Grammar Formalisms

Course Notes 1

Ash Asudeh
University of Canterbury
ash.asudeh@canterbury.ac.nz

December 6–7, 2004

1 Theoretical Linguistics

- *Linguistics* := The study of human language
- *Linguistic theory* :=
  1. A scientific theory of human languages (Bloomfield)
  2. A scientific theory of knowledge of language (Chomsky)
- Pollard and Sag on mathematical theories:

  > In any mathematical theory about an empirical domain, the phenomena of interest are *modelled* by mathematical structures, certain aspects of which are conventionally understood as corresponding to observables of the domain. The theory itself does not talk directly about the empirical phenomena; instead, it talks about, or is *interpreted by*, the modelling structures. Thus the predictive power of the theory arises from the conventional correspondence between the model and the emprirical domain. (Pollard and Sag 1994:6)

- *Grammar formalism* := a mathematically precise notation for formalizing a theory of grammar
- Chomsky on the need to formalize linguistic theory:

  > Precisely constructed models for linguistic structure can play an important role, both negative and positive, in the process of discovery itself. By pushing a precise but inadequate formuolation to an unacceptable conlusion, we can often expose the exact source of this inadequacy and, consequently, gain a deeper understanding of the linguistic data. More positively, a formalized theory may automatically provide solutions for many problems other than those for which it was explicitly designed. Obscure and intuition-bound notions can neither lead to absurd conclusions nor provide new and correct ones, and hence they fail to be useful in two important respects. I think that some of those linguists who have questioned the value of precise and technical development of linguistic theory have failed to recognize the productive potential in the method of rigorously stating a proposed theory and applying it strictly to linguistic material with no attempt to avaoid unacceptable conclusions by *ad hoc* adjustments or loose formulation. (Chomsky 1957:5)

- In grammar engineering and other computational applications of linguistics, we have no choice: a formalism is required for the simple reason that we need a language for encoding grammars that is a) understandable by humans and b) understandable by computers:
  - Grammar development in machine language is not feasible for humans.
  - Computers can't understand natural language statements of grammars, no matter how perspicuous to humans.
- **Grammaticality**:

(1) a. Colorless green ideas sleep furiously. (Chomsky 1957)
    b. *Furiously sleep ideas green colorless.

(2) a. Revolutionary new idea appear infrequently. (Sells 1985)
    b. *Infrequently appear ideas new revolutionary.

- *Syntax* := The sub-field in linguistics concerned with how words combine into larger units — phrases — and how phrases in turn combine into yet larger units:
  - word order

    (3) a. Her tall brother is handsome.
        b. *Tall her is handsome brother.

    (4) a. Aren't I allowed to attend?
        b. *I aren't allowed to attend.

  - agreement

    (5) Every girl / *girls is / *are female / *females.

    (6) All *girl / girls *is / are female / females.

  - heads and complementation

    (7) a. Thora **fears** the vacuum cleaner.
        b. Thora **fears** that the vacuum cleaner will get her.
        c. *Thora **fears** of the vacuum cleaner.

    (8) a. *Thora is **afraid** the vacuum cleaner.
        b. Thora is **afraid** that the vacuum cleaner will get her.
        c. Thora is **afraid** of the vacuum cleaner.

    (9) a. The vacuum cleaner **frightens** Thora.
        b. *The vacuum cleaner **frightens** Thora that it will get her.
        c. *The vacuum cleaner **frightens** of Thora.

- ◦ bounded dependencies

(10) a. **Mary** said Ken pinched **her**.
b. ***Mary** said Ken pinched **herself**.
c. Mary said **Ken** pinched **himself**.

- ◦ (11) a. David said **Ken and Mary** tried to **arrive together**.
b. *Ken and Mary said **David** tried to **arrive together**.

- ◦ unbounded dependencies

(12) What did Kim claim that Sandy suspected that Robin stole?

- *Semantics* := The sub-field in linguistics concerned with meaning.
  - ◦ Truth conditions

    (13) "Snow is white" is true if and only if snow is white.
  - ◦ Entailment

    (14) Every student wants good grades ⇒ That student wants good grades
  - ◦ Presupposition

    (15) Jones has stopped smoking *presupposes that* Jones smoked

1.1 Modern Perspectives

- *Generative grammar* := A system of rules that defines in a formally precise way (i.e. 'generates' [in the mathematical sense – AA]) a set of [structures] that represent the well-formed sentences of a given language. (Sag et al. 2003:525)
  - ◦ This only covers syntax, but generative grammar extends to other sub-systems of linguistic information (semantics, phonology, morphology) under the requirement that explicit systems of rules for those sub-systems are stated.
- Two conceptions of language (Soames 1984):
  - ◦ Languages are abstract mathematical systems in the world. (mathematical/Platonic conception)
  - ◦ Languages are internalized systems of knowledge in the minds of speakers. (psychological/cognitive conception)
- The currently dominant view in linguistics and cognitive science is the second, cognitive/"Knowledge of Language" conception (discussed cogently in Chomsky 1986:1–50).
  - ◦ Three basic questions for linguistics (Chomsky 1986:3):
    1. What constitutes knowledge of language?
    2. How is knowledge of language acquired?
    3. How is knowledge of language put to use?
  - ◦ Chomsky's goals and assumptions (Green and Morgan 2001:2–6):
    1. The mind is innately structured.
    2. The mind is modular.
    3. There is a distinct module for language.
    4. Language acquisition is the central puzzle for linguistic theory.
    5. Syntax is formal.
    6. Knowledge of language is itself modular.

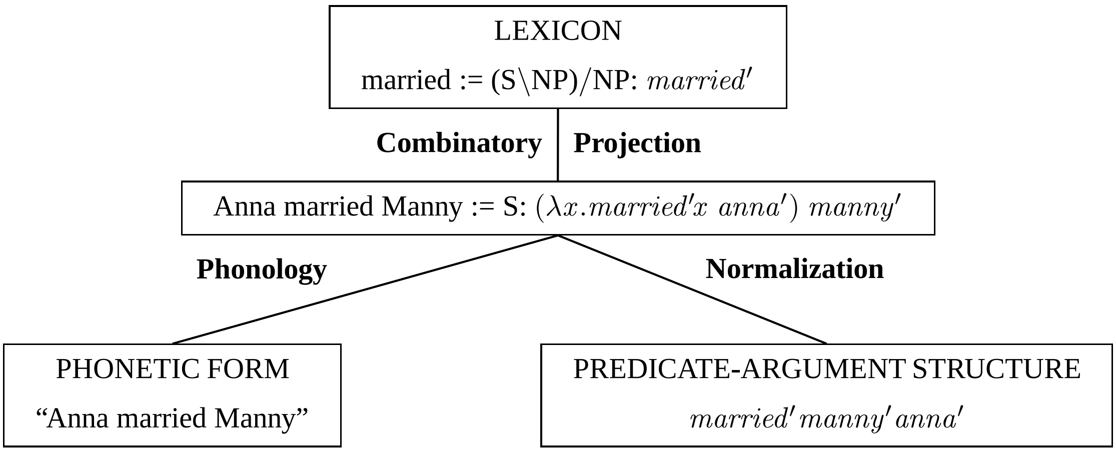
1.2 Goals of the Field

- Acquisition and learnability
- Typology and universals
- Explanatory wide coverage
- Human language processing
- Computation linguistics
  - Mathematical linguistics
  - Parsing
  - Generation
  - Statistical NLP

2 Grammatical Architectures

2.1 Categorial Grammar

- Form-meaning pairing (PHONETIC FORM–PREDICATE-ARGUMENT STRUCTURE), mediated by syntactic categories, projected from lexicon



2.2 Head-Driven Phrase Structure Grammar

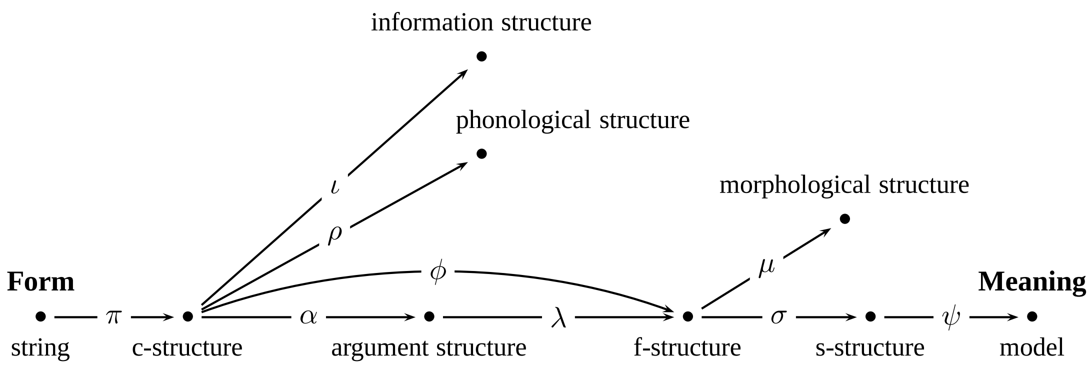
- Sign-based architecture (term *sign* used in the sense of Saussure 1919/1959)

(16)
$$\begin{bmatrix} \text{PHONOLOGY} & \langle \ldots \rangle \\ \text{SYNSEM} & \begin{bmatrix} \text{CATEGORY} & \textit{category} \\ \text{CONTENT} & \textit{content} \\ \text{CONTEXT} & \textit{context} \end{bmatrix} \end{bmatrix}_{\textit{sign}}$$

- Directed acyclic graphs *represent* sorted feature structures, the objects in the theory that model linguistic phenomena (signs)
- Attribute-value matrices *describe* feature structures

2.3 Lexical Functional Grammar

- Parallel projection architecture (Kaplan 1987, Halvorsen and Kaplan 1988, Kaplan 1989)
  - ○ Separate levels of representation (projections) model different aspects of linguistic information
    - · Constituent structure (c-structure): precedence, dominance, constituency
    - · Functional structure (f-structure): grammatical functions, predication, subcategorization, bounded and unbounded dependencies
  - ○ Each projection modelled by logics and data structures appropriate for capturing the information it models
    - · C-structure: trees, described by phrase structure rules
    - · F-structure: tabular functions (represented as attribute value matrices), described by regular expressions
  - ○ Projection functions map from projection to successive projection



References

Chomsky, Noam. 1957. *Syntactic Structures*. The Hague: Mouton.

—. 1986. *Knowledge of Language: Its nature, origin, and use*. New York: Praeger.

Dalrymple, Mary, Ronald M. Kaplan, John T. Maxwell, and Annie Zaenen, eds. 1995. *Formal issues in Lexical-Functional Grammar*. Stanford, CA: CSLI Publications.

Green, Geogia M., and Jerry L. Morgan. 2001. *A Practical Guide to Syntactic Analysis*. Stanford, CA: CSLI Publications, 2nd edn.

Halvorsen, Per-Kristian, and Ronald M. Kaplan. 1988. Projections and Semantic Description in Lexical-Functional Grammar. In *Proceedings of the International Conference on Fifth Generation Computer Systems*, 1116–1122. Institute for New Generation Systems, Tokyo.

Kaplan, Ronald M. 1987. Three Seductions of Computational Psycholinguistics. In P. Whitelock, M. M. Wood, H. L. Somers, R. Johnson, and P. Bennett, eds., *Linguistic Theory and Computer Applications*, 149–181. London: Academic Press. Reprinted in Dalrymple et al. (1995:339–367).

—. 1989. The Formal Architecture of Lexical-Functional Grammar. In Chu-Ren Huang and Keh-Jiann Chen, eds., *Proceedings of ROCLING II*, 3–18. Reprinted in Dalrymple et al. (1995:7–27).

Pollard, Carl, and Ivan A. Sag. 1994. *Head-driven Phrase Structure Grammar*. Chicago, IL and Stanford, CA: The University of Chicago Press and CSLI Publications.

Sag, Ivan A., Thomas Wasow, and Emily M. Bender. 2003. *Syntactic Theory: A Formal Introduction*. Stanford, CA: CSLI Publications, 2nd edn.

Saussure, Ferdinand de. 1919/1959. *Course in General Linguistics*. New York: McGraw-Hill.

Sells, Peter. 1985. *Lectures on Contemporary Syntactic Theory*. Stanford, CA: CSLI Publications.

Soames, Scott. 1984. Linguistics and Psychology. *Linguistics and Philosophy* 7(2): 155–179.

Grammar Formalisms

ALTSS 2004
Course Notes 2

Ash Asudeh
University of Canterbury
ash.asudeh@canterbury.ac.nz

December 6–7, 2004

1 Syntactic Categories and Basic Combinatorics

- Common syntactic categories:

S sentence, clause

(1) [[That the very bouncy ball will bounce on any surface] surprised her].

N noun

(2) That the very bouncy [ball] will bounce on any [surface] surprised [her].

NP noun phrase

(3) That [the very bouncy ball] will bounce on [any surface] surprised [her].

V verb

(4) That the very bouncy ball will [bounce] on any surface [surprised] her.

VP verb phrase

(5) That the very bouncy ball will [bounce on any surface] [surprised her].

P preposition

(6) That the very bouncy ball will bounce [on] any surface surprised her.

PP prepositional phrase

(7) That the very bouncy ball will bounce [on any surface] surprised her.

A adjective

(8) That the very [bouncy] ball will bounce on any surface surprised her.

AP adjective phrase

(9) That the [very bouncy] ball will bounce on any surface surprised her.

- Functional categories:

I inflection

(10) That the very bouncy ball [will] bounce on any surface surprised her.

IP inflectional phrase

(That the very bouncy ball [will bounce on any surface] [surprised her)].

C complementizer

(11) [That] the very bouncy ball will bounce on any surface surprised her.

CP complementizer phrase

(12) [That the very bouncy ball will bounce on any surface] surprised her.

D determiner

(13) That [the] very bouncy ball will bounce on [any] surface surprised [her].

DP determiner phrase

(14) That [the very bouncy ball] will bounce on [any surface] surprised [her].

