defining links between certain pairs of them), decorated with a certain vocabulary of symbols. Mostly those graphs are very much like trees. *CGEL*'s structures are in fact not always constituent-structure trees in the defined sense familiar from mathematical linguistics, and Figure 1 is not.

We set aside one difference that is not illustrated in Figure 1. Some of the representations in *CGEL*'s chapter 15 show parenthetical constituents as loosely connected into trees by a formally unexplicated relation diagrammed as a dotted line indicating the 'Supplement' relation, intended to signal that they are not fully part of the tree-like structure. As Chris Potts (2005, chapter 6) notes, this in effect introduces a third dimension (in the sense of a third kind of possible adjacency) into syntactic structures. Potts questions the need for any such divergent syntax for supplements. His book provides a convincing alternative in semantic terms, and supports it very convincingly. Potts assumes parentheticals are integrated into trees in the same way as any other phrases, differing crucially in two main ways: their intonational phonology and their semantic interpretation. We therefore ignore the issue of the implications of *CGEL*'s dotted-line diagrams, and concentrate on two much clearer and more substantive differences which are illustrated in Figure 1.

The first difference concerns the decorations: in an ordinary tree it is only the points or nodes that are labelled, with category labels like **NP** and **V**. But in *CGEL* there are (tacitly, but see page 25) also **edge labels**, corresponding to grammatical relations like Subject, Object, Head, Complement, and Adjunct.[1]

[1]To forestall a distracting terminological problem, note that the terms 'grammatical relation' and 'grammatical function' are both used in the literature for relations like 'is the subject of'. But not all grammatical relations are functions in the mathematical sense. From a 'bottom-up' perspective, a relation like 'is Object of' is a function, since there can be only one node that a given node is Object of; but the from a 'top-down' perspective it may not be. The matter will depend on whether any VP node can immediately dominate more than one object — and *CGEL* assumes the answer is yes. The relation Coordinate in *CGEL* is a clear case of a relation whose inverse is not a function: every node that is the root node of a coordination has two or more immediate constituents that bear the Coordinate relation to it, so although 'is a coordinate of' (bottom-up) is a function, 'has as one of its coordinates' (top-down) is not. In this paper, for convenience, we have chosen to represent all grammatical relations in 'top-down' mode: we want to be able to use the functional notation $f(x) = y$ in a way that accords with the relational notation $R(x, y)$, and it would be confusing to switch the order of the arguments. Thus under our formalization *CGEL*'s 'grammatical functions' are not always functions on the domain in the mathematical sense. This may be confusing,

Relation	Abbreviation	Remarks
Adjunct	Adjnc	special case of Dependent
Attributive	Attrib	special case of Modifier
Complement	Comp	special case of Dependent
Coordinate	Co	function of elements in a coordination
Dependent	Dep	equivalent to 'non-Head'
Determiner	Det	special case of Dependent
DirObj	O^{d}	special case of Object
Head	H	equivalent to 'non-Dependent'
IndObj	O^{i}	special case of Object
Marker	Mkr	function of Subordinators and Coordinators
Modifier	Mod	special case of Adjunct
Nucleus	Nuc	special case of Head
Object	Obj	special case of (Internal) Complement
PredComp	PC	special case of Complement
Predicate	Pred	special case of Head (head of Clause)
Predicator	Pred	special case of Head (head of VP)
Postnucleus	Postnuc	special case of Dependent
Prenucleus	Prenuc	special case of Dependent
Subject	Subj	also External Complement
Supplement	Supp	special case of Adjunct

Table 2: Grammatical relations (or functions) in *The Cambridge Grammar*

The second difference is that a single constituent may bear two different grammatical relations simultaneously; that is, it may be at the ends of two different edges with distinct edge labels. This situation is referred to in *CGEL* as **function fusion**.

Our methods of formalization in this paper will be model-theoretic rather employing the tools of generative grammar, which originate in proof theory (see Pullum and Scholz 2005 for further discussion of the distinction). We will formalize syntactic structures of natural language expressions as **relational structures** in the model theorist's sense.

A relational structure is just a set with certain relations defined on it. An undecorated graph can be formalized as a particularly simple relational structure that has just one binary relation holding between those pairs of elements that are directly connected by an edge. We will follow most linguists in calling the elements in a graph **nodes**. Unordered trees are a special case of graphs, and linguists call the fundamental edge relation in trees **dominance**.

Linguists most commonly work with **ordered** trees. An ordered tree has an additional binary relation defined on the set of nodes: in addition to dominance there is a relation of **precedence**.

The decorating of the nodes with labels can be represented in terms of relations whose arity is 1. For example, a three-node tree with an A immediately dominating a B followed by a C can be identified with a relational structure (or more pedantically, an equivalence class consisting of all and only the relational structures isomorphic to the tree in the obvious way) in which there are three nodes n_1, n_2, n_3 and a total of five relations:

— a unary relation corresponding to being labelled A and containing just n_1;
— another unary relation corresponding to being labelled B and containing just n_2;
— a third one corresponding to being labelled C and containing just n_3;
— a binary relation of proper domination relating n_1 to n_2 and n_1 to n_3; and
— a binary relation of immediate precedence (being immediately to the left of) that relates n_2 to n_3.

To represent such structures in symbols (which will be necessary later in order to prove things about them), we are taking an ordered tree labelled from the label inventory Σ to be a relational structure $\mathcal{T} = \langle T, \triangleleft_1, <_1, \triangleleft_2, <_2, P_\sigma \rangle_{\sigma \in \Sigma}$, where

but it is a purely terminological matter, so we will ignore it, and talk about particular grammatical relations and grammatical functions interchangeably. Use of the latter term will never be intended to signal the property of being functional — i.e., to entail $(\forall x, y)[(R_i(x, y)) \rightarrow ((\forall z)[R_i(x, z) \rightarrow z = y])]$.

