The NP Analysis of NP

Adam Meyers

New York University

1 Introduction

This paper supports the Government and Binding Unification Grammar (GBUG) analysis of NP in which the determiner is the predicate licensor and the noun is the head (The NP analysis).

In my account, the HEAD $H$ of a phrase $XP$ is the CONSTITUENT-LICENSOR of $XP$. The CONSTITUENT-LICENSOR is obligatory, barring ellipsis, and bears the features that are subject to semantic selection and subcategorization restrictions imposed on $XP$ (cf. (Zwicky, 1985) and (Zwicky, 1993)). The predicate licensor $P$ of a phrase $XP$ is the constituent which semantically selects the other constituents of $XP$, which determines the (exceptional) word order properties of $XP$, and which anchors $XP$. A constituent $P$ ANCHORS $XP$ if $P$ justifies the occurrence of $XP$. Anchoring is like “projecting from a head” in previous P & P approaches, except that the anchor of a constituent need not be the head. The term anchoring is borrowed from Lexicalized Tag Adjoining Grammar and Stratified Feature Grammar. My conceptions of Constituent LICensors, Predicate LICensors, and other licensors are closely related to head-like distinctions made in Categorial Grammar, (Zwicky, 1985), (Hudson, 1987), (Rothstein, 1991) and articles in (Corbett et al., 1993).

There are many different ways in which one constituent is the dominant element or licensor of a phrase. In GBUG each type of licensor is distinguished by its licensing relation: heads are licensors of constituent relations, theta role assigners are licensors of theta relations, modifiers are licensors of modification relations, etc. The head of a phrase $XP$ determines the category of $XP$, and thus the distribution of $XP$ in phrase structure. Choosing which type of licensor is the head makes very specific predictions about phrase structure. The error of the DP analysis is to assume that the head of the phrase must be the licensor of every type of licensing relation. This error is not obvious in the phrase structure tree model, because that model makes it difficult to concisely represent multiple licensing relations.

Proponents of the DP analysis and related hypotheses have the burden of either: (1) showing that the determiner does somehow determine the distribution of
nominal constituents, or (2) providing a different conception of head that justifies assigning these nominal constituents the category *determiner*.

### 2 Identifying the Head of the Noun Phrase

Some differences in choice of head have effects that should be considered relevant for any phrase structure-based theory. On the simplest view, the following examples are phrases of type *noun* or *gerund*. This would enable us to note that, for example, the verb *read* takes an NP complement, as in *John reads long books*. If we assume that there are many different types of phrases in Examples (1) to (3), we would require more complex lexical entries for transitive verbs.

1. a. the very long books (DP or NP)
   b. very long books (DegP or NP)
   c. long books (AP or NP)
   d. books (NP)

2. a. Mary’s eating the sandwich (DP or GerundP)
   b. Mary’s sandwich (DP or NP)
   c. The sandwich (DP or NP)
   d. The big sandwich (DP or NP)
   e. The very big sandwich (DP or NP)
   f. The three biggest sandwiches (DP or NP)

3. a. dirty water (AP or NP)
   b. primeval man (AP or NP)
   c. linguistics articles (AP or NP)

Many P & P linguists, including (Abney, 1987), propose that there is more than one type of head or head-like constituent in each of these phrases. The determiner is the type of head in Examples (1) and (2) that determines the category of the phrase—Abney calls this the *FUNCTIONAL* head. The noun is another type of head, which verbs select for—Abney calls this the *THEMATICAL* head. However, this approach leaves a number of unanswered questions.

If phrases are no longer subcategorized for, then the term *phrase* must drastically change in meaning. Abney’s DP analysis does not explain how the DPs, DegPs, APs, NPs and GerundPs above have the properties attributed to their head categories. For example, which property of the phrase in Example (2c) is determiner-like?

If all NPs are really DPs, the DP analysis provides no explanation of why some nouns need no determiners, as exemplified in Example (3). One can imagine trying to insert “invisible” determiners into every determiner-less DP in order to maintain that all NPs are really DPs. If we assume “invisible” determiners, we must also assume “invisible” degree words and adjectives in order to maintain the DegP and AP hypotheses. Another possibility would be to assume that APs, DegPs and
NPs are allowed to occur wherever DPs occur. However, either of these tacks would make the DP analysis unfalsifiable.

On the basis of phrase structure alone, the NP analysis is clearly preferred over the DP analysis. As discussed below, other considerations discussed in the literature can be explained by positing other types of licensing relations.

