to appear In: Proceedings of WCCFL 8,
Kathavine Hunt 8 Jane Fee, eds.
Stanford Linguistics Association, Stanford. 1989
Running On Empty:

# Null Heads in Head-Driven Grammar

John Nerbonne, Masayo Iida, and William Ladusaw

Hewlett-Packard Laboratories

### 1. Introduction

This paper examines  $\bar{N}$  anaphora in the syntax and lexicon of English.<sup>1</sup>  $\bar{N}$  anaphora includes both *one*-anaphora and null-head anaphora:

(1) a. Scarves were on sale. Al bought a blue one made in China.

b. Al bought several made in China.

(1a) uses the word *one* and is therefore referred to as ONE ANAPHORA. (1b) is missing an expected common noun head (of the object noun phrase in the second sentence), and we therefore refer to it as NULL HEAD ANAPHORA. Our analysis treats these two constructions as syntactically very similar, differing only in whether the anaphoric head constituent is realized as the audible one(s) or a postulated silent CN null (i.e.  $\emptyset$ ).

This is a common and productive construction. It is more common in speech than in prose, and more common in informal styles than in formal ones. But Dahl (1985:129–132) has examples even from formal prose. The productivity of the construction is shown by the fact that  $\bar{N}$  anaphor constructions contain all the variety of structure found in noun phrases.

We also provide a syntax for the closely related partitive construction, which we claim is an instance of a null head construction, though not one which is understood anaphorically.<sup>2</sup>

See Jackendoff (1971, 1977) for discussion of  $\tilde{N}$  gapping. The existence of this structure can be significant in the evaluation of evidence, since the internal syntax of  $\tilde{N}$  gapping differs from that of the  $\tilde{N}$  anaphor construction. But the only examples which are possible in  $\tilde{N}$  Gapping and NOT covered by our analysis of null head anaphora are those where a possessive NP is used as a determiner to an NP WITH posthead modifiers but WITHOUT a CN head. It is possible to add conditions allowing these structures, but the relative infrequency of such examples led us to restrict our attention to the anaphora construction, excluding gapping.

We've had lengthy discussions on the matters presented here with Daniel Flickinger and J. Mark Gawron; for useful comments, criticisms, and other help we would like to thank them, Deborah Dahl, Carl Pollard, Derek Proudian, Geoff Pullum, and Arnold Zwicky.

A second closely related construction is the N gapping construction, which we do not cover entirely. The gapping construction, normally found in coordination and only where antecedents are sentence-internal, allows a wider range of complements and adjuncts in construction with possessives:

<sup>a. Al's relatives from Akron are here.
b. Al's relatives from Akron and
b. Bo's from Biloxi are both here.</sup> 

### (2) Al knows several of the girls.

Our analysis builds on Stockwell, Schachter and Partee (1973:167-190) and on Jackendoff (1977:103-123). We postpone comparisons until we have laid further groundwork.

We give the details of our analysis in HEAD-DRIVEN PHRASE STRUCTURE GRAMMAR, which has been developed by Pollard and Sag (1987). The choice reflects our interest in a monostratal, surfacist approach to syntactic analysis. In brief, our analysis treats all  $\bar{N}$  anaphoric NPs as having the syntax of regular NPs but for the presence of one of two anaphoric CNs: one or null. These anaphoric CNs have specific conditions on which types of complements and modifiers they may take, and on the form of their specifiers.

### 2. Preliminaries

### 2.1. Terminology

Quirk et al. (1985:361) dub INDEPENDENT those possessive pronouns which appear without following head nouns, e.g. mine. We generalize this term to include not only possessive, but also nonpossessive forms, e.g. none. Those determiners and possessive pronouns which require head nouns will analogously be referred to as DEPENDENT. Table (3) presents an illustrative division of determiners, together with the syntactic features our treatment will employ.

1	Q	١
ţ	v	,

Class	Features	Examples
Dependent	[+dependent]	my, her, no, a, · · ·
Independent	[-dependent]	mine, hers, none, one, · · ·
Indiscriminate	[±dependent]	some, which, one, two,

#### 2.2. Three one's

According to all analyses, including ours, the word *one* is seriously ambiguous. There is a generic pronoun sense illustrated in (4) which we ignore completely:

#### (4) One never knows...

There are three additional senses which are quite relevant, however. First, one is a numeral as in (5); second, it is an anaphoric common noun as in (6); and third, it is the independent form of the determiner a as in (7). In this third guise, one bears the same relation to a that none bears to no. We refer to these as the NUMERAL ONE, the COMMON NOUN ONE, and the INDEPENDENT DETERMINER ONE, respectively.