1.1 Categorial Grammar

1.1.1 Categories

- Set of basic categories; typically:

  S clause

  N common noun

  NP noun phrase, proper noun, pronoun

- All other categories are compositionally made up out of basic categories:

  (15) Intransitive verb: S|N

  (16) Transitive verb: (S|N)|N

  (17) Determiner phrase: NP|N

  (18) Adjective: N|N

1.1.2 Combinatorics

- (Functional) application:
  - Non-directional (result|argument):
    A|B, B $\rightarrow$ A
  - Directional
    - Forward application (result/argument):
      A/B B $\rightarrow$ A
    - Backward application, “Lambek style”, “result on top” (argument\result):
      B B\A $\rightarrow$ A
    - Backward application, “Steedman style”, “leading edge” (result\argument):
      B A\B $\rightarrow$ A

1.1.3 Example

(19)

$$\dfrac{\dfrac{\text{Kim}}{\text{NP}} \quad \dfrac{\dfrac{\text{ate}}{(\text{S}\backslash\text{NP})/\text{NP}} \quad \dfrac{\dfrac{\text{the}}{\text{NP}/\text{N}} \quad \dfrac{\text{banana}}{\text{N}}}{\text{NP}}>}{\text{S}\backslash\text{NP}}>}{\text{S}}<$$

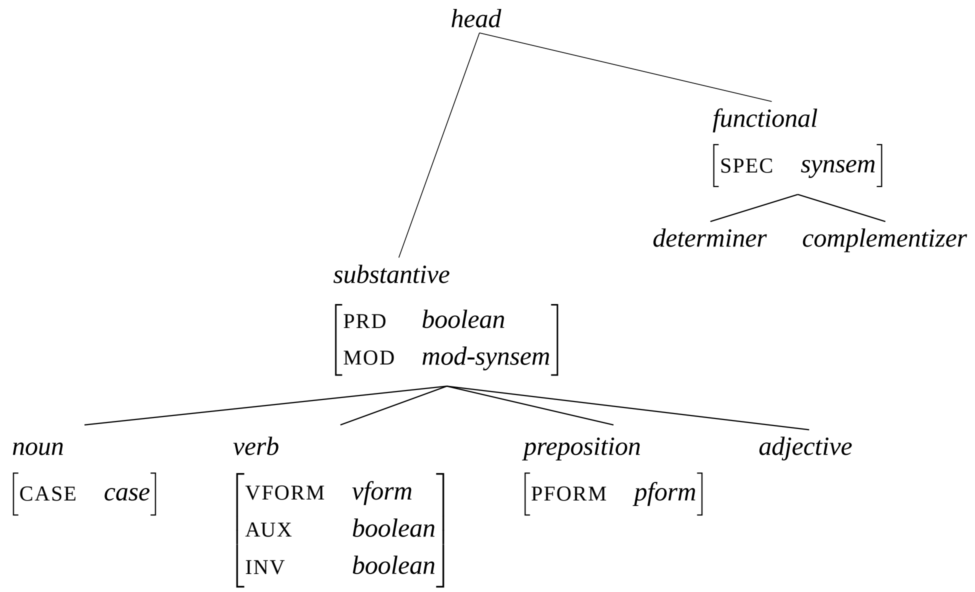
1.2 Head-Driven Phrase Structure Grammar

1.2.1 Categories

- Basic type *category* introduces the following **feature declaration** (Pollard and Sag 1994:398):

$$\textit{category}: \begin{bmatrix} \textsc{head} & \textit{head} \\ \textsc{subcat} & \textit{list(synsem)} \end{bmatrix}$$

- Subtypes of *head* add further feature declarations as appropriate (Pollard and Sag 1994:396–398):



1.2.2 Combinatorics

- Combinations of syntactic units into larger phrases is governed by a number of independent **principles** and **Immediate Dominance Schemas/Rules**.

- Key principles:

(20) Head Feature Principle (HFP)
In any headed phrase, the HEAD value of the mother and the HEAD value of the head daughter must be identical.

(21) Subcategorization Principle (Pollard and Sag 1994)
In any headed phrase, the list value of DAUGHTERS | HEAD-DAUGHTER ... SUBCAT is the concatenation of the list value of the phrase's SUBCAT with the list consisting of the SYNSEM values (in order) of the elements of the list value of DAUGHTERS | COMPLEMENT-DAUGHTERS.

$$\left[\begin{array}{ll} \text{SYNSEM ... SUBCAT} & \boxed{1} \\ \text{DAUGHTERS} & \left[\begin{array}{ll} \text{HEAD-DTR ... SUBCAT} & \boxed{1} \oplus \boxed{2} \\ \text{COMP-DTRS ... SUBCAT} & \boxed{2} \end{array}\right]_{\textit{head-struc}} \end{array}\right]_{\textit{phrase}}$$

(22) Valence Principle (Sag et al. 2003:106)
Unless the rule says otherwise, the mother's values for the features SPR and COMPS are identical to those of the head daughter.

(23) Immediate Dominance Principle
Every headed phrase must satisfy exactly one of the ID schemata.

- ID schemas/rules for **subjects/specifiers**[1] and **complements**

  - Note: More recent versions of HPSG tend to break SUBCAT up into two lists: SUBJ or SPR and COMPS, where the old SUBCAT is equivalent to SUBJ ⊕ COMPS.

(24) Head-subject schema (Pollard and Sag 1994:402)
The SUBCAT value is the empty list ⟨ ⟩, and the DAUGHTERS value is an object of sort *head-comp-struc* whose HEAD-DAUGHTER value is a phrase whose COMPLEMENT-DAUGHTERS value is a list of length one.

(25) Head-specifier rule (adapted from Sag et al. 2003:501)

$$\left[\begin{array}{ll} \textit{phrase} & \\ \text{SPR} & \langle\ \rangle \end{array}\right] \rightarrow \boxed{1} \quad \mathbf{H}\left[\begin{array}{ll} \text{SPR} & \langle \boxed{1} \rangle \\ \text{COMPS} & \langle\ \rangle \end{array}\right]$$

(26) Head-complement schema (Pollard and Sag 1994:402)
The SUBCAT value is a list of length one, and the DAUGHTERS value is an object of sort *head-comp-struc* whose HEAD-DAUGHTER value is a word.

(27) Head-complement rule (adapted from Sag et al. 2003:502)

$$\left[\begin{array}{ll} \textit{phrase} & \\ \text{COMPS} & \langle\ \rangle \end{array}\right] \rightarrow \mathbf{H}\left[\begin{array}{ll} \textit{word} & \\ \text{COMPS} & \langle \boxed{1} \ldots \boxed{n} \rangle \end{array}\right] \quad \boxed{1} \ldots \boxed{n}$$

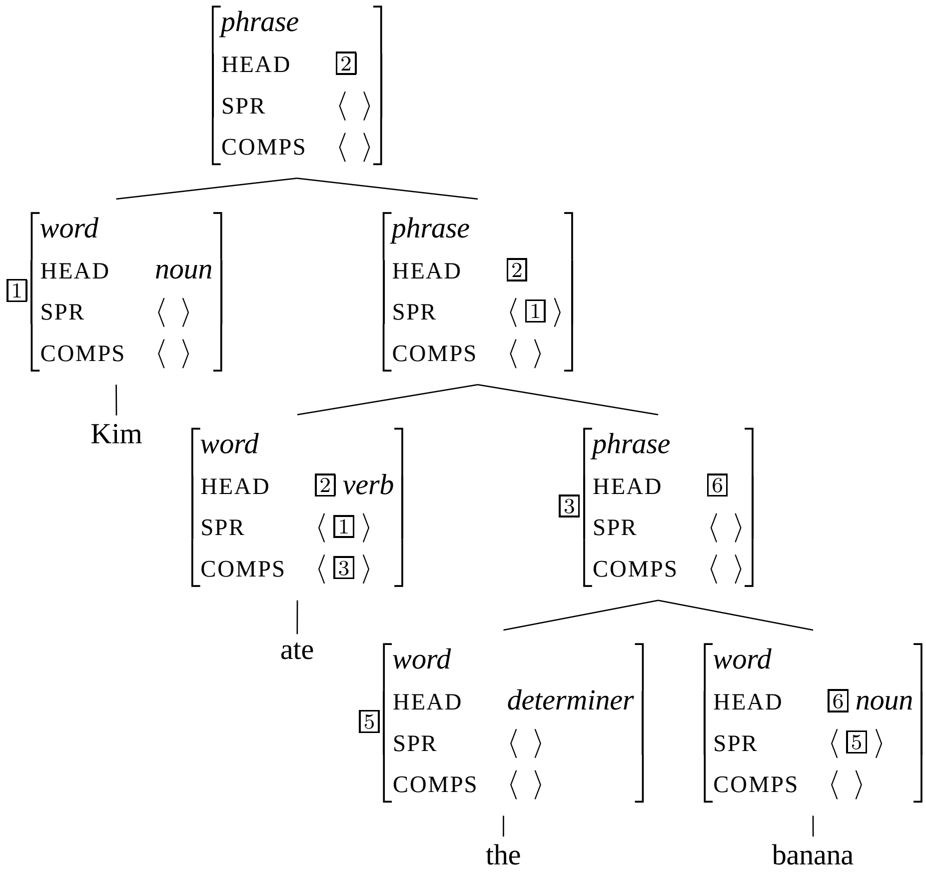
[1] In some versions of HPSG and related formalisms, there is no longer a strong distinction made between subjects and nominal specifiers like determiners or the possessive in *John's destruction of the cake alarmed me.*

1.2.3 Example

(28) Version/style: Pollard and Sag (1994)

```
[phrase
 PHONOLOGY  ⟨ kim, ate, the, banana ⟩
 SYNSEM | LOCAL | CATEGORY [HEAD    [2]
                            SUBCAT  ⟨ ⟩ ]
 DTRS [head-comp-struc
       HEAD-DTR [phrase
                 PHON  ⟨ ate, the, banana ⟩
                 LOC | CAT [HEAD    [2]
                            SUBCAT  ⟨ [1] ⟩ ]
                 DTRS [head-comp-struc
                       HEAD-DTR [word
                                 PHON      ⟨ ate ⟩
                                 CATEGORY  [HEAD    [2]
                                            SUBCAT  ⟨ [1], [3] ⟩ ] ]
                       COMP-DTRS ⟨ [phrase
                                    PHON    ⟨ the, banana ⟩
                                    SYNSEM  [3] [LOC | CAT [HEAD    [6]
                                                            SUBCAT  ⟨ ⟩ ] ]
                                    DTRS [head-comp-struc
                                          HEAD-DTR [word
                                                    PHON    ⟨ banana ⟩
                                                    SYNSEM  [LOC | CAT [HEAD    [6] noun
                                                                        SUBCAT  ⟨ [5] ⟩ ] ] ]
                                          COMP-DTRS ⟨ [word
                                                       PHON    ⟨ the ⟩
                                                       SYNSEM  [5] [LOC | CAT [HEAD    det
                                                                               SUBCAT  ⟨ ⟩ ] ] ] ⟩ ] ] ⟩ ] ]
       COMP-DTRS ⟨ [word
                    PHON    ⟨ kim ⟩
                    SYNSEM  [1] [LOCAL | CATEGORY [HEAD    noun
                                                   SUBCAT  ⟨ ⟩ ] ] ] ⟩ ] ]
```

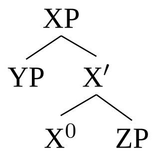
(29) Version/style: Sag et al. (2003)



1.3 Lexical Functional Grammar

1.3.1 Categories

- Basic categories
  - Lexical categories: V^0, N^0, A^0, P^0 (often written without the superscript zero)
  - Functional categories: D^0, I^0, C^0
- Formation of larger categories governed by **X-bar theory** (Chomsky 1970, Jackendoff 1977):



1.3.2 Combinatorics

- Annotated phrase structure rules license larger structures:

(30)

$$\begin{array}{llcc}
\text{IP} & \longrightarrow & \underset{(\uparrow \textsc{subj}) = \downarrow}{\text{DP}} & \underset{\uparrow = \downarrow}{\text{I}'} \\[2ex]
\text{I}' & \longrightarrow & \underset{\uparrow = \downarrow}{\text{I}} & \underset{\uparrow = \downarrow}{\text{VP}} \\[2ex]
\text{VP} & \longrightarrow & \underset{\uparrow = \downarrow}{\text{V}'} & \\[2ex]
\text{V}' & \longrightarrow & \underset{\uparrow = \downarrow}{\text{V}} & \underset{(\uparrow \textsc{obj}) = \downarrow}{\text{DP}}
\end{array}
\qquad
\begin{array}{llcc}
\text{DP} & \longrightarrow & \underset{\uparrow = \downarrow}{\text{D}'} & \\[2ex]
\text{D}' & \longrightarrow & \underset{\uparrow = \downarrow}{\text{D}} & \underset{\uparrow = \downarrow}{\text{NP}} \\[2ex]
\text{NP} & \longrightarrow & \underset{\uparrow = \downarrow}{\text{N}'} & \\[2ex]
\text{N}' & \longrightarrow & \underset{\uparrow = \downarrow}{\text{N}} &
\end{array}$$

- All phrase structure elements are optional and are present only if they dominate lexical material or are required by independent principles of the theory.
- The phrase structure rules license **constituent-structure** (c-structure) trees that are mapped to attribute-value matrices called **functional-structures** (f-structures) via the annotations.
- Interpretation of **metavariables** $\uparrow$ and $\downarrow$:

$\uparrow$:= the f-structure corresponding to the mother of the annotated node
("my mother's f-structure")

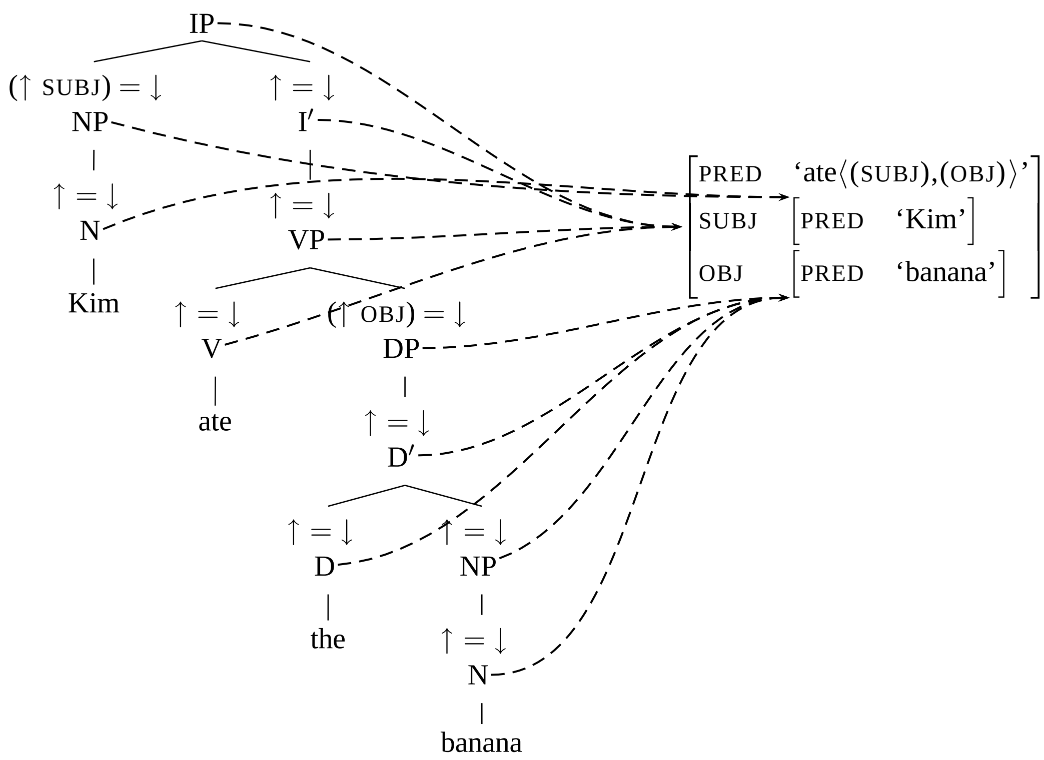
$\downarrow$:= the f-structure corresponding to the annotated node
("my f-structure")

- F-structures are functions in the mathematical sense and must satisfy the following condition:

Consistency (a.k.a Uniqueness Condition; this formulation from Dalrymple 2001:39)
In a given f-structure a particular attribute may have at most one value.