3 Licensing Relations

Government and Binding Unification Grammar (GBUG) is a feature structure (FS) based model for representing analyses of Principles and Parameters (P & P) Theories, including the Extended Standard Theory, Government and Binding Theory and the Minimalist Theory. GBUG is based on previous work in various linguistic frameworks outside of P & P especially Arc-Pair Grammar, Relational Grammar, Stratified Feature Grammar, Head-Driven Phrase Structure Grammar, Categorial Grammar and Lexicalized Tag Adjoining Grammar.

Fundamental P & P relations are modeled in GBUG as licensing relations. These relations include constituent relations (adjunct, specifier, complement), predicate relations (theta, modification, quantification), agreement relations (abstract case and agreement), scope relations (the relation between a quantifier/negative and the constituent over which it takes scope), and possibly others. Each licensing relation X(A,B) is represented in an edge-labeled single-rooted directed acyclic graph as a pair of same source arcs, as shown in Figure 1. A, the value of the X-licensor arc, represents the licensor of relation X, and B, the value of the X-licensee arc, represents the licensee of the relation X. For example, if X is a theta relation, then A is the theta assigner and B is the theta recipient. If X is an adjunct relation then A is the head (adjunct licensor) and B is the adjunct (adjunct licensee). Similarly, Figure 1 is a schema for representing the case relation, the complement relation, the specifier relation, the modification relation and other licensing relations. This schema also clearly identifies case assigners/assignees, heads and specifiers/adjuncts/complements, modifiers and modifiees, etc.
Arc labels in GBUG can simultaneously represent more than one licensor and/or licensee feature, as defined in the grammar. These definitions are intended to capture generalizations and remove unnecessary arcs from the GBUG representation. By definition, the arc label Quantifier represents (is subsumed by) the features Specifier Licensee and Quantification Licensor; the arc label Modifier represents the features Adjunct Licensee and Modification Licensor and the arc label Head-Proj represents the features Adjunct Licensor, Specifier Licensor, Modification Licensee, Quantification Licensee, among others. As diagrammed in Figure 2, GBUG representations of the relations Quantification(X,Y) and Specifier(Y, X) can be combined (or unified) as one FS consisting of same-source Quantifier and Head-Proj arcs. Quantifier is the most general label in the grammar that is subsumed by (is more specific than) both Specifier-Licensee and Quantification-Licensor, and therefore represents their unification. Similarly, Head-Proj is the unification of Quantification-Licensee and Specifier-Licensor. Figure 3 includes representations of the licensing relations: Quantification(a, [good book]), Specifier([good book], a), Modification(good, book) and Adjunct(book, good).

4 Constituent Licensors

The CONSTITUENT LICENSOR, as defined in Definition 1, determines the distribution of the parent constituent to a large extent. This definition is based on distinctions made in (Zwicky, 1985), (Zwicky, 1993) and sources cited there. The feature CONSTITUENT LICENSOR is represented as the feature HEAD-PROJ under the assumption that the constituent licensor is the head of the phrase. Following previous work in generative frameworks, I assume three constituent relations: SPECIFIER, ADJUNCT, and COMPLEMENT. Given that no item can simultaneously be the licensee of more than one of these relations, the licensors of these relations do not need to be differentiated–HEAD-PROJ is the licensor feature for all constituent licensing relations.
**Definition 1** The CONSTITUENT LICENSOR of a phrase XP is the constituent that is obligatory, barring ellipsis, and that bears the features F of XP, such that F are subject to semantic selection and subcategorization restrictions imposed by constituents external to XP.

Under the licensing relation approach, the relations ADJUNCT, SPECIFIER and COMPLEMENT serve to group together other licensing relations into classes. For example, the predicate relation INTERNAL-THETA, the case relation COMP-CASE, are both types of complement relations, just as initial and final 2s are both types of direct object relations in Relational Grammar (RG). The specifier relation combines RG’s notion of subject (external-theta, spec-case) with a variety of other relations including quantifier and some special relations representing “intermediate positions”, which are important in constraints on long distance dependencies and quantifier scope. Adjunct relations correspond to both modifier relations and special relations involving “intermediate positions”.