- (5) one or two hours
- (6) a blue one
- (7) Sam needs a jointer, and he hasn't got one.

# 2.3. Analytical Tasks

These challenges face the grammarian in the analysis of  $\bar{N}$  anaphora constructions:

- 1. The surface constituency of null head anaphors (cf. (1b) above);
- 2. The relation between the dependent determiners (no, my) and their independent counterparts (e.g. none, mine);
- 3. The restrictions on the distribution of the common nouns one and null as illustrated in (8):
- (8) a. Many Ø (made in China)

  \*Many ones (made in China)
  - b. \*The Ø (made in China)
     The one (made in China)
  - c. Which Ø (made in China)
     Which one (made in China)
- 4. The relation of  $\bar{N}$  anaphora to the partitive construction, e.g. many of the scarves.

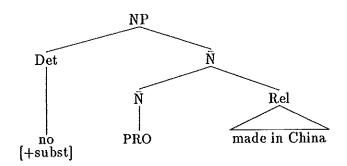
### 3. Jackendoff's Treatment

The  $\bar{N}$  anaphora constructions have been analyzed carefully by Stockwell, Schachter and Partee (1973:167-190) and Jackendoff (1977:103-123), so we begin by sketching those analyses and the points that prompted our modifications. We concentrate on Jackendoff's, since it is the more recent.

Jackendoff proposes a silent pronominal common noun PRO as the head of an NP such as none made in China. The deep structure is the same shape as the one postulated for no objects made in China, and this is the sort of constituent structure we will assign to these phrases (as surface structure). We depart from Jackendoff in the account of how the choice of determiner is affected by the occurrence of the this PRO.

Jackendoff's SUBSTANTIVIZATION TRANSFORMATION (Jackendoff 1977:115) adds '[+substantive]' to a determiner when it is adjacent to PRO; this triggers a rule changing no to none, etc, while having no effect on several. Since the rule is obligatory, noun phrases such as \*no PRO made in China are not generated. Figure (9) shows one of Jackendoff's phrase structures after the substantivization rule has applied.

### (9) Jackendoff's Phrase Structure after Substantivization



The substantivization rule is Jackendoff's account of the alternation between dependent and independent determiners, ensuring that independent determiners always and only appear before null heads, and it seems empirically correct. Nevertheless, this sort of treatment is unavailable in contemporary surface-oriented syntactic theory because it changes the form of a word after it has been inserted into a phrase structure. In particular, the substantivization rule violates Brame's "spelling probition" (Brame 1978:2ff). Our account in section 5 below, replaces Jackendoff's substantivization rule and late spelling rule by a lexical specification on the null CN that its determiners be [+independent].

In both Jackendoff's account and in our own proposal, the occurrence of this PRO is restricted to avoid generating NPs such as \*no blue PRO made in China. Jackendoff appeals to a filtering account which uses the "interpretive rule" (10):

(10) PRO 
$$\rightarrow$$
 UNIT /  $\begin{bmatrix} X \\ +partitive \end{bmatrix}$  —

This rule is proposed to treat PRO in the partitive construction. Roughly, the intent of rule (10) is to provide an interpretation of PRO as UNIT when it is preceded by an independent determiner. Jackendoff (1977:111) is clear about the intent of the rule, i.e. that it apply to items "adjacent to the partitive [independent] word [emphasis added]". Only PROs adjacent to partitive determiners are interpretable by this rule. Sentences which contain uninterpreted PROs are filtered out as ill-formed.

Though Jackendoff does not formulate a distinct N anaphor interpretation rule, an analogue of (10) seems to be intended as an interpretive rule for null head anaphors as well.<sup>4</sup> Trading on the relation between the left environments of the

Jackendoff considers and rejects the objection that the substantivization rule violates the Lexicalist Hypothesis; but he considers a form of Lexicalism weaker than that generally accepted (in the relevant theories) today. He bases his rejection on the grounds that his proposed rule affects a limited class of lexical items and that it has no semantic effects. But these grounds allow even rules such as 'Affix Hopping'.

Rule (10) applies only to partitives, but (10) is the only interpretive rule for PRO in Jackendoff (1971, 1977).

partitive construction and the left environments of the  $\bar{N}$  construction, the interpretation rule would be extended to interpret the null element in the  $\bar{N}$  constructions. That is, the structure can be interpreted *only* where PRO is preceded by an independent determiner; all others are uninterpretable, and therefore ill-formed.