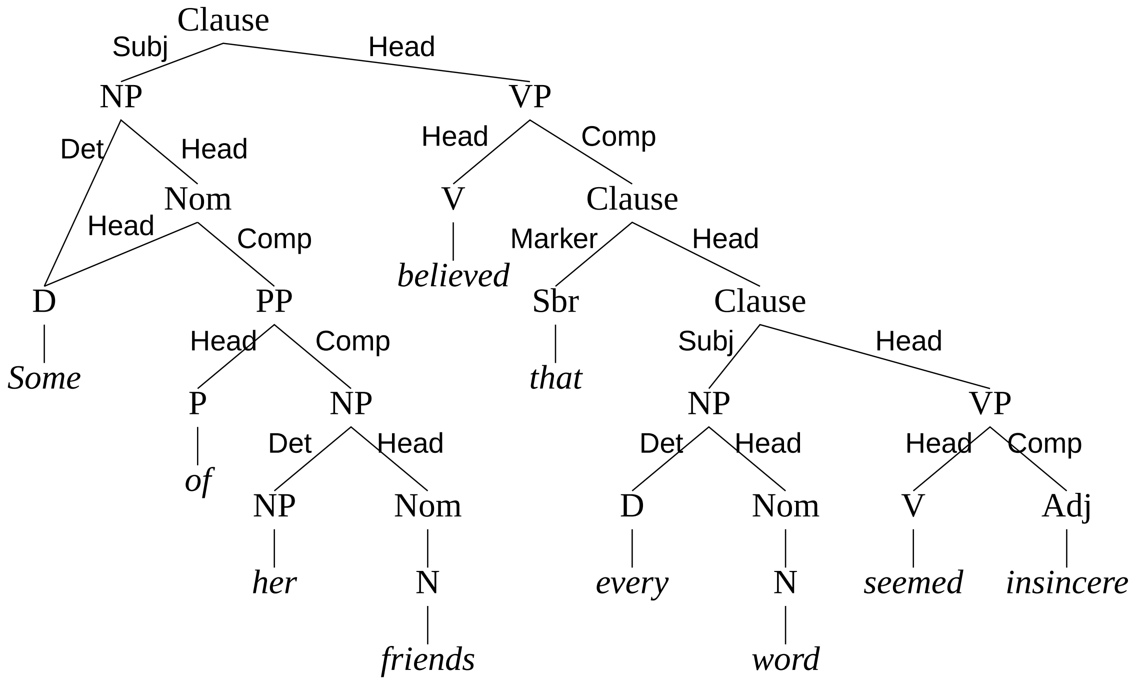

Figure 1: A *CGEL* analysis

— T is the domain (the set of nodes),
— $\lhd_1$ is the immediate precedence relation (the relation 'immediately before'),
— $<_1$ is the irreflexive transitive closure of $\lhd_1$ (hence the relation 'somewhere before'),
— $\lhd_2$ is the immediate dominance relation ('parent of'),
— $<_2$ is its irreflexive transitive closure ('properly dominates'), and
— P_σ for each σ in the set Σ is a unary relation picking out the set of nodes labelled σ.

Every tree has to be such a structure, but of course it also has to satisfy all the usual tree axioms: there must be a single root (a node dominating every node); there must be no tangling of branches (precedence must be inherited down dominance chains); the union of $<_2$ with its inverse must be disjoint from the union of $<_1$ with its inverse; and the union of $<_2$ and $<_1$ with their inverses together with the identity relation must exhaust the domain.

We can also represent edge labels and function fusion in terms of relational structures. All that is involved is adding a set of edge labels. To make our proposal fully precise, we will model a *CGEL* structure as a septuple

$$\mathcal{D} = \langle D, E_f, N_c, \rightsquigarrow, \overset{+}{\rightsquigarrow}, \lhd, \leq \rangle_{f \in F, c \in C}$$

the seven elements of which are as follows:

- D is a finite domain (the set of nodes);
- E_f is a set of subsets of $D \times D$ (one for each $f \in F$) such that $\langle x, y \rangle \in E_f$ iff there is an f-edge from x to y (the set of labeled edge-sets: the Subject edges, the Head edges, and so on);
- N_c is a set of subsets of D (one for each $c \in C$) such that $x \in N_c$ iff x bears label c (this defines the labeling of the nodes);
- $\rightsquigarrow$ is the union of all the E_f, where $f \in F$ (this is the relation determined by the entire set of edges, independent of their edge-labels);
- $\overset{+}{\rightsquigarrow}$ is a subset of $D \times D$, the transitive closure of $\rightsquigarrow$ (this is the relation 'is connected to by a sequence of one or more directed edges', which holds between a node and some other node that is ultimately reachable from it);
- $\lhd$ is a subset of $D \times D$ that totally orders the elements of D that are maximal with respect to $\overset{+}{\rightsquigarrow}$ (this is the 'immediately followed by' relation on words in an expression);

$\leq$ is a subset of $D \times D$, the transitive closure of $\leq$ (this is the 'somewhere earlier than' or 'precedes' relation on the words in an expression).

We require that such a structure is *rooted* with respect to $\overset{+}{\rightsquigarrow}$ (that is, $\mathcal{D} \models (\exists x)(\forall y)[x \overset{+}{\rightsquigarrow} y]$), so there is a node from which every other is reachable, and also that it is acyclic with respect to $\overset{+}{\rightsquigarrow}$ (that is, $\mathcal{D} \models (\forall x, y)[\neg(x \overset{+}{\rightsquigarrow} y \wedge y \overset{+}{\rightsquigarrow} x)]$), so by following the $\rightsquigarrow$ arrows you never come back to a node you have visited before.

From now on, as a convenience of exposition, when we talk about a structure $\mathcal{D}$ we will always assume we can refer to its domain as D, to its set of labeled edges as E_f, to its reachability relation as $\overset{+}{\rightsquigarrow}$, and so on. This saves much tedium: we can say 'in E_f' instead of 'in the edge set $E_f^{\mathcal{D}}$ that is the relevant one for the structure $\mathcal{D}$'.

To make mention of an f-flavoured edge from node x to node y in a structure, we'll use prefix syntax, writing '$E_f(x, y)$'; and similarly, when talking about a node x labeled Z we'll write '$Z(x)$'. But we will generally use infix syntax for the binary relations $\rightsquigarrow$, $\overset{+}{\rightsquigarrow}$, $\triangleleft$, and $\leq$; for example, we'll write '$x \overset{+}{\rightsquigarrow} y$' when talking about there being a chain of edges leading from x to y.

3 Faithful interpretation of logical theories

We turn now to the development of a descriptive metalanguage for talking about trees and *CGEL*-structures, and the central technique we want to illustrate, which is called **faithful interpretation**. We'll begin with a consideration of how to describe trees. Following Rogers (1994, 1998, 2003), we'll be using **monadic second-order logic** (henceforth **MSO**) interpreted on relational structures with tree properties.

MSO is a variant of the predicate calculus in which, in addition to the ordinary (first-order) variables ranging over individuals, there are variables ranging over sets of individuals.[2]

In the usual logician's parlance, a logical **theory** is simply a set of formulae that is closed under logical consequence. The **MSO theory of trees** is the set of all and only the MSO sentences satisfied by any tree structure whatever.

We will call an MSO language **suitable** for a structure $\mathcal{D}$ if the language has monadic predicate symbols for all the symbols labelling the nodes in $\mathcal{D}$, and binary predicate symbols for the relevant binary relations on $\mathcal{D}$, and so on.

A set $\mathcal{T}$ of trees is **definable** in the MSO theory of trees (or **MSO-definable**) iff there is a set of axioms (closed formulae) in the MSO language suitable for $\mathcal{T}$ which is satisfied by all and only those trees in the set.

(In linguistics we are almost always concerned with a proper subset of tree structures, the finite ones, since linguists have almost no dealings with infinite trees. But those who think that insisting on the finiteness of syntactic structures is important, we note that the MSO theory of finite trees is MSO-definable within the MSO theory of all trees, because notions like 'the domain is finite' are expressible in MSO — though not in first-order logic.)