1.3.3 Example

(31)



2 Exercises

1. Consider the following sentence:

   (1) Kim gave Sandy the banana.

   i. Do a Categorial Grammar analysis of this sentence using the Steedman notation.
      a. What new lexical category do you have to assume?
   ii. Do an HPSG analysis using the Sag, Wasow, and Bender notation.
      a. Did you have to make any adjustments to the grammar sketched above?
   iii. Do an LFG analysis.
      a. Did you have to make any adjustments to the grammar sketched above?
      b. Did any issues arise about grammatical functions? If so, how did you deal with them?

2. Consider the following alternation of (1):

   (2) Kim gave the banana to Sandy.

   i. Do a Categorial Grammar analysis of this sentence using the Steedman notation.
      a. What new lexical categories do you have to assume?
      b. Are there any problems/issues with your new lexical items?
   ii. Do an HPSG analysis using the Sag, Wasow, and Bender notation.
      a. Did you have to make any adjustments to the grammar sketched above?
   iii. Do an LFG analysis.
      a. Did you have to make any adjustments to the grammar sketched above?
      b. Did any new issues arise about grammatical functions? If so, how did you deal with them? How does it fit with what you did for (1)?

References

Chomsky, Noam. 1970. Remarks on nominalization. In Richard Jacobs and Peter Rosenbaum, eds., *Readings in English Transformational Grammar*, 184–221. Waltham, MA: Ginn.

Dalrymple, Mary. 2001. *Lexical Functional Grammar*. San Diego, CA: Academic Press.

Jackendoff, Ray. 1977. *$\overline{X}$ Syntax: A Study of Phrase Structure*. Cambridge, MA: MIT Press.

Pollard, Carl, and Ivan A. Sag. 1994. *Head-driven Phrase Structure Grammar*. Chicago, IL and Stanford, CA: The University of Chicago Press and CSLI Publications.

Sag, Ivan A., Thomas Wasow, and Emily M. Bender. 2003. *Syntactic Theory: A Formal Introduction*. Stanford, CA: CSLI Publications, 2nd edn.

Steedman, Mark. 1996. *Surface Structure and Interpretation*. Cambridge, MA: MIT Press.

Grammar Formalisms

Course Notes 3

Ash Asudeh
University of Canterbury
ash.asudeh@canterbury.ac.nz

December 6–7, 2004

1 The Role of the Lexicon

1.1 Background

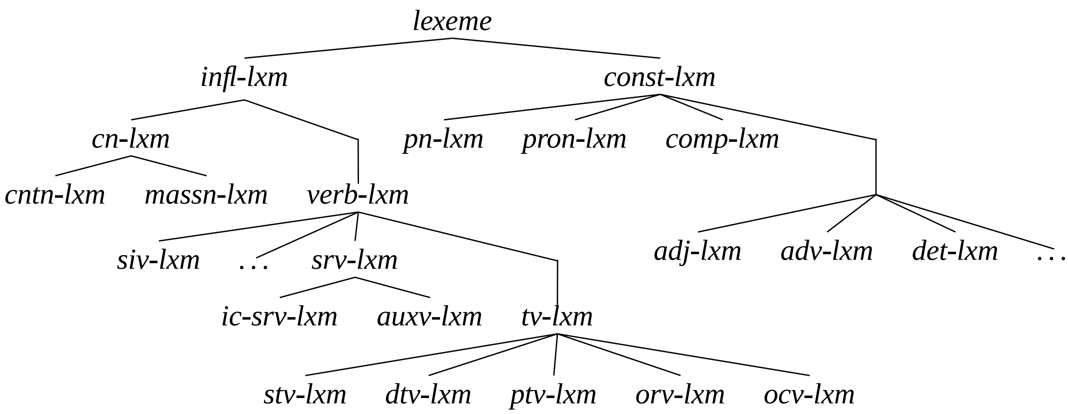
- Original generative view of the lexicon (largely abandoned): no rule-governed behaviour in the lexicon; repository of exceptions
- Current, generally accepted view: there are generalizations that need to be encoded lexically; lexicon is structured body of knowledge, not just repository of exceptions
    - *In the majority of current generative theories, the lexicon expresses linguistically significant generalizations, is highly structured, and delimits combinatoric possibilities to a great extent, through specifications in lexical entries.*
- Lexicalist theories, such as CG, HPSG and LFG, have developed increasingly sophisticated views of the lexicon:
    - Many transformations replaced by lexical rules (some seminal works are Bresnan 1978, Pollard and Sag 1987, Flickinger 1987)
    - Lexical items assigned rich representations (e.g., Kaplan and Bresnan 1982)‘
    - Various methods developed for factoring out common information from lexical items: lexical redundancy rules, type hierarchies, templates/macros
    - ⇒ **A lot of modern grammar engineering consists of formulating lexical entries.**

1.2 Categorial Grammar

- Lexical categories also encode (basic) combinatorics (Bar-Hillel 1953, Lambek 1958). This idea has also been adapted in HPSG (SUBCAT lists) and LFG (PRED features and Completeness and Coherence; see below).
- In Combinatory Categorial Grammar (Steedman 2000 is a recent overview), grammars are supplemented with combinators (you can think of these as analogous to schemas/rules) that add further combinatoric possibilities beyond purely lexically-specified ones.

1.3 Head-Driven Phrase Structure Grammar

- Type-hierarchies (Pollard and Sag 1987, Flickinger 1987)



(part of type hierarchy in Sag et al. 2003:492)

- Lexical rules (Pollard and Sag 1987, Flickinger 1987, Meurers 1999, 2001):

(1) Plural Noun Lexical Rule (adapted from Sag et al. 2003:503)

$$\begin{bmatrix} \textit{i-rule} & \\ \text{INPUT} & \left\langle \boxed{1}, \textit{cn-lxm} \right\rangle \\ \text{OUTPUT} & \left\langle \text{F}_{NPL}\langle \boxed{1} \rangle, \begin{bmatrix} \text{HEAD} \begin{bmatrix} \text{AGR} \begin{bmatrix} \text{NUM} & \textit{plural} \end{bmatrix} \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix}$$

1.4 Lexical Functional Grammar

- Templates (Dalrymple et al. to appear)

PRESENT	=	(↑ TENSE)=PRES
3PERSONSUBJ	=	(↑ SUBJ PERS)=3
SINGSUBJ	=	(↑ SUBJ NUM)=SG
3SG	=	@(3PERSONSUBJ)
	=	@(SINGSUBJ)
PRES3SG	=	@(PRESENT)
	=	@(3SG)
TRANSITIVE(P)	=	(↑ PRED)=‘P⟨(SUBJ),(OBJ)⟩’
INTRANSITIVE(P)	=	(↑ PRED)=‘P⟨(SUBJ)⟩’
TRANSITIVE-OR-INTRANSITIVE(P)	=	{ @(TRANSITIVE P) | @(INTRANSITIVE P) }

bakes	V	@(TRANSITIVE-OR-INTRANSITIVE `bake`) @(PRES3SG)	template form
bakes	V	{ (↑ PRED)=‘bake⟨(SUBJ),(OBJ)⟩’ | (↑ PRED)=‘bake⟨(SUBJ)⟩’ } (↑ SUBJ PERS)=3 (↑ SUBJ NUM)=SG	realized form

(2) a. John bakes bread. b. John bakes.

2 Heads

- Heads are distinguished lexical items that determine properties of larger phrases in which they occur, such as:
  - Category; e.g. a head verb together withs its complements forms a verb phrase (VP).
  - Agreement; e.g. the noun phrase *furry dogs* is plural, because its head noun is plural.
  - Complementation; e.g., the verb *hand* requires two complements (ditransitive), the verb *devour* requires one complement (transitive), and the verb *arrive* takes none (intransitive).

2.1 Categorial Grammar

- Categorial grammar does not have a native notion of head, in the sense of theoretically distinguishing a particular element.
- CG categories correspond tightly to functors and arguments (Steedman 1996). Functors can then be derivatively identified as heads. However, this is not completely straightforward in more powerful CGs. For example, type raising changes an argument into a higher-order functor on the functor that would normally apply to the argument. Which is the head? (advanced answer: the functor in the lowest type)
- The head of a sentence can be identified by finding the (verb) category whose leading edge is the final result of the derivation. The head of sentence (19) in course notes.2 is *ate*.

2.2 Head-Driven Phrase Structure Grammar

- The notion of head is unsurprisingly very important to HPSG.
- There is a feature HEAD that directly encodes the notion. The value of HEAD has many subtypes (see course notes.2, p.4).
- The majority of HPSG structures are headed structures. The head value of the entire structure is identified as token-identical to the head value of the head daughter, as we saw in course notes.2 (28).
- It is possible to follow paths of heads from lexical items all the way to the largest phrase/structure that they head by examining the structure-sharing of the HEAD value between mothers and daughters.

2.3 Lexical Functional Grammar

- LFG has a native notion of category/c-structure head, determined by X-bar theory (Chomsky 1970, Jackendoff 1977), and a derivative notion of f-structure head based on the $\uparrow$ and $\downarrow$ annotations.
- The f-structure head of a phrase is the lexical item that initiates the path of $\uparrow = \downarrow$ annotations that terminate at the top of the phrase.
- Functional categories in LFG are typically analyzed as **co-heads**. They bear the $\uparrow = \downarrow$ annotation and so does their sister. They therefore contribute information to the same f-structure as their sister.

3 Agreement

- Subject-verb

(3) Kim eats the banana.

(4) *Kim eat the banana.

(5) *Kim and Sandy eats the banana.

(6) Kim and Sandy eat the banana.

- Determiner/adjective-noun

(7) A cop chased Sandy.

(8) *A cops chased Sandy.

(9) Two cops chased Kim.

(10) *Two cop chased Kim.

3.1 Categorial Grammar

- A common method for specifying agreement in CG is by further annotating categories (Bach 1983, Steedman 1996):

(11) *eats* := $(S\backslash NP_{3S})/NP$

- Lack of specification for agreement features is understood as underspecification. The category for *eats* states that it requires a third person singular subject, but makes no requirement on its object.

3.1.1 Examples

(12)

Kim	eats	the	banana
NP_{3S}	$(S\backslash NP_{3S})/NP$	NP/N	N

the banana: NP/N, N $\Rightarrow$ NP ($>$)

eats the banana: $(S\backslash NP_{3S})/NP$, NP $\Rightarrow$ $S\backslash NP_{3S}$ ($>$)

Kim eats the banana: NP_{3S}, $S\backslash NP_{3S}$ $\Rightarrow$ S ($<$)

(13)

Cockroaches	eats	the	banana
NP_{3P}	$(S\backslash NP_{3S})/NP$	NP/N	N

the banana: NP/N, N $\Rightarrow$ NP ($>$)

eats the banana: $(S\backslash NP_{3S})/NP$, NP $\Rightarrow$ $S\backslash NP_{3S}$ ($>$)

Cockroaches eats the banana: NP_{3P}, $S\backslash NP_{3S}$ $\Rightarrow$ **FAIL**

3.2 Head-Driven Phrase Structure Grammar

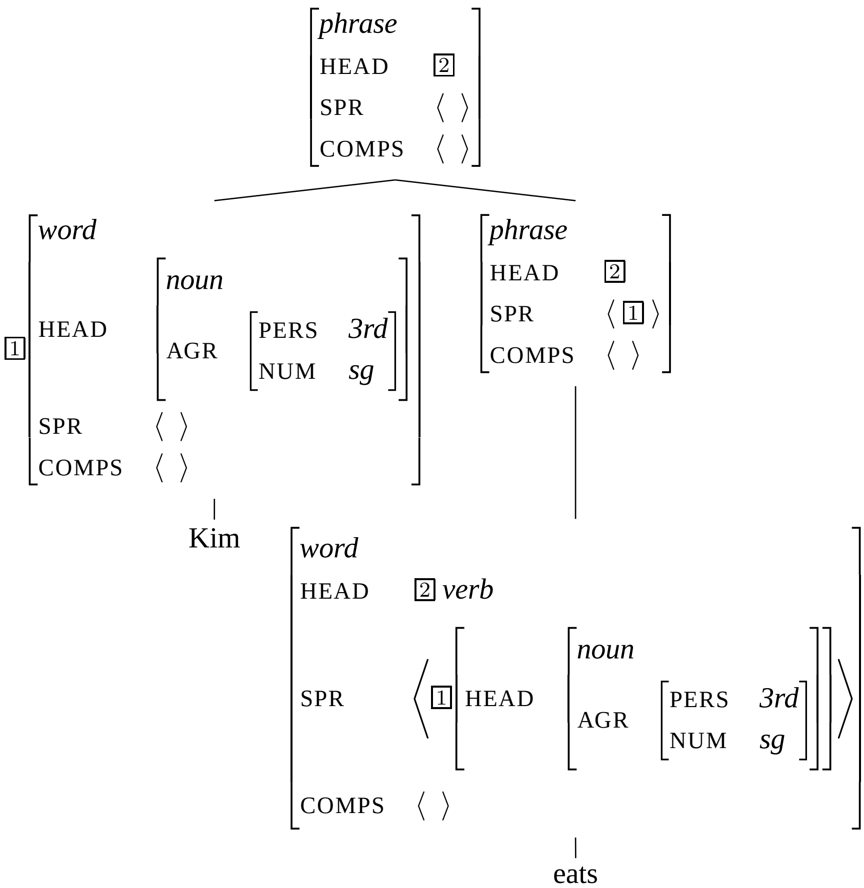
- Agreement captured by stating restriction on category in valence lists:

(14)
$$\begin{bmatrix} \textit{word} & \\ \text{PHONOLOGY} & \langle\, eats \,\rangle \\ \text{SPR} & \left\langle \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{noun} & \\ \text{AGR} & \begin{bmatrix} \textit{3sing} & \\ \text{PER} & \textit{3rd} \\ \text{NUM} & \textit{sg} \end{bmatrix} \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix}$$

- Agreeing item must bear features that can unify appropriately.