Under many other P & P approaches, different constituent licensees refer to different positions in phrase structure relative to the head. These approaches assume that constituents need not be labeled with respect to their functions as specifier licensees (specifiers), adjunct licensees (adjuncts) or complement licensees (complements) because their constituent licensee function follows from their positions in phrase structure relative to the head. As noted by (Pullum, 1985) and others, most
Table 1: Sample “Structural” Definitions of Constituent Licensees

<table>
<thead>
<tr>
<th>Constituent Licensee</th>
<th>Sister-of</th>
<th>Daughter-of</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specifier</td>
<td>$X$</td>
<td>$X$</td>
</tr>
<tr>
<td>Adjunct</td>
<td>$X$</td>
<td>$X$</td>
</tr>
<tr>
<td>Complement</td>
<td>$X^0$</td>
<td>$X$</td>
</tr>
</tbody>
</table>

P & P analyses fail to state clearly which positions carry which licensee function (i.e., which version of $X$ theory they assume). Table 1 lists one set of possible definitions for identifying which kind of non-head constituent an item is, e.g., a specifier is defined as a nonhead which is the sister of $X$ and the daughter of $X$. However, the terms complement, specifier and adjunct are used much too broadly to be defined in this way. Example (4) contains several fairly standard P & P constituent structure analyses of small clauses, ditransitive complements, NPs, and others. If structural definitions describing Example (4) are added to Table 1, the same structural positions correspond to more than one constituent licensee. The Example (4), and possibly other analyses would have to be eliminated in order to arrive at unambiguous structural definitions based on current $X$ theory. The positions $[sister-of(X), daughter-of(X)]$ can be occupied by ditransitive complements as in (4a) or adjectival modifiers of nouns (adjuncts), as in (4b). The position $[sister-of(X), daughter-of(X)]$ can be occupied by either small clause subjects as in (4c) and (4d) (specifiers), relative clauses as in (4e) (adjuncts) or adverbial modifiers of VP in (4f) (adjuncts). If we eliminate non-branching nodes as proposed in, for example, (Baltin, 1989) and (Chomsky, 1994), then either specifiers or adjuncts may occupy the position $[sister-of(X), daughter-of(X)]$, as in Examples (4b) and (g). Therefore, without making severe changes in P & P analyses of phrase structure, it is not possible to unambiguously define specifier, complement and adjunct in terms of structural positions.

(4)  
   a. $[V [V [V give [N John]] [N a book]]])$
   b. $[N three [N ugly [N green books]]])$
   c. I saw $[A [N John] [A angry]]$
   d. I consider $[N [N Mary] [N a nice person]]$
   e. $[N [N the book] [C that Mary bought]]$
   f. I $[V [V ate my mush] slowly]$
   g. $[N ugly [N green books]]$
5 Predicate Licensors

In GBUG the PREDICATE LICENSOR (PRED) is the constituent that ANCHORS or licenses a nonterminal constituent, as in Lexicalized Tree Adjoining Grammars (Lavelli and Satta, 1991) and Stratified Feature Grammar (SFG) (see (Johnson and Moss, 1993) and (Johnson and Moss, 1994)). Preds are like functors of Categorial Grammar, (Keenan, 1974) and SFG. Under this view a determiner is the predicate licensor of the N (or NP) which it anchors. This N consists of the determiner and its predicate licensee, a constituent of type noun. In Figure 3, the feature Quantifier represents (among other things) the fact that its value, the determiner a, is a pred with a predicate licensee of good book, the value of the Head-Proj feature. Similarly the modifier good is a pred with its modifiee book as the predicate licensee. Theta assigning heads are preds with their theta recipients as predicate licensees. Unlike most previous P & P approaches, the pred need not be the head.

Preds impose selection restrictions and word order restrictions on sister constituents. For example, pregnant requires that its modifiee be capable of being pregnant, as in (5). Example (5a) is semantically anomalous due to real-world assumptions precluding books from being pregnant. Referents of Example (5b) are female, though ophthalmologists need not be. The pred must determine word order in the phrase given that exceptional word order is linked to idiosyncratic preds, including degree words, adjectives, etc., as shown in (6).

(5) a. the pregnant book
   b. pregnant ophthalmologists

(6) a. angry enough
   b. president elect

Quantifiers are harder to pin down than most other preds with respect to their selectional and word order properties. Consider the examples in (7). In (a) and (b), violation of requirements imposed on the quantifiee result in ungrammaticality. In (c) and (d) are semantically anomalous unless we change our assumptions about the meaning of the quantifiee–if we treat if and Adam Meyers as being common noun-like.¹⁰ I will assume that the ill-formedness of (7a) through (7d) are all selection restriction violations, even though (a) and (b) are completely ungrammatical, rather than merely anomalous.