A **faithful interpretation** of one logical theory into another is, in informal terms, a mapping from the first set of statements to the other that preserves the key aspects of meaning.

To be more precise, suppose L_1 is a logical language suitable for some class of structures $\mathbb{D}_1$ and L_2 is a language suitable for another class of structures, $\mathbb{D}_2$. A faithful interpretation of an L_1 theory into an L_2 theory is a uniform mapping of the domain of each structure in $\mathbb{D}_1$ to a subset of the domain of a structure in $\mathbb{D}_2$ in such a way that when $\mathcal{D}_1 \in \mathbb{D}_1$ is mapped to $\mathcal{D}_2 \in \mathbb{D}_2$,

(i) the range of the mapping is a non-empty L_2-definable subset of D_2;
(ii) each of the predicates of the signature of L_1 is definable on that domain in L_2; and
(iii) the mapping and the defined predicates are such that satisfiability of formulae in L_1 is preserved in L_2.

[2]Remarkably little use is made of quantification over sets in most linguistic description, we have found. The reason for fixing officially on MSO rather than first-order logic has to do with a remarkable equivalence result, that of Doner 1970, which we refer to later on.

Satisfiability is simply the property of having a model. A formula is satisfiable iff there is a structure $\mathcal{D}$ that makes the formula true when evaluated in $\mathcal{D}$. (Notice that it is possible for it to be unknown whether or not there is such a $\mathcal{D}$.)

To explain the foregoing more intuitively, the idea is to set up a mapping from *CGEL* structures to trees in such a way that (i) the nodes of the *CGEL* structure are mapped to nodes in the tree in a way definable in the metalanguage used for talking about trees; (ii) for any predicate in the metalanguage for the *CGEL* structures, the metalanguage for trees can define the set of nodes that predicate applies to; and (iii) when a formula of the metalanguage for the *CGEL* structures is satisfiable, the translation of that formula into the metalanguage for talking about trees is satisfiable too.

4 Embedding theories of *CGEL* structures into theories of trees

The definitions of the domain and predicates of $\mathbb{D}_1$ (the *CGEL* structures) in L_2 (the language of trees) provide a syntactic translation of formulae in L_1 into formulae of L_2. Quantification is relativized to the range of the mapping, and each predicate symbol in L_1 is replaced with its definition in L_2. Since the interpretation preserves satisfiability, a sentence of L_1 will be in the first theory iff its translation into L_2 is in the second. In this way decidability of the first theory can be reduced to decidability in the second.

Translating theories into theories (providing a faithful interpretation of the MSO theory of classes of *CGEL* structures into the MSO theory of rooted, directed, ordered trees) thus permits us to establish the decidability of theories based on *CGEL* structures. But we can go further than that: we can actually interpret theories of *CGEL* structures in trees via an interpretation that preserves the linear order of the maximal points (the frontiers, or terminal strings) of *CGEL* structures, so that we preserve exactly the set of strings that a grammar describes.

The definitions we have given allow a *CGEL* structure to be mapped to any MSO-definable subset of the domain of some tree, but we will focus our attention on translations in which the mapping from the domain of the *CGEL* structure to the domain of the tree is a **bijection** (a one-to-one correspondence). In that case we can take the mapping to be identity, and build the trees on exactly the same set of nodes as the *CGEL* structure. This makes the range of the mapping trivially MSO-definable, so we easily attain compliance with condition (i) above.

The strategy we use is to build a **spanning tree** of the graph corresponding to the edges of the *CGEL* structure (all of the edges that make up the $\rightsquigarrow$ relation). This will be a tree that, intuitively does not contradict the *CGEL* structure in its properties, or lose any of its crucial information. More precisely, we define the notion of a directed spanning tree thus:

Definition 1 (Directed Spanning Tree) *A directed, unordered, unlabeled tree is a* directed spanning tree *of a CGEL structure iff the domains are the same and the 'parent-of' relation in the tree is included in the 'immediately reachable' relation in the CGEL structure. (That is,* $\mathcal{T} = \langle T, \triangleleft_2, <_2 \rangle$ *is a directed spanning tree of* $\mathcal{D} = \langle D, E_f, N_c, \rightsquigarrow, \overset{+}{\rightsquigarrow}, \triangleleft, \leq \rangle_{f \in F, c \in C}$ *iff (i)* $T = D$ *and (ii)* $\triangleleft_2 \subseteq \rightsquigarrow$*.)*

Since we have preserved the domain of the *CGEL* structure, there is nothing to be done with regard to translation of the node labels — we can simply allow the interpretation of category-label predicates in the trees (**Clause**, **NP**, and so on) to be the very same sets of nodes that interpret those symbols in the *CGEL* structures. We simply build the spanning tree on the nodes of the *CGEL* structure.

The remaining work that must be done, then, is to translate the edge relations. The translation of E_f (an edge labelled with the function f) has three components, which we will represent as a conjunction of formulae using three defined predicates:

— φ_f^O picks out the points which may have an f out-edge;
— φ_f^I picks out the points which may have an f in-edge; and
— $\varphi_f^E(x, y)$ identifies all and only those pairs drawn from those sets that actually are joined by an f-edge.

So φ_f^O is the property shared by all those nodes from which an f edge is permitted to depart, and φ_f^I is the property shared by all those nodes at which an f edge is permitted to arrive, and $\varphi_f^E(x, y)$ means that y is an f of x.

Wherever $\varphi_f^E(x, y)$ holds, $\varphi_f^O(x)$ and $\varphi_f^I(y)$ also hold. Thus the source end and destination end of each f-edge are tagged as permitted to be in those locations. Each formula $E_f(x, y)$ in a theory of *CGEL* structures is translated to:

> $\varphi_f^O(x) \ \wedge\ \varphi_f^I(y) \ \wedge\ \varphi_f^E(x, y)$
> ("x is allowed to be the source of an outgoing f-edge and y is allowed to be the destination of an incoming f-edge and there is an f-edge from x to y")

(In practice, either or both of φ^O and φ^I may be trivial; that is, all points may be permitted to be out-edges or in-edges for any edge label. But there is provision for restricting such things as the categories that can bear particular grammatical relations.)

Definition 2 (Reachability preservation) *A directed spanning tree of a CGEL structure* preserves reachability *iff reachability (between a pair of nodes in the CGEL structure) implies proper domination (between those nodes in the spanning tree).*

More formally, we say that a directed spanning tree $\mathcal{T} = \langle T, \triangleleft_2, <_2\rangle$ preserves reachability *with respect to a CGEL structure* $\mathcal{D} = \langle D, E_f, N_c, \rightsquigarrow, \overset{+}{\rightsquigarrow}, \triangleleft, \leq\rangle_{f\in F, c\in C}$ *(where of course* $D = T$*) iff* $(\forall a, b \in T)[(a \overset{+}{\rightsquigarrow} b) \ \Rightarrow\ (a <_2 b)]$.

So if you can get from a to b in the *CGEL* structure by following edges in the direction of the arrow (away from the root), then a properly dominates b in any spanning tree that preserves reachability.