3.2.1 Example

(15)



3.3 Lexical Functional Grammar

- Agreement captured at f-structure. F-structure heads specify the agreement features of their arguments:

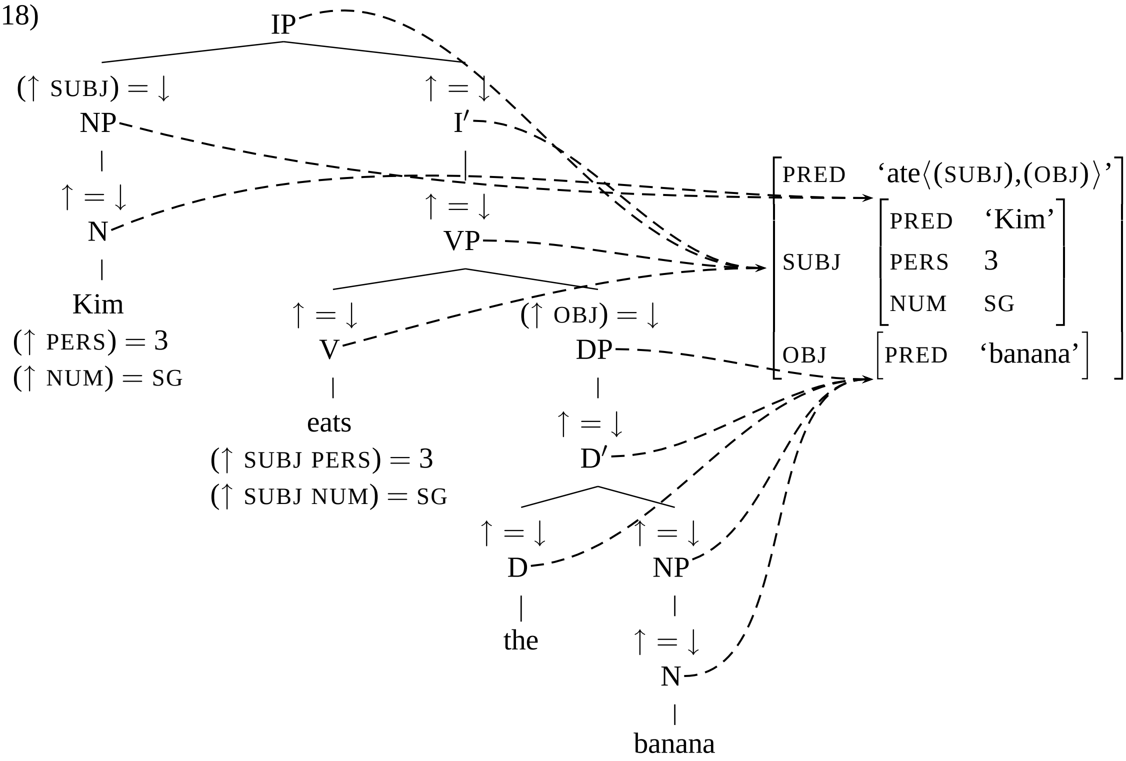
(16) *eats* V $(\uparrow \text{SUBJ PERS}) = 3$
$(\uparrow \text{SUBJ NUM}) = \text{SG}$

- These specifications will only be satisfiable (due to Consistency) if the agreeing item bears consistent specifications:

(17) *Kim* N $(\uparrow \text{PERS}) = 3$
$(\uparrow \text{NUM}) = \text{SG}$

3.3.1 Example

(18)



4 Complementation

- Heads select various aspects of their complements:
  - Number (a.k.a. valency)

(19) a. Thora handed Isak the toy.
b. *Thora handed Isak.
c. *Thora handed the toy.

(20) a. Thora devoured the cookie.
b. *Thora devoured Isak the cookie.
c. *Thora devoured.

(21) a. The train arrived.
b. *The train arrived the passengers.

  - Category

(22)	That Thora slept through the night surprised us.	CP subject
(23)	Thora surprised us.	NP subject
(24)	*Under the bed surprised us.	PP subject
(25)	*To find a leprechaun surprised us.	IP subject
(26)	*Very rare surprised us.	AP subject

  - Grammatical function

(27)	Thora seemed sleepy.	subject, NP
(28)	Under the bed seemed dusty.	subject, PP
(29)	Very rare seems to be how George likes his steak.	subject, AP
(30)	To find a leprechaun seems incredibly unlikely.	subject, IP
(31)	That Thora slept through the night seemed surprising.	, subject, CP

  - Grammatical features
    - Mood

(32)	Thora suspected that Ida had hidden the cookie.	declarative
(33)	*Thora suspected if Ida had hidden the cookie.	
(34)	*Thora enquired that she could have a cookie.	interrogative
(35)	Thora enquired if she could have a cookie.	

4.1 Categorial Grammar

- Number and category of complements directly encoded in lexical categories.
- Grammatical function typically derived from argument position in predicate-argument structure.
- Grammatical features either captured in semantics (predicate-argument structure) or through feature specifications on categories (similarly to agreement).

4.2 Head-Driven Phrase Structure Grammar

- Number of complements encoded on SUBCAT/VALENCE lists.
- Category of complement captured by stating restriction on category in valence lists:

(36)
$$\begin{bmatrix} \textit{word} & \\ \text{PHONOLOGY} & \langle\ surprise\ \rangle \\ \text{SPR} & \left\langle \begin{bmatrix} \text{HEAD} & \textit{noun} \vee \textit{comp} \end{bmatrix} \right\rangle \\ \text{COMPS} & \left\langle \begin{bmatrix} \text{HEAD} & \textit{noun} \end{bmatrix} \right\rangle \end{bmatrix}$$

- Grammatical function derived relationaly from position on SUBCAT list, according to obliqueness hierarchy:
  subject $\prec$ direct object $\prec$ indirect object $\prec$ oblique $\prec$ other complements
- Grammatical features selected through valence lists, analogously to agreement and category.

4.3 Lexical Functional Grammar

- Number of complements and their grammatical functions encoded in PRED feature:

  devour $(\uparrow \text{PRED}) =$ 'devour$\langle$(SUBJ),(OBJ)$\rangle$'

- The principle of Completeness and Coherence ensure that the subcategorization requirements of the predicate are satisfied:
  - **Completeness** (adapted from Dalrymple 2001:37 and Kaplan and Bresnan 1982)
    An f-structure is *complete* if and only if it contains all the grammatical functions that its predicate governs.
  - **Coherence** (adapted from Dalrymple 2001:39 and Kaplan and Bresnan 1982)
    An f-structure is *coherent* if and only if all the governable grammatical functions that it contains are governed by a local predicate.
  - *Governable grammatical functions* := GFs that can be subcategorized for
  - A predicate governs a grammatical function iff the grammatical function is mentioned in the predicate's PRED feature.
- Grammatical features of complement specified through functional equations:

  $(\uparrow \text{COMP MOOD}) = \text{DECLARATIVE}$

- Category selected through interplay of grammatical function annotations on c-structure rules and satisfaction of PRED subcategorization requirements (i.e., Completeness and Coherence).

5 Exercise

1. Consider the following sentence:

(1) Few cats hand people money.

i. Do a Categorial Grammar analysis of this sentence using the Steedman notation.
   a. Account for all agreement relations and complementation requirements.

ii. Do an HPSG analysis using the Sag, Wasow, and Bender notation.
   a. Account for all agreement relations and complementation requirements
   b. Did you have to make any adjustments to the grammar developed so far?

iii. Do an LFG analysis.
   a. Account for all agreement relations and complementation requirements
   b. Did you have to make any adjustments to the grammar developed so far?

References

Bach, Emmon. 1983. On the Relationship between Word-Grammar and Phrase-Grammar. *Natural Language and Linguistic Theory* 1: 65–89.

Bar-Hillel, Yehoshua. 1953. A quasi-arithmetical notation for syntactic description. *Language* 29: 47–58.

Bresnan, Joan. 1978. A Realistic Trasnformational Grammar. In Morris Halle, Joan Bresnan, and George A. Miller, eds., *Linguistic Theory and Psychological Reality*, 1–59. Cambridge, MA: MIT Press.

Chomsky, Noam. 1970. Remarks on nominalization. In Richard Jacobs and Peter Rosenbaum, eds., *Readings in English Transformational Grammar*, 184–221. Waltham, MA: Ginn.

Dalrymple, Mary. 2001. *Lexical Functional Grammar*. San Diego, CA: Academic Press.

Dalrymple, Mary, Ronald M. Kaplan, and Tracy Holloway King. to appear. Lexical Structure as Generalizations over Descriptions. In Miriam Butt and Tracy Holloway King, eds., *Proceedings of the LFG04 Conference*. Stanford, CA: CSLI Publications.

Flickinger, Daniel. 1987. Lexical Rules in the Hierarchical Lexicon. Ph.D. thesis, Stanford University.

Jackendoff, Ray. 1977. *$\overline{X}$ Syntax: A Study of Phrase Structure*. Cambridge, MA: MIT Press.

Kaplan, Ronald M., and Joan Bresnan. 1982. Lexical-Functional Grammar: A Formal System for Grammatical Representation. In Joan Bresnan, ed., *The Mental Representation of Grammatical Relations*, 173–281. Cambridge, MA: MIT Press.

Lambek, Joachim. 1958. The Mathematics of Sentence Structure. *American Mathematical Monthly* 65(154–70).

Meurers, Walt Detmar. 1999. Lexical Generalizations in the Syntax of German Non-Finite Constructions. Ph.D. thesis, Universität Tübingen, Tübingen, Germany. Volume 145 in Arbeitspapiere des SFB 340.

—. 2001. On expressing lexical generalizations in HPSG. *Nordic Journal of Linguistics* 24(2): 161–217.

Pollard, Carl, and Ivan A. Sag. 1987. *Information-Based Syntax and Semantics*. CSLI Publications.

Sag, Ivan A., Thomas Wasow, and Emily M. Bender. 2003. *Syntactic Theory: A Formal Introduction*. Stanford, CA: CSLI Publications, 2nd edn.

Steedman, Mark. 1996. *Surface Structure and Interpretation*. Cambridge, MA: MIT Press.

—. 2000. Information Structure and the Syntax-Phonology Interface. *Linguistic Inquiry* 31(4).

Grammar Formalisms

ALTSS 2004
Course Notes 4

Ash Asudeh
University of Canterbury
ash.asudeh@canterbury.ac.nz

December 6–7, 2004

1 Modifiers

- Some common modifiers:
  - Adjectives

(1) a **red** car
(2) an **American** car
(3) a **big** cockroach
(4) a **former** senator

  - Adverbs

(5) a **really** big cockroach
(6) She foxtrots **fabulously**.
(7) She foxtrots **daily**.
(8) John **quickly** hid the evidence.
(9) **Obviously**, he is nuts.

  - Prepositional phrases

(10) He arrived **on the train**.
(11) He arrived **at the train station**.
(12) He arrived **in one hour**.
(13) He rode the train **for one hour**.
(14) Is that the man **from France**?

  - Noun phrases

(15) She foxtrots **every day**.
(16) She does not foxtrot **here**.

  - Relative clauses[1]

(17) Surgeons **who are talented** deserve awards.
(18) Surgeons, **who are talented**, deserve awards.

[1] Relative clauses present the added complication of containing an unbounded dependency:
(i) The surgeons who the *Guardian* reported that the *Lancet* declared are talented deserve awards.
We will not account for this complication here.

1.1 Categorial Grammar

- Modifiers are of the general form A/A or A\A:
  - Adjective: N/N
  - Adverb: (S\NP)|(S\NP), S/S, (N/N)/(N/N)
  - Prepositional phrase: (S\NP)\(S\NP), N\N

1.1.1 Examples

(19)

Kim	ate	the	big	yellow	banana	quickly
NP	(S\NP)/NP	NP/N	N/N	N/N	N	(S\NP)|(S\NP)

- yellow + banana: N/N N ⇒ N (>)
- big + [yellow banana]: N/N N ⇒ N (>)
- the + [big yellow banana]: NP/N N ⇒ NP (>)
- ate + [the big yellow banana]: (S\NP)/NP NP ⇒ S\NP (>)
- [ate the big yellow banana] + quickly: S\NP (S\NP)|(S\NP) ⇒ S\NP (<)
- Kim + [ate the big yellow banana quickly]: NP S\NP ⇒ S (<)

(20)

Kim	quickly	ate	the	big	yellow	banana
NP	(S\NP)|(S\NP)	(S\NP)/NP	NP/N	N/N	N/N	N

- yellow + banana: N/N N ⇒ N (>)
- big + [yellow banana]: N/N N ⇒ N (>)
- the + [big yellow banana]: NP/N N ⇒ NP (>)
- ate + [the big yellow banana]: (S\NP)/NP NP ⇒ S\NP (>)
- quickly + [ate the big yellow banana]: (S\NP)|(S\NP) S\NP ⇒ S\NP (>)
- Kim + [quickly ate the big yellow banana]: NP S\NP ⇒ S (<)

1.2 Head-Driven Phrase Structure Grammar

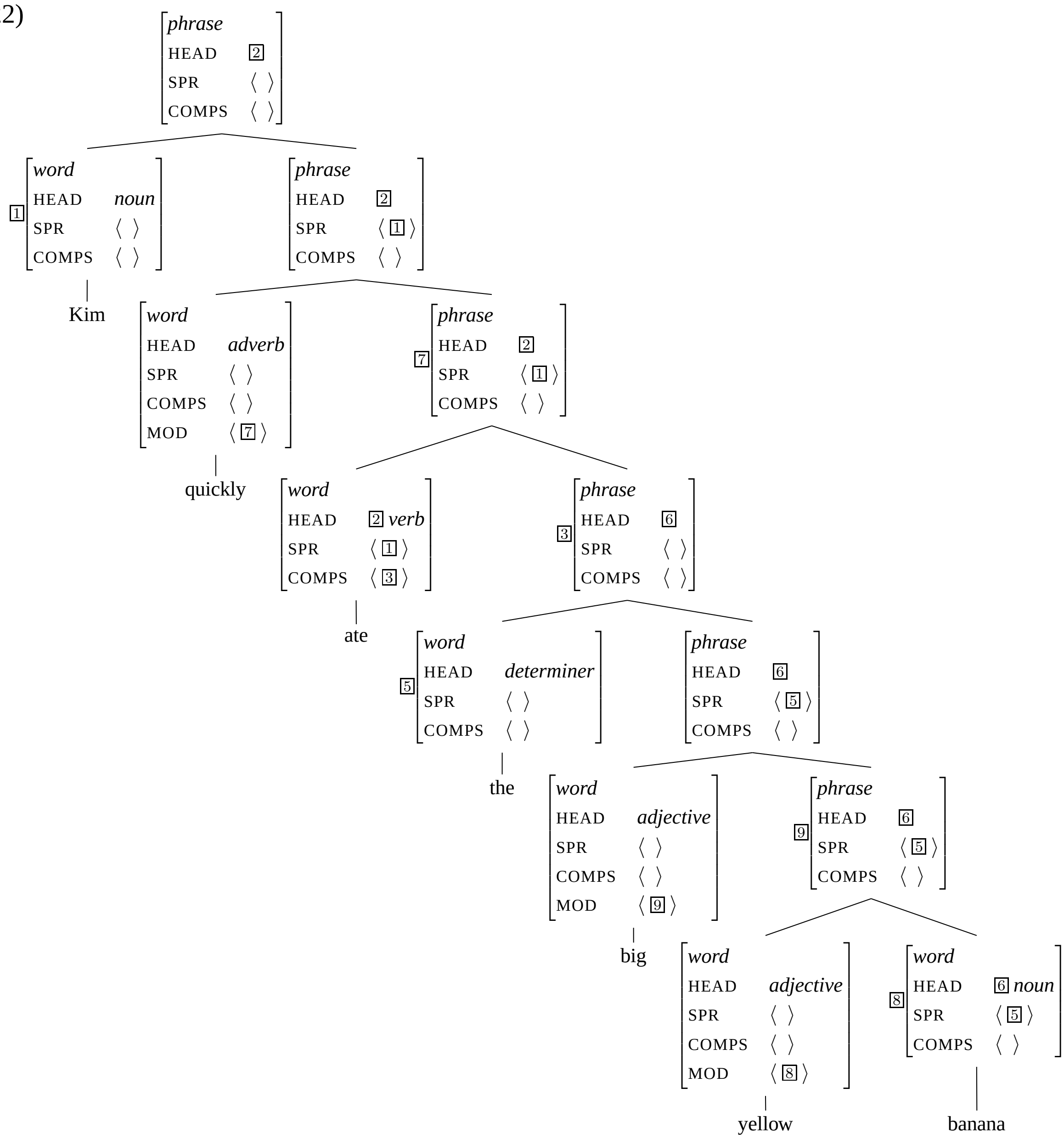
- Add another schema/rule:

(21) Head-modifier rule (adapted from Sag et al. 2003)

$$[\textit{phrase}] \rightarrow \mathbf{H}\boxed{1}\begin{bmatrix}\text{COMP} & \langle\ \rangle\end{bmatrix} \quad , \quad \begin{bmatrix}\text{COMPS} & \langle\ \rangle \\ \text{MOD} & \langle \boxed{1} \rangle\end{bmatrix}$$

1.2.1 Example

(22)



1.3 Lexical Functional Grammar

- Expand our set of c-structure rules to deal with modifiers:

$$\text{IP} \longrightarrow \underset{\downarrow \in (\uparrow \text{ADJ})}{\text{AdvP}} \quad \underset{\uparrow = \downarrow}{\text{IP}}$$

$$\text{VP} \longrightarrow \underset{\downarrow \in (\uparrow \text{ADJ})}{\text{AdvP}} \quad \underset{\uparrow = \downarrow}{\text{VP}} \quad \underset{\downarrow \in (\uparrow \text{ADJ})}{\text{AdvP}}$$

$$\text{N}' \longrightarrow \underset{\downarrow \in (\uparrow \text{ADJ})}{\text{AP}^*} \quad \underset{\uparrow = \downarrow}{\text{N}'}$$

$$\text{AP} \longrightarrow \underset{\uparrow = \downarrow}{\text{A}'}$$

$$\text{A}' \longrightarrow \underset{\downarrow \in (\uparrow \text{ADJ})}{\text{AdvP}} \quad \underset{\uparrow = \downarrow}{\text{A}'}$$

$$\text{A}' \longrightarrow \underset{\uparrow = \downarrow}{\text{A}}$$

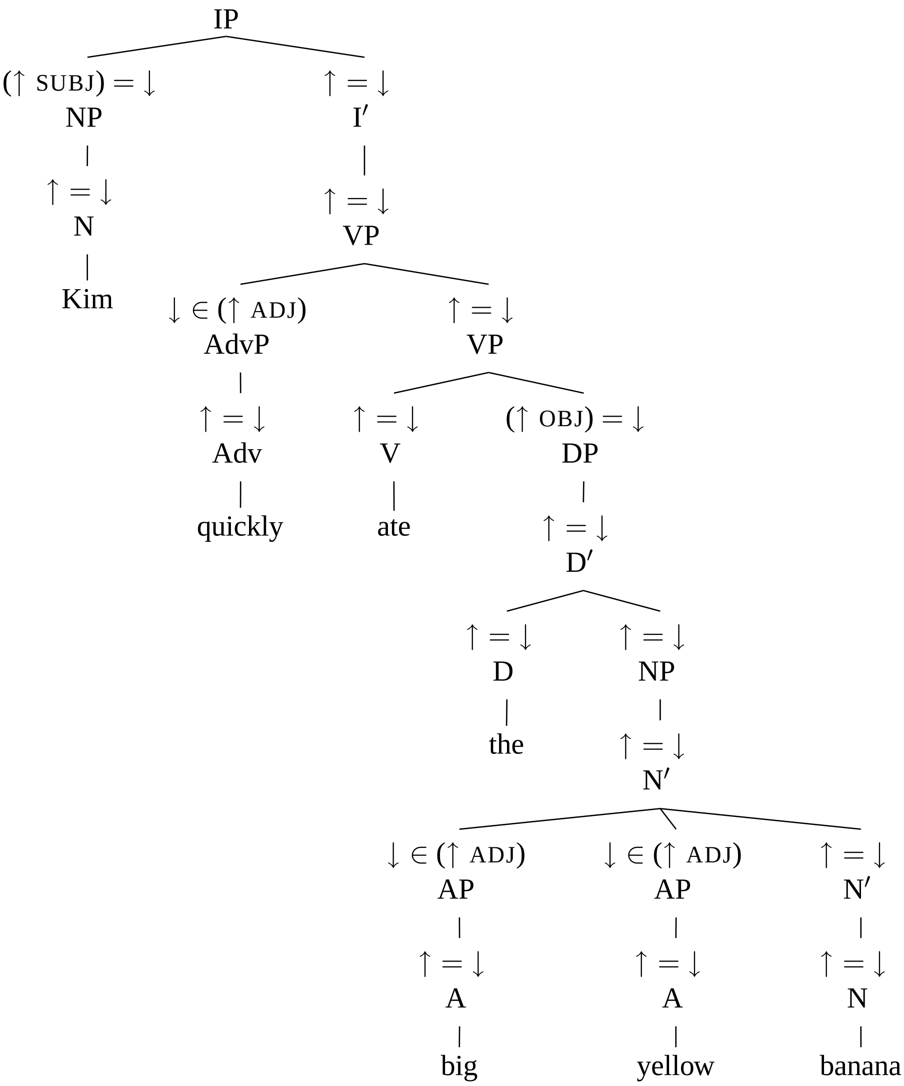
$$\text{AdvP} \longrightarrow \underset{\uparrow = \downarrow}{\text{Adv}'}$$

$$\text{Adv}' \longrightarrow \underset{\uparrow = \downarrow}{\text{Adv}}$$

- A new kind of grammatical function at f-structure, ADJUNCT (abbreviated as ADJ):
  - ADJ has a **set** as a value. The set contains all of the item's modifiers, in a flat, unordered representation.
  - The annotation $\downarrow \in (\uparrow \text{ADJ})$ means that the f-structure of the node bearing the annotation is a member of the adjunct set of the mother's f-structure.
  - Completeness and Coherence do not apply to ADJ, because it is not a subcategorized/governable grammatical function.

1.3.1 Example

(23)



$$\begin{bmatrix} \textsc{pred} & \text{`ate}\langle(\textsc{subj}),(\textsc{obj})\rangle\text{'} \\ \textsc{subj} & \begin{bmatrix} \textsc{pred} & \text{`Kim'} \end{bmatrix} \\ \textsc{obj} & \begin{bmatrix} \textsc{pred} & \text{`banana'} \\ \textsc{adj} & \left\{ \begin{bmatrix} \textsc{pred} & \text{`big'} \end{bmatrix}, \begin{bmatrix} \textsc{pred} & \text{`yellow'} \end{bmatrix} \right\} \end{bmatrix} \\ \textsc{adj} & \left\{ \begin{bmatrix} \textsc{pred} & \text{`quickly'} \end{bmatrix} \right\} \end{bmatrix}$$

2 Exercises

1. Consider the following sentence:

   (1) Kim ate the very big banana.

   i. Do a Categorial Grammar analysis of this sentence using the Steedman notation.
      a. What new lexical category do you have to assume?
   ii. Do an HPSG analysis using the Sag, Wasow, and Bender notation.
      a. Did you have to make any adjustments to the grammar developed so far?
   iii. Do an LFG analysis.
      a. Did you have to make any adjustments to the grammar developed so far?

2. Consider the following sentence:

   (2) Kim ate the banana on the train.

   i. Do a Categorial Grammar analysis of this sentence using the Steedman notation.
      a. What new lexical category do you have to assume?
      b. Does this category bear any relationship to any of the lexical categories you had to develop for course notes.2, exercise 2? If so, describe the relationship.
   ii. Do an HPSG analysis using the Sag, Wasow, and Bender notation.
      a. Did you have to make any adjustments to the grammar developed so far?
   iii. Do an LFG analysis.
      a. Did you have to make any adjustments to the grammar developed so far?
   iv. Is the sentence ambiguous? If it is, can you account for both parses in each framework (CG, HPSG, LFG)?

References

Sag, Ivan A., Thomas Wasow, and Emily M. Bender. 2003. *Syntactic Theory: A Formal Introduction*. Stanford, CA: CSLI Publications, 2nd edn.

Steedman, Mark. 1996. *Surface Structure and Interpretation*. Cambridge, MA: MIT Press.

Grammar Formalisms

ALTSS 2004
Course Notes 5

Ash Asudeh
University of Canterbury
ash.asudeh@canterbury.ac.nz

December 6–7, 2004

1 Grammar Engineering

1.1 Categorial Grammar

- Grok

  Comments grammar implementation platform, final release 24/02/2003

  License GNU Library or Lesser General Public License (LGPL)

  Availability downloadable

  URL http://grok.sourceforge.net/

1.2 HPSG

- Linguistic Knowledge Building (LKB; Copestake 2002)

  Comments state-of-the-art, grammar and lexicon development environment

  License Open Source

  Availability downloadable

  URL http://www.delph-in.net/lkb/
  (also see http://cslipublications.stanford.edu/lkb.html; contains some useful links)

- ALE (Attribute Logic Engine)

  Comments grammar implementation platform for typed-feature structure grammars, especially HPSG; semi-maintained (no major versions since 1999)

  License GNU Lesser General Public License

  Availability downloadable

  URL http://www.cs.toronto.edu/∼gpenn/ale.html

- TRALE

  Comments grammar-implementation platform based on ALE and ConTroll; not publicly available yet

  License ?

  Availability contact developers for project-internal release (see web site)

  URL http://www.sfs.uni-tuebingen.de/hpsg/archive/projects/trale/

- ConTroll

  Comments legacy project (finished 1997, principal results incorporated in TRALE); implements logical foundations of HPSG; no parser

  License free to "people and institutions which make all their research results public"

  Availability downloadable Prolog source code

  URL http://www.sfs.uni-tuebingen.de/controll/

- See Copestake 2002:156–157 for more references to systems for HPSG and other frameworks.

1.3 LFG

- Xerox Linguistics Environment (XLE; Butt et al. 1999)

  Comments state-of-the-art, grammar and lexicon development environment

  License free for education

  Availability downloadable with username/password after license paperwork has been filed

  URL http://www2.parc.com/istl/groups/nltt/xle/

- Grammar Writer's Workbench (a.k.a Medley)

  Comments legacy system (replaced by XLE), semi-maintained, development environment

  License free for research and education

  Availability downloadable

  URL http://www2.parc.com/istl/groups/nltt/medley/

References

Butt, Miriam, Tracy Holloway King, María-Eugenia Niño, and Frédérique Segond. 1999. *A Grammar Writer's Cookbook*. Stanford, CA: CSLI Publications.

Copestake, Ann. 2002. *Implementing Typed Feature Structure Grammars*. Stanford, CA: CSLI Publications.

Grammar Formalisms

Selected References

Ash Asudeh
University of Canterbury
ash.asudeh@canterbury.ac.nz

December 6–7, 2004

Note: These references are not meant to be exhaustive in any way. They are rather meant to serve as useful starting points (mainly books) for finding out more information.

Foundations of Generative Linguistics

Chomsky, Noam. 1957. *Syntactic Structures*. The Hague: Mouton.

—. 1965. *Aspects of the Theory of Syntax*. MIT Press.

—. 1970. Remarks on nominalization. In Richard Jacobs and Peter Rosenbaum, eds., *Readings in English Transformational Grammar*, 184–221. Waltham, MA: Ginn.

Grammar Formalisms

Green, Geogia M., and Jerry L. Morgan. 2001. *A Practical Guide to Syntactic Analysis*. Stanford, CA: CSLI Publications, 2nd edn.

Sells, Peter. 1985. *Lectures on Contemporary Syntactic Theory*. Stanford, CA: CSLI Publications.

Shieber, Stuart M. 1986. *An introduction to unification-based approaches to grammar*. Stanford, CA: CSLI Publications.

Grammar Engineering

Butt, Miriam, Tracy Holloway King, María-Eugenia Niño, and Frédérique Segond. 1999. *A Grammar Writer's Cookbook*. Stanford, CA: CSLI Publications.

Copestake, Ann. 2002. *Implementing Typed Feature Structure Grammars*. Stanford, CA: CSLI Publications.

Categorial Grammar

Introductory

Wood, Mary McGee. 1993. *Categorial Grammars*. London: Routledge.

Foundational

Bar-Hillel, Yehoshua. 1953. A quasi-arithmetical notation for syntactic description. *Language* 29: 47–58.

Lambek, Joachim. 1958. The Mathematics of Sentence Structure. *American Mathematical Monthly* 65(154–70).

Further Developments

van Benthem, Johan. 1991. *Language in Action: Categories, Lambdas, and Dynamic Logic*. Amsterdam: North-Holland. Reprinted in 1995 by the MIT Press, Cambridge, MA.

Buszkowski, Wojciech, Witold Marciszewski, and Johan van Benthem, eds. 1988. *Categorial Grammar*. Amsterdam: John Benjamins.

Carpenter, Bob. 1997. *Type-Logical Semantics*. Cambridge, MA: MIT Press.

Moortgat, Michael. 1997. Categorial type logics. In Johan van Benthem and Alice ter Meulen, eds., *Handbook of Logic and Language*, 93–177. Cambridge, MA: MIT Press. Co-published with Elsevier Science B.V., Amsterdam, The Netherlands.

Morrill, Glyn. 1994. *Type Logical Grammar*. Dordrecht: Kluwer.

Oehrle, Richard T., Emmon Bach, and Deirdre Wheeler, eds. 1988. *Categorial Grammars and Natural Language Structures*. Dordrecht: Reidel.

Steedman, Mark. 1996. *Surface Structure and Interpretation*. Cambridge, MA: MIT Press.

—. 2000. *The Syntactic Process*. Cambridge, MA: MIT Press.

Head-Driven Phrase Structure Grammar

Introductory

Sag, Ivan A., Thomas Wasow, and Emily M. Bender. 2003. *Syntactic Theory: A Formal Introduction*. Stanford, CA: CSLI Publications, 2nd edn.

Borsley, Robert D. 1996. *Modern Phrase Structure Grammar*. Oxford: Blackwell.

Foundational

Pollard, Carl, and Ivan A. Sag. 1987. *Information-Based Syntax and Semantics*. CSLI Publications.

—. 1994. *Head-driven Phrase Structure Grammar*. Chicago, IL and Stanford, CA: The University of Chicago Press and CSLI Publications.

Further Developments

Ginzburg, Jonathan, and Ivan Sag. 2001. *Interrogative Investigations*. Stanford, CA: CSLI Publications.

Levine, Robert, and Georgia Green, eds. 1999. *Studies in Contemporary Phrase Structure Grammar*. Cambridge: Cambridge University Press.

Webelhuth, Gert, Jean-Pierre Koenig, and Andreas Kathol, eds. 1999. *Lexical and Constructional Aspects of Linguistic Explanation*. Stanford, CA: CSLI.

Lexical Functional Grammar

Foundational

Bresnan, Joan, ed. 1982. *The Mental Representation of Grammatical Relations*. Cambridge, MA: MIT Press.

Dalrymple, Mary, Ronald M. Kaplan, John T. Maxwell, and Annie Zaenen, eds. 1995. *Formal issues in Lexical-Functional Grammar*. Stanford, CA: CSLI Publications.

Introductory

Falk, Yehuda. 2001. *Lexical-Functional Grammar: An Introduction to Parallel Constraint-Based Syntax*. Stanford, CA: CSLI Publications.

Kroeger, Paul. 2004. *Analyzing Syntax: A Lexical-Functional Approach*. Cambridge: Cambridge University Press.

Further Developments

Bresnan, Joan. 2000. Optimal Syntax. In Joost Dekkers, Frank van der Leeuw, and Jeroen van de Weijer, eds., *Optimality Theory: Phonology, Syntax, and Acquisition*, 334–385. Oxford: Oxford University Press.