We object to the filtering account for two general reasons: first, our view of semantics provides no analogue to his interpretive rule (10); and second, the filtering account effectively functions as a second grammatical stratum, a complication we eschew for theoretical reasons. In addition, Jackendoff's proposal is empirically deficient. It licenses PRO by rules which are sensitive to the properties of preceding words. But the null CN can be licensed by preceding phrases (comparative and superlative adjective phrases) as well, as (11) shows:<sup>5</sup>

### (11) Scarves were on sale. The most expensive was from China.

In our account, there is no licensing of the null  $\bar{N}$ ; rather, the  $\bar{N}$  selects for an appropriate range of (phrasal or lexical) specifiers and complements.

Stockwell, Schachter, and Partee (1973:175) take a different tack: they explain the  $\bar{N}$  anaphor construction as a pronominalization process which deletes one(s) after certain determiners. They allow that the deletion is obligatory after some determiners, and optional after others, and formulate their treatment not as an interpretive filter, but rather through more standard morphosyntactic rules. We might think of their rule as fixing the interpretation of a feature [+tacit]. In this case, we effectively have a rule of allomorphy accounting for the distribution of one and  $\emptyset$ , but this would be a rule of allomorphy which must be sensitive to adjacent phrases, not merely adjacent words so that the occurrence of the null-head anaphor in a sentence like (11) is correctly predicted. This violates the well-established practice of morphologists codified in Zwicky's 'Trigger Constraint' (Zwicky 1984:389), according to which allomorphic rules are never sensitive to anything more than adjacent words.

Our own account of the distribution of *one* and  $\emptyset$  is presented in Section 6 below; it relies on common nouns governing the feature [ $\pm$ dependent] on their specifiers, and is therefore a purely syntactic account. As such, it covers phrases, as well as words, and it requires no revision in these conceptions of morphological operations. For us, the common noun governs the determiner in the same way that some Latin verbs govern dative, accusative or ablative objects. This is a syntactic dependency, not one of superficial adjacency.

In summary, we extend the analyses of Jackendoff (1977) and Stockwell, Schachter, and Partee (1973) with respect to the tasks (2) and (3) of Section 2.3 above, (2) being that of accounting for the alternation between dependent and independent determiners; and (3) that of describing the distribution of one and  $\emptyset$ . We have moreover restricted the theoretical mise en scene to surface-oriented grammatical apparatus and standard morphology. Finally, we have computationally implemented both the grammar (lexical and syntactic analyses), and the interpretation (semantics and pragmatics).

It would go beyond the bounds of this paper to demonstrate in detail that the superlative form most+ADJ is phrasal, not lexical, so that e.g. most expensive is a phrase, not simply a word customarily written with a space. But briefly, the argument would point out (i) that most may take scope over clearly phrasal material, e.g. conjunctions (most interesting and beautiful); and (ii) that parentheticals can interrupt the putative prefix-root combination in most constructions, but not in any other affix-root or root-affix combinations.

## 4. Surface Constituency

We propose that null-headed constructions contain a phonologically null common noun (following Jackendoff), thus assigning to (1b) a syntactic structure exactly analogous to that in (1a). We present four arguments in favor of this analysis, all essentially simplicity arguments. First, the analysis explains the category (NP) of the phrase several made in China as an  $\bar{\mathbf{X}}$  projection. This is explained in our analysis, where the null head is a common noun.

Second, it explains why we find essentially the same range of posthead  $\bar{N}$  complements and adjuncts in this construction that we find in normal  $\bar{N}$  constructions. Table (12) contains a suggestive list of adjunct types found in both standard noun phrases and in null-headed phrases. The list can be expanded, using examples with complements such as Many proposals to dismiss Jones were defeated. Several to retain him were tabled or The people in this country feel a great interest in economy, but not much in politics. We explain this fact in our analysis since null common nouns select for the same range of posthead complements and adjuncts as their more pronounceable counterparts.

(12)

Type	Examples
Prepositional Phrase	many (people) from Boston
Adverbial	many (people) there
Relative Clause	many (people) who ski
Thatless Relative	many (people) Tom knows
Reduced Relative	many (people) seen there

We note here that we do not find the same range of complements and adjuncts in null-headed NP's with possessive determiners, a fact illustrated in example (a) in footnote 2 above and by (13).