Notice, since $\mathcal{T}$ is a spanning tree of $\mathcal{D}$ it will be the case that for all $a, b \in D\,(= T)$ we have $(a <_2 b) \ \Rightarrow\ (a \overset{+}{\rightsquigarrow} b)$. (This follows because the $\triangleleft_2$ edges of $\mathcal{T}$ are a subset of the $\rightsquigarrow$ edges of $\mathcal{D}$.) Hence the implication is actually an equivalence.

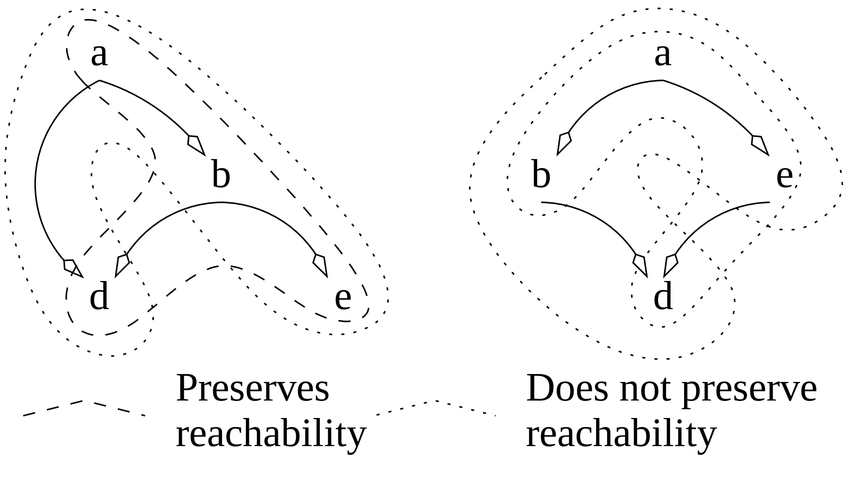

Figure 2: Two directed ordered acyclic graphs each having two spanning trees

Now, every *CGEL* structure has at least one directed spanning tree, but since the tree is directed and $\triangleleft_2$ may be a proper subset of $\rightsquigarrow$ it may be the case that no directed spanning tree for a given *CGEL* structure preserves reachability (see Figure 2.)

If reachability can be preserved, then the translation of the edge relations can be **direction-preserving**: if the edge (in the *CGEL* structure) goes from a to b, then in the spanning tree the image of a will dominate that of b. Hence, the set of pairs that satisfy $\varphi_f^E(x, y)$ will be a subset of the set interpreting $<_2$: the only place we need to look to find the nodes at the end of edges leading from a is in the subtree rooted at a in the tree.

5 Property I: branch ordering

Things now begin to move fairly swiftly. In what follows we will give proofs in sketch form at best, and sometimes not at all. We begin by considering the implications of an arbitrary structure $\mathcal{D}$ having the following property (that is, the property of satisfying the statement that for convenience of reference we will call Property I):

$$\boxed{\textbf{Property I}: \quad (\forall x, y, z)[(x \overset{+}{\rightsquigarrow} z \wedge y \overset{+}{\rightsquigarrow} z) \rightarrow (x \approx y \vee x \overset{+}{\rightsquigarrow} y \vee y \overset{+}{\rightsquigarrow} x)]}$$

Property I requires that for any three points in $\mathcal{D}$, if the third can be reached from either of the first two, then either the first two are identical or one can be reached from the other. In essence, 'branches' of $\mathcal{D}$ (sets of nodes falling between two given nodes with respect to $\overset{+}{\rightsquigarrow}$) are totally ordered. Our first theorem is this:

Theorem 1 *A CGEL structure* $\mathcal{D} = \langle D, E_f, N_c, \rightsquigarrow, \overset{+}{\rightsquigarrow}, \lhd, \leq \rangle_{f \in F, c \in C}$ *has a directed spanning tree that preserves reachability iff it has Property I.*

Proof (sketch): To show that Property I implies that $\mathcal{D}$ has a directed spanning tree that preserves reachability, by construction, given a *CGEL* structure

$\mathcal{D} = \langle D, E_f, N_c, \rightsquigarrow, \overset{+}{\rightsquigarrow}, \lhd, \leq \rangle_{f \in F, c \in C}$

let $\mathcal{T}^{\mathcal{D}}$ denote the structure $\langle D, \lhd_2{}^{\mathcal{D}}, <_2{}^{\mathcal{D}}, N_c \rangle_{c \in C}$ defined by the following conditions:

$$\lhd_2{}^{\mathcal{D}} \quad = \quad \{\langle a, b \rangle \mid a \rightsquigarrow b \quad \text{and } (\forall c \in D)[a \overset{+}{\rightsquigarrow} c \ \rightarrow \ \neg(c \overset{+}{\rightsquigarrow} b)]\ \}$$

and $<_2{}^{\mathcal{D}}$ is the transitive closure of the $\lhd_2{}^{\mathcal{D}}$ relation. That is, there is a 'parent of' edge between each point a and all points which are minimal (with respect to $\overset{+}{\rightsquigarrow}$) among those reachable from a in $\mathcal{D}$. In effect this discards all edges of $\mathcal{D}$ for which there is a (directed) path (using any of the edge types) containing two or more edges between the same pair of points. This preserves reachability, because an edge will be dropped only if there is such a path.

The fact that reachability is preserved, along with the fact that the *CGEL* structure is acyclic, transitive, and rooted, implies that $<_2{}^{\mathcal{D}}$ is asymmetric and transitive, and that $\mathcal{T}^{\mathcal{D}}$ has a unique minimum with respect to the $<_2$ relation. It is straightforward to show that Property I also ensures that each point other than the root has a unique predecessor with respect to $<_2{}^{\mathcal{D}}$.

For the other direction we use contraposition. If $\mathcal{D}$ fails to exhibit Property I then no subgraph that covers the vertices of $\mathcal{D}$ can be a tree. ■

All the function-fusion analyses given in *CGEL* appear to yield structures that satisfy Property I. In fact they appear to satisfy the following statement, which is much stronger:

$$\boxed{\textbf{Property I}^{+}: \quad (\forall x, y, z)[(x \overset{+}{\rightsquigarrow} z \ \wedge \ y \overset{+}{\rightsquigarrow} z) \ \rightarrow \ (x \approx y \ \vee \ x \lhd_2 y \ \vee \ y \lhd_2 x)]}$$

This requires reachability between x and y in a single step: function fusion at a node z involves branches which connect either to a single parent node or to a pair of nodes in which one is child of the other (see Payne et al. 2007:566–584).