—. 2001. *Lexical-Functional Syntax*. Oxford: Blackwell.

Dalrymple, Mary, ed. 1999. *Semantics and Syntax in Lexical Functional Grammar: The Resource Logic Approach*. Cambridge, MA: MIT Press.

Dalrymple, Mary. 2001. *Lexical Functional Grammar*. San Diego, CA: Academic Press.

Kuhn, Jonas. 2003. *Optimality-Theoretic Syntax: A Declarative Approach*. Stanford, CA: CSLI Publications.

Syntactic theory: A formal introduction (2nd. ed)

Ivan A. Sag, Thomas Wasow, Emily M. Bender.
Stanford, CA: CSLI Publications. 2003.

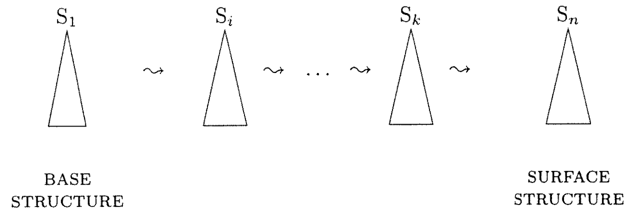
Used with permission.

Appendix B
Related Grammatical Theories

As noted in Chapter 2, the theory of grammar developed in this text is most closely related to the framework known as 'Head-driven Phrase Structure Grammar', or HPSG. HPSG is one of a number of frameworks for grammatical analysis that have been developed within the Chomskyan paradigm, broadly conceived. The intellectual tradition our theory represents is eclectic in its orientation, synthesizing ideas from several approaches to the study of language. To clarify these connections, we provide here a brief and incomplete survey of related theories of grammar. We hope not only to provide some information about the intellectual roots of our approach to syntax, but also to explicate its relationship to other contemporary theories.

The basic concept of generative grammar is simply a system of rules that defines in a formally precise way (i.e. 'generates') a set of sequences (strings over some vocabulary of words or 'formatives') that represent the well-formed sentences of a given language.[1] Thus both of the systems considered in Chapter 2 – the regular expression (finite-state) and context-free phrase structure grammars – are generative grammars, as is the grammar summarized in Appendix A.

Generative syntax began in the 1950s when Noam Chomsky and others he influenced developed and formalized a theory of grammar based on the notion of 'transformation'.[2] The architecture of a transformational generative grammar defines sentence well-formedness indirectly: first, base (or 'underlying' or 'kernel') structures are generated via a system of phrase structure rules;[3] and then transformational rules apply successively to map these phrase structures into other phrase structures. The sentences of the language, then, are just those that can be derived by applying transformational rules to the base structures according to a particular regime, e.g. a regime of ordered transformations, at least some of which are obligatory. A transformational derivation thus involves a sequence of phrase structures, the first of which is a base structure and the last of which is a phrase structure (usually called a 'surface structure') whose word string corresponds to a sentence of the language:



Transformational generative grammar (which has dominated the mainstream of syntactic theory from the 1960s through to the present) has changed significantly over the years. Yet, despite considerable evolution within this framework, the notion of transformational derivation has been present in one guise or another in virtually every formulation of transformational grammar.[4] Similarly, other commonalities remain in the practice of transformational grammarians, such as the treatment of sublexical entities (e.g. inflectional affixes) as independent syntactic elements, that is, as syntactic primitives on a par with words.[5]

In contrast to the transformational tradition, there is another approach to generative grammar, equally committed (if not more so) to the original goal of developing precisely formulated grammars. This tradition has two distinctive properties:

(i) **Constraint-Based Architecture**: Grammars are based on the notion of constraint satisfaction, rather than transformational derivation.

(ii) **Strong Lexicalism**: Words, formed in accordance with an independent lexical theory (or 'module'), are the atoms of the syntax. Their internal structure is invisible to syntactic constraints.

These two design properties together form the basis of the 'constraint-based lexicalist' (CBL) approach to generative grammar. In CBL approaches, surface structures are

[1] We follow here Chomsky's (1975,1966) original usage of this term. The term 'generative grammar' is sometimes also used to refer to 'generative-transformational grammar' (see below). There is no small irony in this usage, given the characteristic practice of transformational work of the last two decades. This body of literature, though technical in appearance, has systematically eschewed the development of consistent, broad coverage systems whose predictions can be verified empirically. This stands in marked contrast to the formal and analytic precision that is emblematic of most of the nontransformational generative approaches we survey below, whose mathematical properties (including their generative capacity, both stringset and structural) has in many cases been explored in detail, and whose empirical coverage has frequently been tested in terms of large-scale, computational grammars.

[2] Chomsky's notion of transformation grew out of the theory of linguistic transformations invented in the 1950s by Zellig Harris. A number of relevant papers are collected in Harris 1970. On Harris' often overlooked contribution to theoretical linguistics and related areas, see Nevin 2003 and Nevin and Johnson 2003.

[3] In some versions of this approach, lexical insertion into the structures generated by the CFG is accomplished by a separate specialized mechanism.

[4] There are exceptions, though, e.g. Koster 1987 and Brody 1995. Chomsky has always maintained that it is 'not easy' to provide empirical (or theory-independent) evidence that would lead one to prefer a transformational theory over simpler alternatives. Despite this repeated claim (Chomsky (1981: 90f; 1995: 223f)), Chomsky has included transformational operations in every version of grammatical theory he has developed since the 1950s. In more recent work, Chomsky (2002) claims that transformations are a 'conceptual necessity'; see Levine and Sag 2003 for some further discussion.

[5] Hence Chomsky's introduction of the term (syntactic) 'formative' to encompass stems, noninflecting words, and inflectional affixes.

generated directly, though ancillary kinds of syntactic representation may be cogenerated (see below).

The principle of strict lexicalism has its origin in the pioneering work of Chomsky (1970), who challenged previous attempts to derive nominalizations (e.g. *the enemy's destruction of the city*) from clauses (e.g. *the enemy destroyed the city*) via syntactic transformations. In the mid and late 1970s, many alternatives to transformational analyses were developed. There are two particularly significant developments in this period. The first is Bresnan's 'Realistic' Transformational Grammar (widely circulated in unpublished form; a version was published as Bresnan 1978), which for the first time provided a cogent treatment of numerous phenomena (e.g. passivization) in lexical rather than transformational terms. Bresnan's dramatic first step inspired a number of people, notably Brame (1979) and Gazdar (1981) [first drafted in 1980], to take the further step of purging transformations from syntactic theory altogether. Second, the emergence of the framework of Montague Grammar provided new techniques for characterizing meanings directly in terms of surface structure, thereby eventually eliminating any semantic motivation for syntactic transformations. In many versions of transformational grammar, active and passive sentences were derived from a common underlying structure, leading to the (controversial) suggestion that many aspects of meaning are preserved by transformational derivations. With the advent of more sophisticated methods of semantic analysis, distinct surface structures could be assigned formally distinct but equivalent semantic interpretations, thus accounting for the semantics in a principled fashion without appeal to transformations.

'Realistic' Transformational Grammar and Montague Grammar together set the stage for the emergence of fully nontransformational generative frameworks in the late 1970s and early 1980s. Most notable among these are Lexical Functional Grammar (LFG), Generalized Phrase Structure Grammar (GPSG), Categorial Grammar (CG), and Dependency Grammar, each of which we summarize below. The subsequent history of CBL generative grammar witnessed not only considerable development in each of these frameworks, but also the introduction of other new approaches, notably, Construction Grammar (CxG). Of immediate relevance also is the evolution of GSPG, through the integration of ideas from various other frameworks, into the framework of HPSG, from which many analyses and the general orientation of the present text are directly drawn.

Not all influential theories of grammar fit comfortably into our dichotomy between transformational and CBL approaches. In particular, three frameworks that need to be mentioned are Relational Grammar (RG), Tree Adjoining Grammar (TAG), and Optimality Theory (OT). These are all discussed briefly following our survey of CBL theories.

The primary focus of this appendix is a survey of CBL approaches. However, most CBL frameworks grew out of early transformational work, and more recent developments within transformational grammar have continued to exert an influence upon CBL research. For this reason, our survey begins with a very brief historical sketch of relevant aspects of transformational generative grammar.

B.1 Historical Sketch of Transformational Grammar, ca. 1955 to the present

An early version of transformational generative grammar was presented in Chomsky's 1957 book, *Syntactic Structures*, which was a condensed version of portions of *The Logical Structure of Linguistic Theory*, a magnum opus completed in 1955, but not published until twenty years later. The analyses presented there and in other transformational works of the period included explicit formal statements of rules intended to license all and only the well-formed sentences of the language under discussion (usually English). This emphasis on the precise formulation of hypotheses is perhaps the greatest influence of early transformational grammar on the approach presented here (along with such tools as context-free grammar, which were also invented by Chomsky during this period).

As noted above, a key claim of transformational grammar (in all its versions) is that an empirically adequate grammar requires that sentences be associated not with a single tree structure, but with a sequence of trees, each related to the next by a transformation. The initial trees in Chomsky's *Syntactic Structures* theory were to be generated by a CFG. For example, passive sentences (such as *The cat was chased by the dog*) were derived from the same underlying structures as their active counterparts (*The dog chased the cat*) by means of a passivization transformation that permuted the order of the two NPs and inserted the words *be* and *by* in the appropriate places.

The most celebrated analysis in this theory is its treatment of the English auxiliary system (roughly, the material covered in Chapter 13 of this text). Chomsky (1957) proposed that tense was, in the underlying syntactic structure, a 'formative' separate from the verb on which it ultimately appears. A movement transformation was posited to account for inversion in questions (deriving, e.g. *Is the sun shining?* from the same underlying structure as *The sun is shining*); and an insertion transformation placed *not* in the appropriate position for sentence negation. Both these transformations in some instances have the effect of stranding tense – that is, leaving it in a position not adjacent to any verb. For these cases, Chomsky posited a transformation to insert *do* as a carrier of tense. Several other uses of auxiliary *do* (e.g. in ellipsis) were also treated on this view as instances of tense stranding. This unified account of apparently disparate uses of *do*, together with the formal explicitness of the presentation, won many converts to transformational grammar.

Katz and Postal (1964) and Chomsky (1965) introduced a number of major changes into transformational grammar, and Chomsky dubbed the resulting theory 'the Standard Theory'. It differed from earlier transformational grammar in several ways, some rather technical. Among the important innovations of this theory were the use of recursive phrase structure rules (allowing for the elimination of transformations that combined multiple trees into one) and the introduction of syntactic features to account for subcategorization (valence).

Perhaps the most important conceptual change was the addition of a semantic component to the theory of transformational grammar. In this theory, the initial tree in each sentence's derivation, known as its 'deep structure', transparently represented all the information necessary for semantic interpretation. In particular, it was claimed that there is a simple mapping between the semantic roles played by arguments to a verb (intuitively, who did what to whom) and the deep structure grammatical relations (subject, object, etc.). In the final tree of the derivation (the surface structure), the words and phrases were arranged as the sentence would actually be pronounced. On this theory, then, transformations were thought to be the primary link between sound and meaning in natural language.

The Standard Theory had great intuitive appeal and attracted much attention from neighboring disciplines. In particular, many philosophers were attracted by the idea that deep structures might provide something very much like the 'logical form' of sentences needed for precise analysis of their role in inference. Likewise, psychologists hoped that the transformational derivations were a first approximation to the mental processes involved in the production and comprehension of utterances. Initial experiments gave credibility to this idea, in that they showed a correlation between the psychological complexity of a sentence and the number of transformations posited in its derivation. Further research on this idea (usually referred to as the 'derivational theory of complexity') failed to support it, however, and by the early 1970s it had been largely abandoned (Fodor et al. 1974).

Most contemporary grammatical theories have preserved the most important innovations of the Standard Theory, namely, syntactic features, recursive phrase structure, and some sort of semantic component. On the other hand, no current theory maintains the centrality of transformations in mediating between sound and meaning.

The first major challenge to Chomsky's views within the generative paradigm was a movement known as 'Generative Semantics'; its leading figures included George Lakoff, James McCawley, Paul Postal, and John ('Haj') Ross. They carried the central idea of the Standard Theory to its logical conclusion, claiming that deep structures should themselves be viewed as representations of meaning, and denying that syntactic and semantic rules should be considered distinct components of a grammar. That is, on the Generative Semantics view, something was considered a possible input to the transformational rules just in case it represented a proposition that made sense. Hence all languages could be derived from the same underlying source, differing only in how the underlying representations get transformed into sounds.

The underlying trees of Generative Semantics were far larger and more elaborate than those of the Standard Theory (though the inventory of grammatical categories was much reduced). Virtually all the work involved in describing the relationships between form and meaning in language was done in this theory by transformations, though these rules were rarely formulated explicitly.

Generative Semantics enjoyed wide currency for a few years and served as the vehicle for the exploration of a wide range of fascinating phenomena in many languages. Although the theory itself had a short life span (for reasons that have been debated by historians of linguistics[6]), many of the constructions first discovered by generative semanticists continue to figure prominently in theoretical discussions. Moreover, some recent analyses have borne striking resemblances to earlier Generative Semantics proposals, as has often been observed.

Unlike the generative semanticists, Chomsky and some others (notably, Ray Jack-

[6]See Newmeyer 1986, Harris 1993, and Huck and Goldsmith 1995.

endoff) quickly abandoned the idea that pairs of sentences with identical deep structures must be synonymous. In particular, they argued that transformations that reordered quantified NPs could change the scopes of the quantifiers (e.g. *Many people read few books* was claimed to have a range of interpretations different from *Few books are read by many people*). Hence they claimed that structures other than deep structures must play a role in semantic interpretation.

Instead of the complex underlying trees and elaborate transformational derivations of Generative Semantics, the framework that Chomsky dubbed the 'Extended Standard Theory' (EST) posited a relatively impoverished theory of transformations; instead, it enriched other components of the theory to carry much of the descriptive burden. In addition to the new types of semantic rules alluded to above, schematization over phrase structure rules and an enriched conception of the lexicon – including lexical rules – were introduced. These innovations have been carried over into much contemporary work, including the theory developed in this text. The approach of EST led to a highly 'modular' theory of grammar, with a variety of distinct types of mechanisms to account for different kinds of empirical phenomena.

EST also saw the introduction of 'empty categories' – that is, elements that occupy positions in a tree but which have no phonetic realization. These included a type of null pronoun used in control constructions (e.g. the subject of *leave* in *We tried to leave*) and 'traces' of elements that have been moved.

A central concern of EST and much subsequent work has been to constrain the power of the theory – that is, to restrict the class of grammars that the theory makes available. The primary rationale for seeking such constraints has been to account for the possibility of language acquisition, which (as noted in Chapters 1 and 9) Chomsky regards as the central question for linguistics.

EST was superseded by what came to be known as Government and Binding Theory, or GB. GB was first laid out in Chomsky's (1981) book, *Lectures on Government and Binding*.[7] It develops the modular style of EST, dividing the theory of grammar into a set of subtheories, each with its own set of principles, assumed to be universal. Although GB still used transformational derivations to analyze sentences, it reduced the transformational component to a single rule (referred to as 'Move α'), which could move anything anywhere. The idea was that general principles would filter out most derivations, preventing the massive overgeneration that would otherwise result from such an underconstrained transformational operation.