(7) a. *every one dog
   b. *this a dog
   c. ~Scientists study the if
   d. ~I met the Adam Meyers

According to (Greenberg, 1963) (p. 86), demonstratives and numerals need not have the same word orders relative to the noun. Assuming that these are both types of determiners, this suggests that the determiner, not the noun determines word
order. However, variation within the same type of determiner (articles, numerals, etc.) seems rare in the world’s languages. One of Abney’s leading arguments for the DP analysis is based on his analysis of the poss-ing construction, e.g., Example (2a). Abney claims that determiners are heads of DPs, selecting either NP or VP (GerundP) complements.

In the NP analysis, possessive determiners are preds that select constituents of type gerund or type noun to produce the GerundPs and NPs in Example (2). My term *Pred* differs from Abney’s term *functional head* in that preds do not determine the category of the superordinate phrase, e.g., the determiner is a pred anchoring (or justifying) a superordinate NP. The approach presented here avoids the pitfalls of the DP analysis, while maintaining the insight that determiners select their quantifier. Degree words select their degree licensees (and their degree complements) and adjectives select their modifiees.

Some versions of the DP analysis, like Abney’s, propose two or more different types of head. (Radford, 1993) assumes that both types of head the immediate head (functional head) and the ultimate head (thematic head) both determine the category of the nonterminal. (Rothstein, 1991) proposes that there are three types of head, only two of which determine the category of their nonterminal. Rothstein’s Type I heads are, in GBUG terms, both preds and heads, e.g., verbal heads with respect to their theta recipients. Rothstein’s Type III heads are preds that are not heads—they do not determine the category type, e.g., adjective modifiers. Her type II heads are essentially functional heads. Type II heads, according to Rothstein, determine the category of the dominating nonterminal and also bind theta positions in the grids of their complements. Rothstein uses this theta grid property to account for the obligatoriness of determiners. In the next section, the obligatoriness of determiners is accounted for, without assuming they are heads. On this basis, I assume that Type II heads do not exist.

6 Obligatory Determiners with Non-Generic Singular Count Nouns

The head of a phrase is often defined as being the “distributional equivalent” of the whole phrase ((Bloomfield, 1933), p. 233). However, as noted by (Zwicky, 1985), (Hudson, 1987) and others, this is not an easy characteristic to apply to nominal phrases given that determiners are sometimes obligatory. In this section, it is proposed that maximal projections of phrases are defined by constraints rather than particular numbers of bar levels. These constraints may vary by lexical item and/or language. In English, all constituents of type noun are complete NPs if they are subsumed by the FS in Figure 4. Noun heads are “distributionally equivalent” to NPs in the sense that they determine the distribution of the phrase with respect to selection and subcategorization. The reason why some noun heads cannot replace their maximal projections in sentences is that they lack a Quant-Val feature value subsumed by *indef*. This analysis is worked out in greater detail in (Meyers, 1994), (Meyers,
1995a) and has its origins in some unpublished work by David Johnson and myself.

In my analysis, each determiner and adjective anchors a phrase of type noun, as represented in Figures 5 and 6. The phrase anchored by these lexical items structure shares the value of the feature Quant-Val with the head (like head features of Generalized Phrase Structure Grammar). The atomic feature value (node label) indef subsumes the label def. Most pronouns, proper nouns and generic nouns have Quant-Val of Def; plural nouns and mass nouns have Quant-Val of Indef; singular non-generic count nouns have Quant-Val of nil; and determiners anchor noun phrases with various Quant-Val. In combination, Indef and Def unify to Def. Thus, The green books, like The green book is definite. A green book is indefinite. These phrases are subsumed by Figure 4 and are therefore complete noun phrases. Green book has a quant-val of nil, is not subsumed by Figure 4 and is therefore not a complete noun phrase. Determiners are obligatory in phrases headed by nouns with Quant-Val of nil. Without determiners, constituents headed by these nouns would not satisfy the NP completeness constraint in Figure 4.

The fact that determiners are obligatory in noun phrases should not be surprising given the variety of other obligatory complements, specifiers and adjuncts. Adverbial time nouns often require either adjective adjuncts, adverbial adjuncts or
certain determiner specifiers (one, but not the or a) for them to form adverbial time noun phrases, as shown in (8a-c) (cf. (Sager, 1981), p. 87). Subjects are obligatory in finite sentences, as in (8d). Complements of some verbs are obligatory, as in (8e). (Grimshaw, 1990) (pp. 132-133) notes that some passives require either passive by-phrases or some sort of modifier, as in (8f-g). In spite of their obligatoriness, none of the aforementioned items are heads.