Even in the form using Property I, the construction in the proof of Theorem 1 is fully constrained in the sense that, for any node a, if a smaller set of children were selected there would be some point that is reachable from a in $\mathcal{D}$ but not dominated by a in $\mathcal{T}^{\mathcal{D}}$ (since the children of a in the construction are all minimal with respect to $\overset{+}{\rightsquigarrow}$ among those reachable from a) and if a larger set of children were selected $\lhd_2{}^{\mathcal{D}}$ would either not be a subset of $\overset{+}{\rightsquigarrow}$ or there would be some $a \neq c$ and b for which $\langle a, b \rangle, \langle c, b \rangle \in \lhd_2{}^{\mathcal{D}}$. We therefore have this corollary:

Corollary 1 *If* $\mathcal{D}$ *exhibits Property I then there is a* unique *directed spanning tree of* $\mathcal{D}$ *which preserves reachability.*

6 Property II: reachability/precedence exhaustiveness

Thus far we have not said anything about our spanning trees being left-to-right ordered. We now add left-to-right order, and then define a notion of compatibility between trees and *CGEL* structures that respects both reachability and order.

Definition 3 (Directed ordered spanning tree) *A directed, ordered, unlabeled tree is a* ***directed ordered spanning tree*** *of a CGEL structure* $\mathcal{D}$ *iff* $\langle T, \triangleleft_2, <_2 \rangle$ *is a directed spanning tree of* $\mathcal{D}$.

Definition 4 (Compatible directed ordered spanning tree) *A directed ordered spanning tree*

$$\mathcal{T} = \langle T, \triangleleft_1, <_1, \triangleleft_2, <_2 \rangle$$

of a CGEL structure $\mathcal{D}$ *is* ***compatible*** *with* $\mathcal{D}$ *iff it preserves reachability and it* ***preserves order****:*

$$(\forall a, b)[a \leq b \quad \Leftrightarrow \quad a <_1 b \text{ and } a, b \text{ are both leaves of } \mathcal{T}\ ]$$

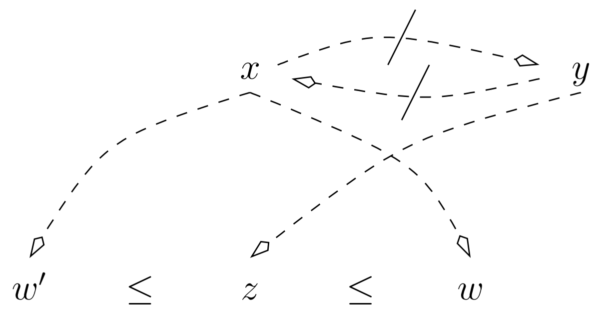

Figure 3: A directed ordered acyclic graph with tangling

Let Leaf(x) be explicitly defined as $(\forall y)[\neg(x \rightsquigarrow y)]$; that is, it means that x is a leaf node: from x you cannot reach anything, because x is not at the source end of any edge. Let $x \overset{*}{\rightsquigarrow} y$ be explicitly defined as $x \overset{+}{\rightsquigarrow} y \vee x \approx y$, i.e., as reachability in zero or more steps. And let $\prec^+_{\mathcal{G}}$ be a relation extending $\leq$ to (a subset of) non-maximal points in $\mathcal{D}$ in the following way:

$x \prec^+_{\mathcal{G}} y$ is explicitly defined as meaning

$$(\forall w, z)[\neg(x \overset{*}{\rightsquigarrow} y) \ \wedge \ \neg(y \overset{*}{\rightsquigarrow} x) \ \wedge \ (\ (x \overset{*}{\rightsquigarrow} w \wedge \text{Leaf}(w) \wedge y \overset{*}{\rightsquigarrow} z \wedge \text{Leaf}(z)) \ \rightarrow \ (w \leq z)\ )\ ]$$

Thus $x \prec^+_{\mathcal{G}} y$ means that there are no w and z such that (i) you can't reach y from x in zero or more steps, and (ii) you can't reach x from y in zero or more steps, and (iii) if from x you can reach the leaf w and from y you can reach the leaf z then w precedes z.

Note that $\prec^+_{\mathcal{G}}$ is asymmetric and transitive and that every structure $\mathcal{D}$ will satisfy $(\forall x, y)[(x \prec^+_{\mathcal{G}} y \vee y \prec^+_{\mathcal{G}} x) \ \rightarrow \ \neg(x \overset{+}{\rightsquigarrow} y)]$. Let **Property II** be this property of an arbitrary structure $\mathcal{D}$:

Property II: $(\forall x, y)[x \approx y \vee x \overset{+}{\rightsquigarrow} y \vee y \overset{+}{\rightsquigarrow} x \vee x \prec^+_{\mathcal{G}} y \vee y \prec^+_{\mathcal{G}} x]$

Property II requires that, for any two distinct points in $\mathcal{D}$, either one can be reached from the other or one precedes the other. That means that $\overset{+}{\rightsquigarrow}$ and $\prec^+_{\mathcal{G}}$ together totally order D, in the same sense that $<_2$ and $<_1$ totally order a tree. Since $\prec^+_{\mathcal{G}}$ is defined only for nodes that do not interleave the maximal nodes reachable from them, the property, in essence, requires that the edges of $\mathcal{D}$ do not "tangle" with respect to the ordering of its maximal nodes (as in Figure 3). This ensures that an analogue of the usual no-tangling property in trees will hold, leading to our next theorem:

Theorem 2 *A CGEL structure* $\mathcal{D} = \langle D, E_f, N_c, \rightsquigarrow, \overset{+}{\rightsquigarrow}, \triangleleft, \leq \rangle_{f \in F, c \in C}$ *has a compatible directed ordered spanning tree iff it satisfies Property I and Property II.*

The proof is a routine application of the definitions, and we omit it.

7 Property III: function segregation

Properties I and II suffice to guarantee that a *CGEL* structure can be reduced to a directed spanning tree in a way that respects the orderings imposed by the edges and the ordering of its maximal points. But the edge set of the spanning tree is, in general, a proper subset of the edges of the *CGEL* tree and not even a homogeneous subset with respect to to the edge labels. We face the question of whether it is possible to encode the edges of the *CGEL* structure in the spanning tree.

It is quite clear that the answer is yes. The approach we adopt here can be conceptualized as sliding the edge labels down the lines onto the nodes that the edges connect to. The structure for a noun phrase with Determiner–Head fusion could be conceptualized as having a single Determinative node at the head of two distinct edges, as in Figure (4a).

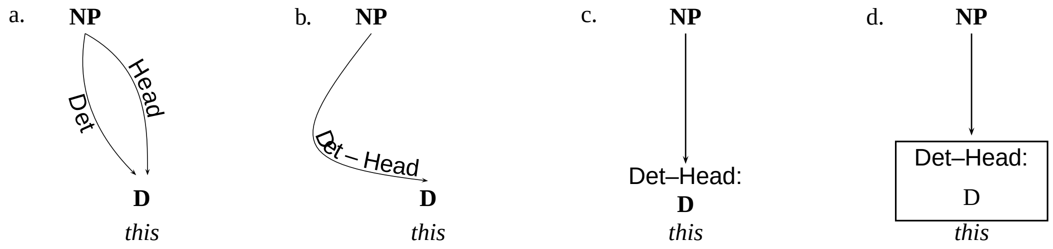

Figure 4: Four ways to represent the structure of the Det–Head noun phrase *this*

But the two function names might just as well be associated with the same edge, and we could slide the double function label down onto the end near the node label, as shown in Figure 4 (b).