Elaborating on earlier work in EST, GB analyses posited a taxonomy of empty categories. Binding Theory, which was a major topic of research within GB, was applied not only to overt pronouns, but also to empty categories. Movement was formulated so as to leave behind traces (a kind of empty category), which were bound by the moved element. Binding Theory thus attempted to find relations between constraints on movement and constraints on possible pronoun-antecedent relations. Since movement was used to deal with a wide range of phenomena (including filler-gap dependencies, the active-passive relation, raising, extraposition, and auxiliary inversion), linking all of these to the binding principles yielded a highly interconnected system.

[7]For an introductory presentation, see Haegeman 1994.

The primary focus in GB and subsequent transformational research has been the postulation of a theory of universal grammar. GB claimed that many of the principles that make up the theory of grammar are parameterized, in the sense that they vary within a narrow range.[8] Learning a language, on this view, consists of fixing a small set of parameters (plus learning vocabulary). That is, GB claimed that all languages are essentially alike, with only a few restricted parameters of possible variation.

Many linguists since the early 1980s have framed their grammatical studies in terms of this framework, yielding a large literature that represents a analyses of a much wider range of languages and phenomena than any of the other theories listed here. But the analyses developed within GB are often inconsistent with one another. In addition, these analyses are seldom formulated with a precision comparable to that assumed in this text. For these reasons (and the further absence of any theory of what could count as a possible 'parameter'[9]), particular GB analyses and the general claims about cross-linguistic parametric variation are often quite difficult to evaluate. Nonetheless, it is clear that GB analyses tend to share certain noteworthy characteristics, including the following:

- Highly articulated phrase structures (linguistically significant distinctions and relations are encoded into tree configurations);
- Use of movement (that is, the transformation 'Move α');
- Extensive use of empty categories;
- A rich set of universal principles, some of which are parameterized;
- Avoidance of language-particular rules (properties specific to a language are to be expressed in terms of values of universally available parameters);
- Deductive structure (small changes in a grammar should have far-reaching consequences for the language, so that stipulation is minimized).

The theory we have presented here has been influenced by GB in a number of ways. These include very general goals, such as striving for a theory whose components interact deductively. They also include more specific design features, such as the general form of the Binding Theory (though not the detailed statement of the binding principles). Finally, there are specific points of our grammar that were first proposed within GB analyses, such as treating complementizers as heads that could take sentences as their complements.

As of this writing, the most recent incarnation of transformational grammar is the 'Minimalist Program'. As its name implies, MP is a program for research, rather than a theory of syntax. Further, it is a program that seeks to discover whether natural language grammar (and in particular syntax) is an 'optimal' system in the sense of requiring the fewest theoretical constructs. The general tactic is to lay out what appears to be an optimally simple theory of syntax and then test whether it can account for the data of natural languages.

The model that the MP is currently exploring is one in which the only information associated with formatives is information relevant to either the pronunciation (that is, the level of Phonological Form) or the meaning (that is, the level of Logical Form). A

[8]Another name for this approach to syntax is 'Principles and Parameters'.

[9]For a recent discussion of this issue, see Baker 2001 and especially Trask 2002.

small set of syntactic operations combine the formatives into syntactic structures and transform those structures along two paths, creating a Phonological Form and a Logical Form, which are meant to be interpreted by other cognitive systems. Constraints posited in the theory may refer to these end levels, but not any intermediate stages. In addition, competing derivations based on the same set of formatives can be compared and evaluated with respect to some economy metrics. The derivations that are preferred by the economy metrics yield grammatical strings, those that are dispreferred yield ungrammatical strings. More recent work has sought to localize such effects, applying economy constraints at each step of a transformational derivation. This conception of grammar, in which the properties of competing transformational derivations are crucial in determining sentence well-formedness,[10] represents a radical departure from the original goals and methodology of generative grammar and has no direct connection with the theoretical orientation of the present work.

The seminal work on MP is Chomsky 1995. For an elementary exposition of MP, see Radford 1997; also useful are Webelhuth 1995 and Epstein and Seely 2002. For an extensive comparison of MP with CBL approaches in general, see Johnson and Lappin 1999. For a critical perspective on recent Minimalist work, see Bender 2002. For a sociologically intriguing discussion of the evolution of MP from GB, see Lappin et al. 2000a, 2000b, 2001.

B.2 Constraint-Based Lexicalist Grammar

Some CBL theories of grammar were developed by transformational grammarians who began to question the reasons that had been given for positing transformations. Others have a rather different genealogy, growing out of work on the artificial languages of logic and computer science. Combining the precision and rigor of the latter tradition with the attention to empirical generalizations that dominates the former has been extremely fruitful, though by no means easy.

The following subsections provide very brief sketches of some of the most important CBL theories of grammar.[11]

B.2.1 Categorial Grammar (1974 to the present)

Categorial Grammar (CG) has a long history dating back to the 1930s, but it was developed primarily by mathematical logicians before the early 1970s. CG first received widespread attention from linguists when the logician Richard Montague used it as the syntactic framework to go with his new approach to analyzing natural language semantics.[12] Contemporary work on CG maintains Montague's strong commitment to semantic

[10]Generative semanticists in the early 1970s briefly discussed the need for what were then termed 'transderivational constraints', but the idea was not pursued for long.

[11]The dates given in parentheses are roughly the periods during which a substantial number of researchers have been or are still active in developing the theory.

[12]Montague's intensional logic and his precise framework for studying the relation between expressions and their meaning had a considerable influence on work in linguistics. His famous remark (Montague 1970:373) that 'there is in my opinion no important theoretical difference between natural languages and the artificial languages of logicians' is sometimes referred to (following Bach 1989) as 'Montague's Hypothesis'.

compositionality – that is, to the idea of building up the structures and interpretations of sentences at the same time, using coordinated syntactic and semantic rules (the 'Rule-to-Rule Hypothesis').

The central idea of CG is that the combinatory properties of grammatical categories can be directly encoded in their forms. A CG has a small number of basic categories, and a large (potentially infinite) class of categories defined in terms of how they combine with other categories. For example, we might designate S and NP as basic and define a category of expressions that combine with an NP on their left to form an S, which can be annotated NP\S. This category would be equivalent to our VP (including those consisting simply of an intransitive verb). Another category, annotated (NP\S)/NP, would be used for expressions that combine with an NP on their right to form an NP\S; this corresponds roughly to our category of strictly transitive verbs, but might also be used for more complex expressions like *let stand*, in sentences like *The decision let stand the lower court's ruling*. Because categories are defined in terms of their members' potential for combining with other constituents, CG is often seen as a variety of Dependency Grammar (q.v.).

The simplest forms of CG only combine categories by simple concatenation, with concomitant elimination of a 'slash' in the combined category. For example, two expressions of categories NP and NP\S can be concatenated to form a new expression of category S. Such a CG is provably equivalent to a simple context-free phrase-structure grammar. Richer versions of CG include other rules of category combination, for example allowing two expressions of categories A/B and B/C to combine to create an expression of category A/C (so-called 'composition'). They may also allow what is called 'type raising', for example, reanalyzing an expression of category A as belonging to a more complex category like B/(B/A) or B/(A\B). Such enrichments of the basic machinery of CG have made possible appealing analyses of a range of syntactic, semantic, and phonological phenomena in natural languages.

CG divides into two major schools of thought, known as 'type logical grammar' and 'combinatory categorial grammar'. The former treats the combination of categories as a system of logic, and the derivations of sentences are treated as a kind of mathematical proof. The latter focuses more on linguistic questions, positing operations that are motivated by natural language phenomena, even if they seem unnatural as logical operations.

The primary attractions of CG have always been its conceptual simplicity and the fact that it is well suited to the formulation of tightly linked syntactic and semantic analyses. It also provides particularly elegant and appealing accounts of coordination. In a number of works (notably Steedman 1996), Steedman has argued that CG is well-suited to explaining a range of facts about English prosody (that is, pitch accent and intonation) and about how people process sentences.

One characteristic of all but the simplest versions of CG has been variously cited as a strength and a weakness: CG typically allows many more different analyses of any given string than other theories of grammar. For example, in a theory with composition and type-raising, a simple transitive sentence like *Pat likes oatmeal* can be derived either by combining the verb with its object first, and then with its subject, or by combining the verb with its subject first, and then with its object. With longer sentences, the number

of possible derivations grows extremely rapidly.

More information about CG is available online at: http://www.cs.man.ac.uk/ai/CG. For an introduction to categorial grammars, see Wood 1993 and Carpenter 1997. A brief overview is available online at http://cognet.mit.edu/MITECS/Entry/steedman. Steedman (1996, 2000) summarizes the combinatory categorial grammar tradition. For the competing 'type-logical' approach, see Morrill 1994 and Moortgat 1997.

The influence of CG on the theory developed in this text is quite clear. The valence features of HPSG do much the same work as complex categories do in CG. The nodes of trees in HPSG are labeled with feature structures that contain all the information in CG categories (plus quite a bit more). Our grammar rules cancel elements off our valence lists in a way analogous to the combinatory rules of CG. Hence many CG analyses can be translated fairly straightforwardly into HPSG.

B.2.2 Construction Grammar (1988 to the present)

Construction Grammar (CxG) is a label applied to a family of theories that take grammatical constructions as essential units out of which sentences are built. Constructions, on this view, encompass not only (descriptions of) words and phrases, but also idioms and other types of collocations, characterizable at various levels of generality. An important leading idea of work in CxG is that there is no important difference between 'core' and 'peripheral' constructions.[13] Construction grammarians argue that, by contrast, grammatical theory can and should include numerous principles whose domain of application includes both kinds of construction, no matter how that distinction is made. A common misconception about CxG is that it is just a theory of 'marginal', 'peripheral' constructions. To the contrary, one of its fundamental claims is that linguistic rules (constraints) operate at diverse levels of generality, without there being any theoretical difference (other than the scope of their application) between construction-specific constraints, constraints of intermediate grain (e.g. constraints on all headed constructions) and the most general principles of UG.

CxG researchers countenance neither empty categories nor transformations, two formal devices central to transformational work, finding them empirically unmotivated and descriptively unnecessary. As in HPSG (q.v.), CxG is based on the notion of constraint satisfaction and constraint inheritance is used to express generalizations that cut across diverse kinds of objects that pattern into clusters obeying a 'family resemblance'. Chapter 16 of this book outlines a formal version of CxG. As is clear from the discussion there, the central features of the theory developed in our first fifteen chapters translate naturally into a constructional framework. The framework outlined there is closest to that of Fillmore et al. (forthcoming). There are many other approaches that fall under the CxG umbrella, however, some of which eschew formal approaches to grammar. ,

There are two principal substantive motivations for CxG. The first is to account for grammatically determined, non-truth-conditional aspects of meaning – including such 'pragmatic' factors as conventional implicature and presupposition – in the formal rep-

[13]This is a distinction introduced (Chomsky 1986b) to distinguish constructions subject to UG from language-particular idiosyncrasies. Unfortunately, no criteria have (to our knowledge) ever been offered to distinguish between these two sorts of constructions, rendering the empirical content of the 'core-periphery' distinction, once again, quite difficult to evaluate.

resentations of sentences and of the grammatical constructions that license them. The second is to account for the full range of idiomatic and semi-idiomatic constructions, phenomena that are pervasive in the world's languages. Although every generative approach to the grammar of natural languages is committed in principle to full coverage of the facts of all languages, as well as in the extraction of intralanguage and interlanguage generalizations (the latter usually considered to constitute the stuff of universal grammar), varying approaches differ in their relative emphasis on the full coverage of language facts versus the development of a parsimonious theory of universal grammar. CxG falls at the end of this scale, emphasizing its concern for empirical coverage and the need base accounts of UG on well worked-out empirical descriptions.

Kay 1995 is a brief, acessible overview of the basics of CxG. For online information, see the official website at: http://www.constructiongrammar.org/; there is further information at: http://www.icsi.berkeley.edu/~kay/bcg/ConGram.html. A version of CxG showing the influence of approaches to grammar that refer to themselves as 'cognitive' (e.g. Lakoff 1987, Langacker 1987) is developed by Goldberg (1995). The original, generative, view of CxG (Fillmore et al. 1988) is further developed in recent work by Michaelis and Lambrecht (1996), Kay and Fillmore (1999), Kay (2002), and Fillmore et al. (forthcoming).

B.2.3 Dependency Grammar (1959 to the present)

Work on transformational grammar rests on two crucial (but controversial) assumptions about sentence structure: that it is organized hierarchically into 'phrases' (hence 'phrase structure'), and that grammatical relations such as 'subject' and 'object' are derivative, to be defined in terms of phrase structure configurations. The assumption of phrase structure is a distinctively American contribution to linguistics, having been suggested by Bloomfield (1933). Bloomfield suggested that sentences should be analyzed by a process of segmentation and classification: segment the sentence into its main parts, classify these parts, then repeat the process for each part, and so on until the parts are 'morphemes', the indivisible atoms of grammar. Thus *Cool students write short essays* divides into the noun phrase *cool students* plus the verb phrase *write short essays*, which in turn divides into the verb *write* plus *short essays*, and so on. This contrasts with the European tradition (which dates back to classical Greece) in which the focus is on individual words and their relationships - for example, *cool* is an 'attributive modifier' of *students*, and *students* is the subject of *write*.

The attraction of phrase structure analysis is its formal clarity, which is revealed by the familiar phrase structure trees. Various linguists (mainly European) have attempted to develop the traditional approach in a similarly formal manner, with the emphasis on the relationships among words rather than on the groupings of words. One of the characteristics of these relationships is that the words concerned are generally not equal, in that one serves to modify the meaning of the other; so *cool students* denote certain students, and *students writing essays* denotes a kind of writing. The relationships are called 'dependencies', with the modifying word depending on the modified (so *cool* depends on *students*, and *students* on *write*), and the approach is called 'Dependency Grammar' (DG) to contrast it with phrase structure grammar.

There are several ways to represent DG analyses diagrammatically, including a system that has been widely used in American schools since the nineteenth century which is often called simply 'sentence diagramming'. The first real attempt to build a theory of DG analysis was Tesnière 1959, but since then developments in the tradition of PSG[14] have been paralleled in DG theories. One of these which is particularly close in other respects to HPSG is 'Word Grammar' (Hudson 1984, 1990, 1998). Online information about DG is available at: http://ufal.mff.cuni.cz/dg/dg.html.

In some respects, HPSG bridges the gap between DG and PSG, for in HPSG all the parts of a phrase depend directly on its head word – phrases are 'head-driven', just as in DG. On the other hand, in HPSG the dependent parts are themselves phrases with their own internal structure consisting of a head word and its dependents.

B.2.4 Generalized Phrase Structure Grammar (1979–1987)

Generalized Phrase Structure Grammar, (known as GPSG) was initiated by Gerald Gazdar in a pair of papers (Gazdar 1981, 1982) that attracted the attention of numerous researchers in the field of syntax. The theory was further developed by him and a number of colleagues in the early 1980s and was codified in the 1985 book, *Generalized Phrase Structure Grammar* (Gazdar et al. 1985), which provides a detailed exposition of the theory.