(13) a. my relatives from Boston

his friends who ski

b. \*mine from Boston

\*his who ski

This is an exceptional fact in our analysis, but one which is easily accommodated by the postulation of a distinct null common noun which selects only for a possessive specifier.

Third, we predict the possibility of extraposition from NP, noted by Selkirk (1977) and Jackendoff (1977:107), since this regularly applies to  $\bar{N}$  complements:

#### (14) Several attended who Tom had never seen before.

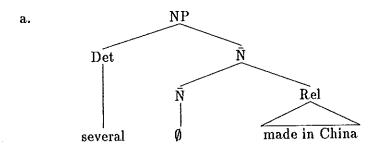
Fourth and finally, the treatment is conservative in its postulation of new lexical classes, syntactic constructions, and grammar principles. This is best appreciated by comparing the structure defended here with its likely alternatives, sketched in Figure (15).

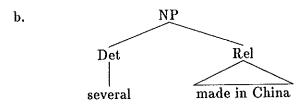
Our own proposed constituent structure analysis of null-headed  $\bar{N}$  anaphors, following Jackendoff, is shown as the first tree diagram. It assembles null-headed noun phrases using the same lexical classes, subcategorizational and selectional restrictions, grammar rules, and and grammar principles that are used for headed noun phrases.

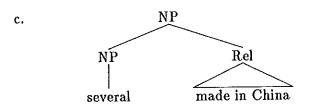
The attempt to analyze null-headed  $\bar{N}$  anaphors without postulating a null common noun leads to examination of structures such as (15b) or (15c). (15b) contains

no proper head for the noun phrase, so that its major category (noun) is essentially anomalous. In both (15b) and (15c) we have no account of why determiners (or saturated NP's) should select essentially the same complements and adjuncts found in standard noun phrases, and this would seem to require novel selectional specifications. Furthermore, in both (15b) and (15c) the possibility of extraposition is entirely unexpected-possibly requiring a new rule just for these structures. In general NP dependents such as nonrestricted relative clauses or certain adverbs (e.g here in Jones here) are not extraposable, and we are not aware of any dependent for the determiner phrase which can be extraposed.

### (15) Candidate Constituent Structures for Null Head



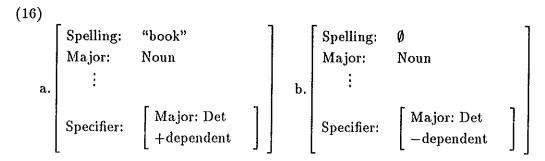




We have assumed the N analysis of NP's throughout this section, in particular in the analysis of relative clauses. Examples such as All of the students and most of the faculty who attended the meeting provide reason to doubt whether all relatives attach to N, and reason to suspect that some relatives attach to NP. In general, these considerations make structure (15c) a more plausible candidate than it might first seem to be.

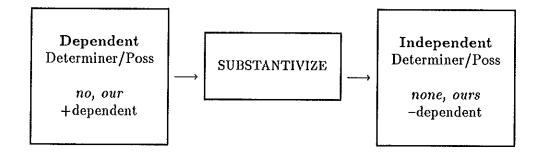
## 5. Dependent and Independent Determiners

The syntactic distinction between dependent and independent determiners is simply the value of the feature [ $\pm$ dependent], as presented in Table (3) above. Pronounceable common nouns require the value [ $\pm$ dependent] on their specifiers, while the null common noun requires [ $\pm$ dependent] as illustrated in (16). Note that we assume that the  $\bar{N}$  s select their modifiers, which will be discussed in more detail shortly.



In order to eliminate redundancy in the lexicon, we also postulate a lexical rule which maps dependent specifiers into their independent counterparts. This applies to both on determiners (no, none) and possessive pronouns (our, ours); and is sketched in Figure 17.<sup>7</sup>

### (17) The Substantivizing Lexical Rule



# 6. The Distribution of one and $\emptyset$

# 6.1. Variation between Determiners

The common noun *one* and null are interchangable in some environments and but not in others. Here are characteristic patterns:

The dependent possessive pronouns are identical in form to genitive elements in the paradigm of personal pronouns (e.g. a friend of mine (= my friend)). If these were not just homonyms, but the same lexical item, then rule (16) would have two distinct clauses, one mapping possessive pronouns to their genitive counterpart, and one mapping dependent to independent feature.