Indeed, sliding it off the arrow head and putting it right above or immediately before the category label of the node, separated by a colon, yields the usual notation used in the diagrams in *CGEL* itself, as in Figure 4 (c).

And this makes it clear that no issue of expressive power is going to arise as a result out of the edge labels, since the description could be re-implemented in a way that took objects like the contents of the box in Figure 4 (d) to be simply categories.

If we just cross-multiplied the 16 categories and 20 functions listed in Table 1 and Table 2 it would yield no more than 320 combined function + category pairs, which is not unmanageable.

However, such cross-multiplication is merely a way to convince oneself that no problem of principle arises here. It is not how *CGEL* conceives of the analysis, and it is not necessary.

MSO is powerful enough to encode the fact that $\langle a, b \rangle$ belongs to E_f in a given structure by adding two sets of new monadic second-order predicates to our metalanguage: O_f and I_f. We assert that a is in the set assigned to O_f (meaning 'there is an f edge that is an out-edge from a') and b is in the set assigned to I_f ('there is an f edge that is an in-edge to b').

These new predicates do not need to be incorporated as extensions to the node label alphabet: they are not primitive to the theory modeled by the *CGEL* structures, and they can be regarded as a distinguished set of monadic second-order variables which will ultimately be bound with a global existential quantifier.

So $E_f(x, y)$ (in the vocabulary suitable for the *CGEL* structure) will be translated as $O_f(x) \wedge \varphi_f(x, y) \wedge I_f(y)$ in the tree-description language, where $\varphi_f(x, y)$ picks out the structural relationships which can hold in the spanning tree between points that are related by E_f in *CGEL*.

The definition of φ_f is specific to the theory expressed by the set of *CGEL* structures. For this approach to work, it must be possible to define this in a way that captures every possible relationship between the two ends of an f edge but which is also unambiguous in picking out the actual members of the sets assigned to I_f and O_f that are related by E_f in the *CGEL* structure.

Note that, if a and b are related by an f edge in the *CGEL* structure then in any compatible ordered directed spanning tree it will be the case that a properly dominates b, since compatible spanning trees preserve reachability. So the I_f node associated with a given a that is the source of an f edge will always be in the subtree dominated by a.

One condition that is sufficient to guarantee that such a φ_f can be uniformly defined is the property of *CGEL* structures we shall call Property III:

$$\boxed{\textbf{Property III}: \quad (\forall x, y, z, q)[(E_f(x, y) \land E_f(z, q) \land x \overset{+}{\rightsquigarrow} z) \rightarrow \neg(z \overset{+}{\rightsquigarrow} y)\ ]}$$

Here's what this means. Suppose there is an edge labelled f from x to y, and another f edge from z to some arbitrary node q. If you can get from x to z along some path down the edges, then you can't get from z to y. Thus there is a sense in which the f edges do not overlap: you never find a path between two nodes related by an f edge that includes a node at the source end of some different f edge. So if there's an f edge from x to y, then x the only O_f node that occurs on any path between x and y.

Now suppose a node a that is the source of an f edge dominates two distinct nodes b and c that are destinations of f edges. For concreteness, assume that b is the node related to a by an f edge. Then the member of O_f corresponding to c (call it d) must fall between a and c with respect to proper domination. Otherwise a is reachable from d and c is reachable from a, and that would violate Property III.

This the f-destination node corresponding to a given f-source node a will always be that f-destination node in the subtree rooted at a for which no f-source node intervenes. Thus $E_f(x, y)$ will translate to this:

$$O_f(x) \quad \land \quad (x <_2 y \land (\forall z)[(x <_2 z \land z <_2 y) \rightarrow \neg O_f(z)) \quad \land \quad I_f(y)$$

The foregoing is enough to establish our main result, presented as Theorem 3:

Theorem 3 *If an MSO-definable set $\mathbb{D}$ of CGEL structures satisfies Properties I, II, and III, then there is a faithful interpretation of the theory of $\mathbb{D}$ into the MSO theory of trees.*

We omit the proof, which is straightforward.

8 Function fusion in English grammar

We have not shown so far that the three properties, in particular Property III, will always be satisfied by *CGEL*-style analyses. To this extent we are offering only a conjecture about the *CGEL* descriptive framework, not a mathematical result. We have not even given a rigorous survey of all the relevant cases of structures proposed in *CGEL* itself. However, we have inspected what appear to be the only relevant analyses, and Property III seems entirely plausible. We provide a brief survey of the relevant constructions (see Huddleston et al. 2005, 98–100 for a convenient informal summary of the facts, and Payne et al. 2007, 566–584 for a detailed theoretical discussion).

Fused functions are found as alternative constructions to **NP**s and **PP**s, and involve single-word realisation of what would otherwise have been linearly adjacent Head and Dependent constituents.

In all the cases we are aware of, actual counterpart constructions exist in which the Head and Dependent constituents are separate. Thus alongside **NP**s like *everyone* (where a single word is both Determiner and Head) there are counterparts like *every person*; alongside **NP**s like *the French* (where *French* is both Modifier and Head) there are counterparts like *The French people*; and so on. In the counterpart construction we have a node n_x of some category X with a child node n_a realising some function F_a, and either a child or a grandchild n_b realising some different function F_a. In all cases the nodes n_a and n_b are linearly adjacent, and (a substantive observation not reflected in the formal account above) in all cases either F_a or F_b is the function Head.

Payne et al. (2007) identifies some further generalisations. Let us temporarily refer to the node realising the Head function as n_h (from what we have said it follows that either $n_h \approx n_a$ or $n_h \approx n_b$), and to the node realising the Head function as n_d (so $n_h \not\approx n_d$). The category of n_h is always in the same projection class as that of n_x (in the sense that X-bar theory seeks to formalise: in the *CGEL* system the

N projection class is {**N**, **Nom**, **NP**}, the **V** projection class is {**V**, **VP**, **Clause**}, the **P** projection class is {**P**, **PP**}, and so on). Moreover, n_d always has a category Y distinct from X.

Now, the difference in the function-fusion construction is simply that a single node realises both F_a and F_b, so $n_h \approx n_d \approx n_a \approx n_b$.

Some examples will make this clearer.

8.1 Plain Det–Head fusion

CGEL uses Det–Head fusion in the structure of **NP**s like *this*, *that*, *many*, *several*, *everyone*, *nobody*, *something*, *none*, etc., as seen represented in various ways in Figure 4, where a single Determinative functions as Determiner of the whole **NP** and Head of its **Nominal** (non-branching **Nominal** constituents are omitted from diagrams in *CGEL*). So we have an **NP** node n_x with a **D** child node realising both the Determiner and Head functions.