(8)  a. [Last night], John went to the movies  
    b. [A night ago], John went to the movies  
    c. [One night], John went to the movies  
    d. [John] laughed  
    e. John built [his home]  
    f. The building was built [by John’s mother]  
    g. The building was built [to please John’s mother]

7 Conclusion

The noun determines the distribution of the noun phrase with respect to subcategorization and selection restrictions. Therefore the noun is the head of the noun phrase. Most previous DP analyses confuse a constituent’s role as predicate licensor, with its role as the head or constituent licensor of the phrase.

The obligatoriness of the determiner with singular nongeneric common nouns has been assumed to imply that the determiner is the head of NP. However, the determiner is only obligatory for English head nouns that are not marked definite or indefinite. The more significant obligatory element is clearly the noun. Thus the
obligatoriness of determiners in certain NPs should be viewed in the same light as the obligatoriness of certain subjects, complements and modifiers.

For the sake of brevity, I have limited the data in this paper to English. There are some interesting cases in other languages in which the determiner plays the role of other types of licensors. For example, numerals are Agreement licensors in Russian NPs. They determine the external agreement properties of the NP and require adjectives and nouns internal to the NP to take particular morphological forms. The reader is directed to (Corbett, 1993), (Payne, 1993), (Meyers, 1994), and the sources cited there for details. However, the point remains the same. The head has the role of determining the category of the superordinate phrase. All accounts of headedness must account for this role regardless of how they represent other licensing relations.

Notes

1See (Meyers, 1994), (Meyers, 1995b), (Meyers, 1995a) for details.

2GBUG FSs are actually Generalized Feature Structures (GFSs). In GFS models, subsumption and unification relations are defined on nonidentical arc and node labels. GFS models are assumed for SFG ((Johnson and Moss, 1993), (Johnson and Moss, 1994) and (Johnson et al., 1993)) and for GBUG. GFS models resemble “typed” FS models, which allow nonidentical node labels (but not nonidentical arc labels) to enter subsumption/unification relations.

3The Quantification relation refers to the predicate relation between the determiner (quantifier) and the adjacent over which the quantifier ranges.

4By convention, each relation R is listed in the text as R(licensor,licensee), the licensor preceding the licensee in the notation.

5The term constituent licensor is approximately the same as the term base, as defined in (Zwicky, 1993).

6(Grimshaw, 1990) discusses A-ADJUNCTS, a class of constituents that includes the passive by-phrase. A-adjuncts have properties of both adjuncts and arguments. In GBUG, the a-adjunct is a SUPPRESSED-THETA licensee, a type of adjunct licensee and theta licensee (not a type of complement).

7Intermediate positions in GBUG are specifier licensee or adjunct licensee arcs. Constraints on these positions are constraints on arcs, rather than constraints on nodes as in most P & P approaches. In some GBUG analyses, intermediate arcs structure-share with other types of arcs. In other GBUG analyses (always in most P & P ap-
proaches) the value of intermediate arcs are intermediate traces. Intermediate po-
ititions (traces) are similar to (intermediate) slash categories in GPSG and HPSG.

8 The constituent that, according to Keenan, “denotes the function” is equivalent
to the functor of Categorial Grammar.

9 These tests for selection restrictions are based on work by Fodor, Lakoff, Mc-
Cawley, Katz, Postal, Weinreich and others in the 1960s and 1970s. See (Meyers,
1994) and (Meyers, 1995a) for further discussion.

10 (7a,b,d) are from (Meyers, 1994). (7c) was originally from (Weinreich, 1966),
p. 463.

11 The Norwegian data in (i) and (ii) suggests that the article determines irreg-
ular word orders (the same phenomena occurs in Swedish, Danish Faroese and Ice-
landic). In most constructions, the definite article is a suffix and the indefinite article
is a separate word that precedes the noun. However, this data is not conclusive ev-
idence because definite determiners can be realized as separate words (den in Nor-
wegian) which precede the noun when the NP contains a restrictive relative clause.
See (Fioretta, 1995) for further details about these constructions.

(i) boken
    bok + en
    ‘the book’

(ii) en bok
    a book
    ‘a book’

12 In the poss-ing construction, the NP within the possessive determiner is an external-
theta licensee selected by the gerund. Thus selection goes in both directions.

References


itors, *Alternative Conceptions of Phrase Structure*. The University of Chicago
Press, Chicago.