(18)	Independent	Dependent	Indiscriminate
	mine Ø	*my Ø	his Ø
	*mine one(s)	*my one(s)	*his one(s)
	*mine tall one(s)	my tall one(s)	his tall one(s)
	*mine tall Ø	*mv tall Ø	*his tall Ø

First, we recall that we must analyze this as a syntactic, not an allomorphic alternation, since the (phrasal) superlative licenses the null head as in (11) above.

Second, note that since this is a generalization concerning the determiners and adjectives that a common noun may appear in construction with, it is essentially a selectional restriction. Two accounts of selection in this domain are current: first, determiners may be analyzed as selecting for  $\bar{N}$  phrases, and second, common nouns may be the locus of selection, as Pollard and Sag (1987: 139ff) propose. We consider these possibilities in turn.

Considering determiners as selecting for  $\bar{N}$  phrases, we focus on a single alternation, the very general pattern in (19):

If determiners are to select  $\bar{N}$  phrases, then the  $\bar{N}$  tall one must be syntactically distinct from the  $\bar{N}$  one, and since  $\bar{N}$  is certainly the head of Adj +  $\bar{N}$ , the distinguishing feature cannot be a head feature. Since no unbounded dependency is involved, this is also a poor candidate for a foot (binding) feature. In fact, there is no independently motivated feature type that might bear the putative distinction. We are therefore skeptical about this treatment.

We turn then to the second option, in which we analyze the head noun as selecting for both adjectives and determiners. In this case, null must select only independent determiners, disallowing both adjectives and dependent determiners. The common noun *one*, on the other hand, must select for an obligatory adjective and a [+dependent] determiner.

The lexical entry for  $\emptyset$  in (16b) can be contrasted with the entry for *one* in (20):

(20)

## 6.2. Exceptional Determiners

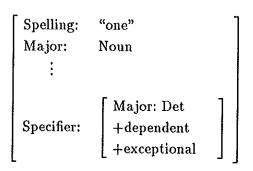
This basic picture is complicated in an interesting fashion by a small class of exceptional elements which are consistent with both null- and *one*-headed Ns, illustrated in (21):

#### (21) Exceptional

which Ø
which ones
which tall ones
\*which tall Ø

We find the same pattern after the determiners each, this, that, these, those, either, neither, and after comparative and superlative adjective phrases, e.g. most interesting. The unexpected behavior is the well-formedness of examples such as which one, where no adjective precedes the one. An obviously available and mechanically correct treatment for these phenomena would involve the postulation of a homonymous one which selects only this class of determiner (we'll call it [+exceptional]) and does not obligatorily require adjective phrases. The lexical entry for the one in combination with exceptional determiners is given below.

(22)



[+exceptional] determiners can combine with either one, thus: which blue one as well as which one. Dependent determiners with no [+exceptional] specification, on the other hand, combine only with one together with adjectives; hence several blue ones, but not \*several ones.

A more interesting tack is to collapse the two lexical entries above into a single lexical item *one*. In this case *one* must select conditionally for an adjective whenever a nonexceptional determiner is chosen.

Spelling: "one"

Major: Noun

:

Specifier: 
$$\begin{bmatrix} Major: Det \\ +dependent \\ \alpha & expectional \end{bmatrix}$$

Adjuncts:  $\{\begin{bmatrix} Major: Adj \\ \beta & obligatory \end{bmatrix} \dots \}$ 

It is not our purpose to put forward a detailed proposal about such mechanisms, but only to note that some such mechanism is clearly of interest. For a further example, we noted in Section 4 above that possessive N anaphors select for no adjuncts, in marked contrast to those with determiners. This fact, too, could be expressed as a constraint on a more complicated version of (23).

Conditional selection appears to be a novel mechanism (though expressible in the Local Constraint Language of JPSG (Hasida 1986:85-87) or Kasper's Functional Unification Grammar (Kasper 1988: 237ff)), but we suggest that it may find application in the specification of optional selection, which is often accompanied by minor semantic changes (cf. transitive vs. ditransitive write, where the understood object of the transitive is letters).

### 6.3. Other Environments

There are other environments in which null common nouns may appear; we have not attempted to catalogue these, and we must resist the temptation now for lack of space. We include the specification below for a null that combines comparative and superlative adjective phrases in order to substantiate our claim to have correctly analyzed the necessarily syntactic conditioning in null head anaphora. We would prefer to collapse this null common noun specification with that in (15) above using the mechanisms outlined in (23), but we suppress this complication here.

<sup>&</sup>lt;sup>8</sup> A detailed proposal ought to answer questions such as: what is the range of logical operators allowed (boolean or quantificational)? Which semantics for boolean operations is assumed (classical, intuitionist, etc.)? Does an entry such as (23) correspond to one word or two?