8.2 Partitive Det–Head fusion

Partitive examples like the one illustrated in Figure 1 are a slight variant of this, where the Determinative is Determiner of the whole **NP** and Head of a branching **Nominal** that also contains the *of*-headed Complement **PP** that accompanies it. A node n_x of category **NP** has a Determiner child n_a of category **D**, and a Head child of category **Nom**, and that **Nom** has a child n_b realising the Head function, and $n_a \approx n_b$.

8.3 Plain Modifier–Head fusion

CGEL also posits Modifier–Head fusion under almost exactly parallel circumstances as Det–Head fusion, in the structure of certain definite **NP**s that have a modifying attributive Adjective but no Noun to serve as Head: *the first* (≡ "the one that is first"), *the youngest of them* ("the youngest child from among their five children"), *the absurd* ("that which is absurd"), *the rich* ("rich people considered as a class"), *the French* ("the people of the French nation considered as a class"), etc.

8.4 Partitive Modifier–Head fusion

Modifier–Head fusion also has a partitive variant, where an *of*-**PP** is associated with the modifier-head, as in *the youngest of them*, which is diagrammed in Figure 5.

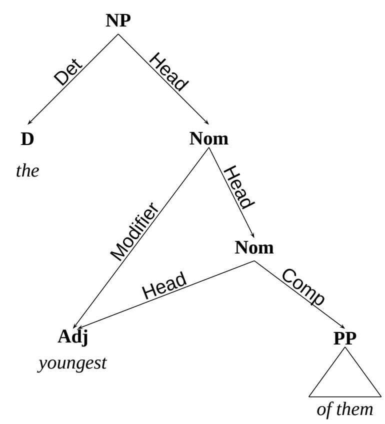

Figure 5: The structure of the **NP** *the youngest of them*

8.5 Det–Head and Mod–Head fusion with DPs and PPs

Entirely parallel to the foregoing cases are certain further cases of Determiner–Head and Modifier–Head fusion posited in Payne et al. (2007). Again these are found at the beginning of **NP**s, but the constituents with fused functions belong not to lexical categories like **D** or **Adj** but to phrasal categories: **DP**s (which are not **NP**s!) in phrases like *at least once* and *more than once* (see p. 590), and **PP**s (see p. 589, n. 31). Otherwise they introduce nothing new. (One attested example with a **DP** in fused Modifier–Head function is shown in Payne et al.'s diagram (24).)

8.6 Head–Prenucleus fusion in fused relatives

What *CGEL* calls **fused relatives** are **NP**s such as *what she wrote*, in which the word *what* is both the pronoun lexical Head of (the head **Nominal** of) the **NP** and the Prenucleus in the relative clause that functions as Modifier of that **Nominal**. The structure of *what she wrote* (as seen in *CGEL*, p. 1073) is shown in Figure 6.

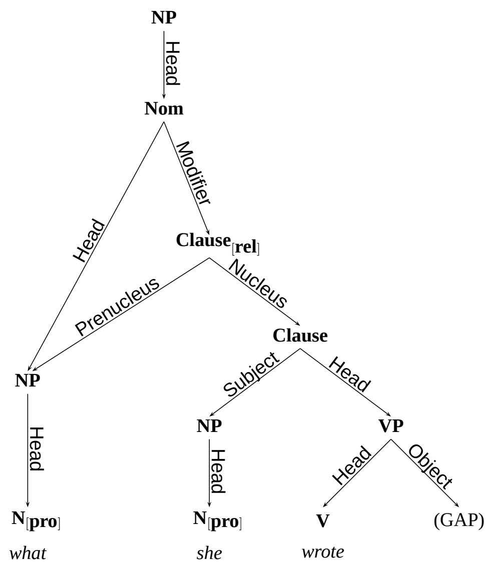

Figure 6: The structure of the **NP** *what she wrote*

For all of these cases, the structures posited in *CGEL* satisfy Properties I – III.

In particular, if anyone could find evidence that a *CGEL*-style analysis of some aspect of English syntax would be best formulated in terms that involve a structure incompatible with Property III, that would be an interesting discovery. We should point out, however, that Property III as given is not crucially necessary: it could be weakened in a number of ways without losing the result that the theory of *CGEL* structures can be faithfully interpreted in the MSO theory of trees. For instance, if the f edges can be refined into some arbitrary finite number of subtypes such that each of the subtypes satisfies Property III, then a similar strategy could be implemented with the subtypes.

The assumptions we have made to obtain the correspondence between *CGEL* structures and trees are thus not at all restrictive. We require only that there be finitely many (sub)types of edge such that the set of paths for any given (sub)type of edge in the *CGEL* structure — the paths picking out all and only the pairs of points that are related by that (sub)type of edge — can be defined using MSO on trees. These can be quite subtle definitions, possibly depending on the labels of the nodes and the edges that occur along the paths, and they may be inductively defined (as subsets of the tree). Similar remarks obtain concerning the description of unbounded dependencies.

9 Logical and language-theoretic consequences

The foregoing survey of function fusion constructions suggests that Property 3 can be assumed to hold in the kind of *CGEL* structures we are interested in, at least when considering the description of English. Theorem 3 depends on that assumption.

If the theorems proved here do indeed hold for all analyses of English constructions, we immediately get a number of interesting ancillary results about the set of all finite structures corresponding to English expressions, i.e., what would be called 'the language' in formal language theory terms. First, as a consequence of well-known facts about MSO on trees, we have this corollary:

Corollary 2 *If an MSO-definable set* $\mathbb{D}$ *of CGEL structures satisfies Properties I, II, and III, then the MSO theory of* $\mathbb{D}$ *is decidable.*

This means that the set of consequences of the axioms is a recursive set — whether a given formula φ is in it can be decided by an algorithm. This immediately (in fact trivially) gives us another corollary:

Corollary 3 *If an MSO-definable set* $\mathbb{D}$ *of CGEL structures satisfies Properties I, II, and III, then satisfiability for the MSO theory of* $\mathbb{D}$ *is decidable.*

This means that for a given MSO grammar statement φ it is possible to determine mechanically, by an algorithm, whether or not any structure could be well formed according to φ. It follows because theoremhood and satisfiability are logically related: φ is a theorem iff $\neg\varphi$ is not satisfiable, and φ is satisfiable iff $\neg\varphi$ is not a theorem. Since for any φ we can find out in finite time whether φ (or $\neg\varphi$) is in the theory of $\mathbb{D}$, we can also find out whether φ (or $\neg\varphi$) is satisfiable.

Next, in virtue of theorem due to John Doner 1970, we have a result linking the logic-based description of structures to a particular kind of automaton:

Corollary 4 *If an MSO-definable set* $\mathbb{D}$ *of CGEL structures satisfies Properties I, II, and III, it is recognizable by a a bottom-up finite-state tree automaton.*

And that, by a well-known theorem of Thatcher 1967, gives us one further result:

Corollary 5 *If an MSO-definable set* $\mathbb{D}$ *of CGEL structures satisfies Properties I, II, and III, then the string yield of* $\mathbb{D}$ *is context-free.*

We have thus proved that if Properties I – III hold of legitimate *CGEL* analyses for English, the set of grammatical strings entailed by a grammar stated in the form of a set of MSO-expressed constraints on *CGEL* structures will be a CFL. And we have obtained this result without reference to any of the usual mathematical linguistic notions: context-free grammars, pumping lemmas, pushdown automata, or generative grammars of any sort.

10 Conclusion and summary

We have shown that it is straightforward to represent the structures assumed in Huddleston et al. 2002 as relational structures, and that under fairly mild and plausible conditions MSO theories interpreted on such structures can be faithfully reinterpreted on tree models in a way that preserves all significant syntactic consequences, specifically, the claims made about constituency, node labeling, and word order.

Since a set of finite tree models definable by an MSO theory is always accepted by some bottom-up finite-state tree automaton (Doner 1970), and the string yield of such a set of trees is a CFL (Thatcher 1967), our conjecture would imply, under the conservative assumption that all of the statements informally made in that work are expressible in MSO, that the comprehensive description of English in *CGEL* supports a positive answer to a question first raised by Chomsky in 1956: whether the set of all grammatical strings of English words is a CFL. (The reason that we call it a conservative assumption that MSO will suffice to state any grammatical generalization of the sort found in *CGEL* should be clear to anyone who spends a little time reflecting on how various statements of that sort can be stated precisely. MSO is really a very rich and flexible language for talking about trees. But we do not, of course, substantiate the claim here; it stands as a conjecture.)

A positive answer to the question of whether English is CF would agree with the answer given in Gazdar 1981 and Gazdar et al. 1985. It would also be in tacit agreement with the claim that parsers of the sort characterized by Marcus (1980) can analyze English — though this was not known until the result by Nozohoor-Farshi (1986, 1987), which showed that Marcus's parsers for transformational grammars could only recognize CFLs.

Even more surprisingly, it is in tacit agreement with nearly all of the GB literature of the 1980s, though this was not known until Rogers 1998 pointed out that all of the content of GB theory seemed to be expressible in the form of MSO constraints on trees.

Various bases for thinking English cannot be a CFL have been presented in the literature of the last fifty years. It was once thought that unbounded dependencies were sufficient to raise doubts, but Gazdar (1981) showed very clearly that this was not likely to be a stumbling block, at least for English (it seems that some Scandinavian languages may be a different matter).

There has not been any convincing argumentation regarding the CFLs as too small a class to allow for the description of English. Pullum and Gazdar 1982 demonstrated the failure of all the best known arguments. Pullum 1985b; 1985a dispatched two others. Alexis Manaster-Ramer came up with what at first appeared to be a convincing argument based on the apparent phrasal reduplication in the construction illustrated by *Cold War or no Cold War* (though he did not offer any detailed discussion), but Pullum and Rawlins 2007 have recently shown that on closer examination the argument is not at all convincing.

There is thus good reason for linguists to continue to take an interest in the wealth of syntactic theories that (whether by accident or design) endorse the claim, because while it is compatible with interesting and plausible descriptions of the system of unbounded dependencies in English and the coordination possibilities of the language, it also meshes in well-understood ways with a very large body of work in computational linguistics concerning context-free parsing.

There do seem to be some human languages that, considered as stringsets, do not correspond to CFLs (Shieber 1985). But it should be kept in mind that it follows from Theorem 13 of Rogers 2003 that non-context-free stringsets describable by tree adjoining grammars (TAGs) can be described using MSO constraints on the '3-dimensional' tree-like structures described there. There is thus no problem of principle about providing MTS descriptions of at least some kinds of non-CFL stringsets using MSO on tree-like models. The question is just when and for what such descriptions will be needed. What we have just seen is that such additional power will apparently not be needed for anything that is expressible in terms of the theory we can take to be implicit in *CGEL*.

References

Doner, John. 1970. Tree acceptors and some of their applications. *Journal of Computer and System Sciences* 4:406–451.

Gazdar, Gerald. 1981. Unbounded dependencies and coordinate structure. *Linguistic Inquiry* 12:155–184.

Gazdar, Gerald, Ewan Klein, Geoffrey K. Pullum, and Ivan A. Sag. 1985. *Generalized phrase structure grammar*. Oxford: Basil Blackwell.

Huddleston, Rodney, Geoffrey K. Pullum, et al. 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.

Huddleston, Rodney, Geoffrey K. Pullum, et al. 2005. *A student's grammar of the English language*. Cambridge: Cambridge University Press.

Marcus, Mitchell. 1980. *A theory of syntactic recognition for natural language*. Cambridge, MA: MIT Press.

Nozohoor-Farshi, R. 1986. On formalizations of Marcus' parser. In *Proceedings of the 11th Conference on Computational Linguistics*, 533–535. Morristown, NJ, USA: Association for Computational Linguistics.

Nozohoor-Farshi, R. 1987. Context-freeness of the language accepted by Marcus' parser. In *Proceedings of the 25th annual meeting on Association for Computational Linguistics*, 117–122. Morristown, NJ, USA: Association for Computational Linguistics.

Payne, John, Rodney Huddleston, and Geoffrey K. Pullum. 2007. Fusion of functions: The syntax of *once*, *twice*, and *thrice*. *Journal of Linguistics* 43:565–603.

Potts, Christopher. 2005. *The logic of conventional implicatures*. Oxford: Oxford University Press.

Pullum, Geoffrey K. 1985a. On two recent attempts to show that English is not a CFL. *Computational Linguistics* 10:182–186.

Pullum, Geoffrey K. 1985b. *Such That* clauses and the context-freeness of English. *Linguistic Inquiry* 16:291–298.

Pullum, Geoffrey K., and Gerald Gazdar. 1982. Natural languages and context-free languages. *Linguistics and Philosophy* 4:471–504.

Pullum, Geoffrey K., and Kyle Rawlins. 2007. Argument or no argument? *Linguistics and Philosophy* 30:277–287.

Pullum, Geoffrey K., and Barbara C. Scholz. 2005. Contrasting applications of logic in natural language syntactic description. In *Proceedings of the 13th International Congress of Logic, Methodology and Philosophy of Science*, ed. Petr Hajek, Luis Valdes-Villanueva, and Dag Westerstahl, 481–503. London: KCL Publications.

Rogers, James. 1994. Studies in the logic of trees with applications to grammar formalisms. Doctoral Dissertation, University of Delaware, Newark, DE.

Rogers, James. 1998. *A descriptive approach to language-theoretic complexity*. Stanford, CA: CSLI Publications.

Rogers, James. 2003. wMSO theories as grammar formalisms. *Theoretical Computer Science* 293:291–320.

Shieber, Stuart. 1985. Evidence against the context-freeness of human language. *Linguistics and Philosophy* 8:333–343.

Thatcher, James W. 1967. Characterizing derivation trees of context-free grammars through a generalization of finite automata theory. *Journal of Computer and System Sciences* 1:317–322.