We suggest that the representation above is useful because it encodes the information in a single word. It is worth noting that, in some feature formalisms, having a single lexical representation such as (23) is provably equivalent to having two, each with the more specific information. But here we urge caution: the range of syntactic structures compared may be the same, but the overall linguistic systems could nonetheless be distinct. There are at least two important nonsyntactic roles which word definitions play: (i) words are the units on which lexical rules operate; and (ii) anaphoric relationships are sensitive to word (sense) identity, and thus word identity.

Spelling: Ø
Major: Noun
:

Adjuncts: {

Major: Adj
Degree: Super V Compar
+obligatory

...}

### 7. The Partitive

The partitive construction uses the same null head as the  $\bar{N}$  anaphora construction. It differs only in NOT requiring an antecedent for proper interpretation:

(25) Al knew several of the girls.

We add just two notes about the syntax of partitives. First, there is a constraint that the NP in an of NP construction be definite; Jackendoff (1977:117) dubs this the PARTITIVE CONSTRAINT. Thus we find:

- (26) a. Al knew several of the girls.
  - b. \*Al knew several of few girls.

Ladusaw (1982:233-35) sketches a satisfactory semantic account of this fact, in which no syntactic mechanisms are required.

There is a second essentially semantic constraint which concerns a sort of agreement between the independent determiner and the head of the partitive NP:

- (27) a. much of the water.
  - b. \*much of the consultants.
  - c. \*many of the water.
  - d. many of the consultants.

But this constraint is likewise semantic—there are singular partitive NP's for which the independent determiner many is appropriate, e.g. many of the group, and there are so-called "universal grinder" construals of non-mass terms with mass specifiers, e.g. much of the house.

# 8. Parsing

An important potential drawback of this proposal is that, in supposing that a lexical entry might be realized as the empty string, it runs the risk of asking a parser to look at so much structure that the parser runs significantly more slowly. It turns out, however, that one can minimize these effects by postulating the possible presence of a null element in as few places as possible.

In our implementation of this analysis, we postulate the possible presence of null after the closed class of independent determiners. For more complete coverage we would have to include adjectives among the elements which license this postulation. Another tack that we considered involved the modification of a parser so that it would postulate nulls as parsimoniously as possible.

### 9. Conclusions

In summary, we build on analyses by both Stockwell, Schachter and Partee and Jackendoff. We extend these in providing implementation, semantics and pragmatics; we follow Jackendoff in assigning to null-headed anaphors a surface constituency Stockwell et al. would have regarded only as proper deep structure. We renovate these analyses in identifying the conditioning of the distribution of one and  $\emptyset$  syntactically, not morpholexically or through string adjacency conditions; and in providing a lexical account of the relation between dependent and independent determiners.

### References

- Brame, Michael. 1978. The base hypothesis and the spelling prohibition. Linguistic Analysis 4(1), 1-30.
- Dahl, Deborah A. 1985. The Structure and Function of One-Anaphora in English. Ph.D thesis, University of Minnesota.
- Hasida, Koiti. 1986. Conditional unification for natural language processing. In Proceedings of COLING '86, 85-87.
- Jackendoff, Ray. 1971. Gapping and related rules. Linguistic Inquiry 2(1), 21-36.
- Jackendoff, Ray. 1977.  $\bar{X}$  Syntax: A Study of Phrase Structure. MIT Press, Cambridge.
- Kasper, Robert T. 1988. Conditional descriptions in functional unification grammar. In Proceedings of the 26th Annual Meeting of the Association for Computational Linguistics, Stanford, 233-240.
- Ladusaw, William A. 1982. Semantic constraints on the English partitive construction. In Proceedings of the First West Coast Conference on Formal Linguistics, 231-242.
- Pollard, Carl and Ivan Sag. 1987. An Information-Based Theory of Syntax and Semantics. University of Chicago Press, Chicago.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, and Jan Svartik. 1985. A Comprehensive Grammar of English Language. Longman, London.
- Stockwell, Robert P., Paul Schachter, and Barbara H. Partee. 1973. The Major Syntactic Structures of English. Holt, Rhinehart and Winston, New York.
- Zwicky, Arnold M. 1984. Welsh soft mutation and the case of object NPs. In Joseph Drogo, Veena Mishva, and David Testen (eds.), Proceedings of the 20th Meeting of the Chicago Linguistic Society, Chicago, 387-402.

,
(
ţ