The central idea of GPSG is that standard context-free phrase structure grammars can be enhanced in ways that do not enrich their generative capacity, but which do make them suitable for the description of natural language syntax. The implicit claim of work in GPSG was that the tenable arguments against CFG as a theory of syntax were arguments about efficiency or elegance of notation, and not about coverage in principle.

Among the important ideas that originated in GPSG are the separation of CFG rules into (i) rules of immediate dominance ('ID rules'), which specify only which phrases can appear as daughters in a local syntactic tree, and (ii) rules of linear precedence ('LP rules'), which specify general constraints determining the order of daughters in any local tree. This factorization of the two functions of traditional CFG rules is preserved in HPSG, though we have not employed it in the formulation of grammar rules in this text.

A second idea stemming from work in GPSG is the treatment of long-distance dependency constructions, including filler-gap constructions (such as topicalization, *wh*-questions, and relative clauses). The GPSG treatment of these phenomena involved locally encoding the absence of a given constituent via a feature specification. The remarkable result of the transformationless GPSG analysis of filler-gap dependencies was that it succeeded where transformational theories had failed, namely in deriving the Coordinate Structure Constraint and its 'across-the-board' exceptions (see Chapter 14). This feature-based analysis of filler-gap dependencies is preserved in HPSG, and we have carried it over virtually intact to the current text.

[14] We use the term 'PSG' to refer to research developing extensions and generalizations of Context-Free Phrase Structure Grammars, e.g. GPSG (q.v.) and HPSG (q.v.).

B.2.5 Head-Driven Phrase Structure Grammar (1984 to the present)

HPSG evolved directly from attempts to modify GPSG in the interdisciplinary environment of Stanford's Center for the Study of Language and Information (CSLI), the site of several experimental computational systems for language processing. From its inception, HPSG has been developed as a conscious effort to synthesize ideas from a variety of perspectives, including those of Situation Semantics (which originated at CSLI at about the same time as HPSG), data type theory, and a variety of other linguistic frameworks under development in the early and mid-1980s. The name 'Head-driven Phrase Structure Grammar' was chosen to reflect the increasingly recognized importance (as compared with, say, GPSG) of information encoded in the lexical heads of syntactic phrases. Dependency relations are lexically encoded, as they are in Dependency Grammar (q.v.), Categorial Grammar (q.v.) and LFG (q.v.). The theoretical aspects of HPSG have been developed in considerable detail in three books (Pollard and Sag 1987, Pollard and Sag 1994, and Ginzburg and Sag 2000) and a number of major articles.

Some of the key ideas of work in HPSG are: (1) a sign-based architecture (see Chapter 16); (2) the organization of linguistic information via types, type hierarchies, and constraint inheritance; (3) the projection of phrases via general principles from rich lexical information; (4) the organization of such lexical information via a system of lexical types; and (5) locality of selection, agreement, case assignment and semantic role assignment, as guaranteed by the organization of feature structures, e.g. into *synsem* objects that appear on ARG-ST (or SUBCAT) lists, and (6) the factorization of phrasal properties into construction-specific and more general constraints. These properties have all been discussed at various places in this text.

Since the inception of HPSG, researchers have been involved with its computational implementations. From 1980 until 1991, Hewlett-Packard Laboratories in Palo Alto, California supported one such project, which involved the first two authors of this text and a number of colleagues and students. It was with this project that many of us learned for the first time how far the rhetoric of theoretical linguistics can be from the reality of working grammars. At the time of this writing, implementations of HPSG and HPSG-like grammars are being developed at numerous universities and industrial research laboratories around the world, including sites in North America, Western and Eastern Europe, Japan, Korea, Taiwan, and Australia. The LinGO initiative, an ongoing collaboration that includes partners in the US, Europe, and Asia, makes available (without cost) a number of open-source, HPSG-related, computational resources. These resources, which include grammars, lexicons, and the LKB Grammar Engineering Platform,[15] are available online at: http://lingo.stanford.edu. Another HPSG-related implementation effort, one that includes collaborators from Europe, the US, and Canada, is the TRALE system (and its predecessor – the ConTroll-System, developed at the linguistics department at the University of Tübingen), which is described at http://www.sfs.uni-tuebingen.de/hpsg/sysen.html. The TRALE system is an extension of the Attribute-Logic Engine (ALE) system, which can be freely downloaded from http://www.cs.toronto.edu/~gpenn/ale.html. For general information about HPSG, see http://hpsg.stanford.edu and http://www.ling.ohio-state.edu/research/hpsg/.

[15] For an accessible introduction to the LKB system and related issues, see Copestake 2002.

B.2.6 Lexical Functional Grammar (1979 to the present)

The theory of Lexical Functional Grammar, commonly referred to as 'LFG', shares with Relational Grammar (q.v.) the idea that relational concepts like 'subject' are of central importance and cannot be defined in terms of tree structures. But it also treats phrase structure as an essential part of grammatical description and has focussed on the development of a universal theory of how constituent structures are associated with grammatical relations.

In LFG, each phrase is associated with multiple structures of distinct types, with each structure expressing a different sort of information about the phrase. The two representations that have been the center of attention in most LFG literature are: the 'functional structure', which expresses the relational information that is analogous in certain respects to our ARG-ST and in other respects to the valence features SPR and COMPS; and the 'constituent structure', which is a tree diagram very similar to the surface structures of the Standard Theory. General principles and construction-specific constraints define the possible pairings of functional and constituent structures. In addition, LFG recognizes a number of further levels of representation. Perhaps most notable among these are σ-structure, which represents linguistically significant aspects of meaning, and a-structure, which serves to link syntactic arguments with aspects of their meanings. Thus the analogue of the HPSG sign presented in Chapter 16 is a tuple of LFG structures, possibly the four-tuple consisting of a sentence's c-structure, f-structure, a-structure, and σ-structure.

There are no transformations in LFG. Much of the descriptive work done in earlier theories by transformations is handled by an enriched lexicon, an idea pioneered by LFG researchers. For example, the active-passive relation in English is treated as a lexical relation between two forms of verbs. In early LFG, this was codified in terms of lexical rules similar to those presented in this text. Subsequent work has sought to develop a more abstract conception of lexical relations in terms of 'lexical mapping theory'. LMT provides for constraints on the relation between f-structures and a-structures, that is, constraints associated with particular arguments that partially determine their grammatical function. It also contains mechanisms whereby arguments can be suppressed in the course of lexical derivation. In LFG, information from lexical entries and phrasal annotations is unified to produce the functional structures of complex expressions.

As the above description makes clear, LFG and HPSG bear many resemblances. In particular, HPSG has been able to incorporate many insights from work in LFG, most notably: the significant use of dependency-based analyses of numerous phenomena (rather than accounts based on constituent structure) and the general constraint-based approach to grammatical description. There are crucial differences between LFG and HPSG as well, e.g. the use of types and type-based constraint inheritance, which plays no role in the LFG literature. Similarly, the idea of cancelling arguments from valence lists (which stems from Categorial Grammar (q.v.)) is absent from LFG; instead, dependencies are managed in terms of the notions of 'completeness' and 'coherence', which ensure that functional structures contain sufficient and consistent information. LFG's constraining equations provide a way of handling certain kinds of grammatical requirements that has no analog in HPSG. Conversely, the defeasible constraints that are found in some versions of HPSG have no direct counterpart in LFG.

The differences in practice between the HPSG and LFG communities can lead to rather different analyses of the same phenomena; yet these analyses are often compatible with either framework. For an overview of current developments in LFG, see Dalrymple et al. 1995, Bresnan 2001, Dalrymple 2001, and the LFG website at http://www-lfg.stanford.edu/lfg/. The Parallel Grammar Project, based at the Palo Alto Research Center (PARC), is an ongoing collaboration that has developed a variety of computational resources for NLP based on LFG. Information about the ParGram Project, which includes researchers at a number of sites in the US, Europe and Japan, can be found online at http://www2.parc.com/istl/groups/nltt/pargram/.

B.3 Three Other Grammatical Frameworks

B.3.1 Relational Grammar (1974 to the present)

We can now return to the second controversial claim that transformational grammar made about sentence structure, namely, that grammatical relations are derivative. In early theories of generative grammar, transformations were defined in terms of structural properties of tree diagrams. To the extent that traditional notions like 'subject' and 'direct object' were employed in these theories, they were regarded simply as shorthand for relations between linguistic elements definable in terms of the geometry of trees. Relational Grammar (RG), developed by Paul Postal, David Perlmutter, David Johnson and others, adopts primitives that are conceptually very close to the traditional relational notions of subject, direct object, and indirect object. In this respect there is a strong affinity between RG and Dependency Grammar (q.v.). The grammatical rules of RG are formulated in relational terms, replacing the earlier formulations based on tree configurations. For example, the passive rule is stated in terms of promoting the direct object to subject, rather than as a structural rearrangement of NPs.

This approach allows rules to be given very general formulations that apply across languages. The characterization of passivization as promotion of the object does not depend on whether subjecthood and objecthood are indicated by word order or by other means, such as case marking on the nouns or some marking on the verb.

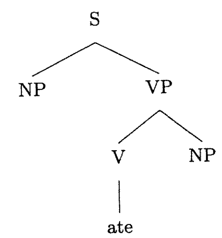
Although the influence of RG on the theory presented here may not be obvious, it is real. The notions of 'specifier' and 'complement' employed in this text are generalizations of 'subject' and 'object'. Languages use different grammatical devices to mark these relations (word order, case marking, agreement, etc.), and so a theory whose primitives are too closely linked to these devices would be unable to express cross-linguistic similarities. A number of contemporary theories, including LFG (q.v.) and HPSG (q.v.), have adopted this central insight of RG.

The RG framework has been applied to the description of a much wider variety of languages than were earlier generative theories (which tended to concentrate on the familiar European languages, East Asian languages, and a few others). For a brief online overview of the leading ideas of RG, see: http://cognet.mit.edu/MITECS/Entry/aissen. Various results of work in this framework are anthologized in Perlmutter 1983 and Postal and Joseph 1990. 'Arc Pair Grammar' (Johnson and Postal 1980) is an axiomatization and elaboration of many of the central ideas of Relational Grammar. Arc Pair Grammar was the first formalized constraint-based theory of grammar to be developed in the generative

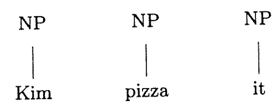
tradition. Although few details of the theory have been carried over into contemporary CBL theories, it pioneered many important innovations.[16]

B.3.2 Tree-Adjoining Grammar (1975 to the present)

Where other theories of grammar have posited representations of grammatical dependencies instead of or in addition to tree structure, Tree-Adjoining Grammar (TAG) (developed by Aravind Joshi, Leon S. Levy, Masako Takahashi, K. Vijay-Shanker, David Weir, and others, beginning in the early 1970s) has pursued the possibility of making trees into better representations of grammatical dependencies. The building blocks of a TAG are not rules and lexical entries, but elementary trees, anchored in lexical entries. In order to use trees to represent dependencies, TAG introduces an 'extended domain of locality'. In CFGs, the domain of locality is just one level in the tree, as each grammar rule can only describe a subtree of depth one. Trees of depth one can only represent some grammatical dependencies – e.g. between a verb and its object, but not between a verb and its subject. The elementary trees of a TAG are larger, such as this tree representing the verb *ate* and its dependents:[17]



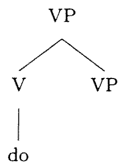
By the operation of 'substitution', the NP nodes in the elementary tree for *ate* could be replaced with any of the following trees:



With this extended domain of locality, TAG elementary trees can represent valence dependencies, but only if all arguments are realized locally. In order to capture non-local realization of dependents (e.g. raising, long-distance dependencies), TAG introduces another operation for combining elementary trees, called 'adjunction'. In adjunction, one elementary tree is spliced into another. Adjunction is defined so as to be possible only if the tree to be spliced in has a node on its frontier that is of the same category as its mother. An example of such a tree is the elementary tree for the raising verb *do*:

[16] For example, the graph-theoretic foundations that Johnson and Postal developed for Arc Pair Grammar are essentially the same as those employed in much work in HPSG (although the definition of feature structures as functions adopted in this book is a slightly different conception).

[17] Notice that the information in this tree is almost identical to that found in the ARG-ST (or SUBCAT) lists in HPSG analyses. The extension of the domain of locality to include subjects has been part of lexicalist theories since the invention of Categorial Grammar (q.v.) in the 1930s.



The operation of adjunction allows TAG to 'factor recursion from the domain of dependencies', while still using trees to represent both.

This presentation of the core ideas of TAG has been informal, but the formal properties of TAG and a number of its variants have been extremely well-studied. Indeed, research on TAG, its variants, and its relationship to other formalisms constitutes an important area of work in formal language theory. Further, there are broad-coverage implemented grammars developed in TAG, including the XTAG grammar (The XTAG Research Group 2000) and some implemented HPSG grammars have been compiled into TAGs (Kasper et al. 1995). Online information about the XTAG Project is available at http://www.cis.upenn.edu/~xtag/home.html. There is also work on modeling psycholinguistic results (including processing complexity) using TAG, including Joshi 1990 and Joshi et al. 2000. For an overview of TAG, see Joshi 2003. Other important work in the literature on TAG includes Joshi et al. 1975, Vijay-Shanker 1987, Weir 1987, and Abeillé and Rambow 2000.

B.3.3 Optimality Theory (1993 to the present)

Optimality Theory, or OT, was first developed as a phonological framework (Prince and Smolensky 1993), and has recently been adapted to syntactic analysis (see, for example, Barbosa et al. 1998, Bresnan 2000, Legendre et al. 2001, and Sells 2001). For an overview, see Kager 1999, Archangeli and Langendoen 1997, and the papers collected at the Rutgers Optimality Archive (http://roa.rutgers.edu/).

OT posits a universal set of defeasible constraints. The grammar of a language consists of a ranking of the constraints. Determining whether a given string of words is a well-formed sentence involves comparing it with other candidate expressions of the same proposition. The candidate whose highest-ranking constraint violation is lower than that of any other candidate is grammatical. For example, if constraint A outranks constraint B, which outranks constraint C, and if candidate sentence 1 violates A whereas candidate sentence 2 violates B and C, then sentence 2 is preferred over sentence 1, and sentence 1 is ungrammatical. If no other candidate sentence wins such a competition against sentence 2, then sentence 2 is licensed by the grammar.

The idea of constraints that can be violated is also incorporated in the theory presented in this book, since defeasible constraints specified in type hierarchies can be overridden. Moreover, a hierarchy of types with defeasible constraints defines a partial ordering on those constraints, with those introduced lower in the hierarchy taking precedence over those introduced at higher levels. Although there are substantive differences, certain central properties of OT can also be found in inheritance hierarchies with defeasible constraints.[18]

[18] Such hierarchies are explored in some detail in the artificial intelligence literature of the 1970s.

OT follows much earlier work in generative grammar in positing rich systems of universal grammar. However, the idea that determinations of well-formedness necessarily involve comparing structures or derivations is a break with past views, as we already noted in discussing the Minimalist Program (q.v.). Another common characteristic of MP and OT is the use of defeasible constraints. As noted above, such constraint mechanisms of various sorts have been proposed from time to time within some theories, including the theory presented in this book. This is not surprising, since idiosyncratic exceptions to general patterns are commonplace in natural languages. Defeasible constraint mechanisms are now employed fairly widely in various theories of syntax (though they are still controversial in certain circles). It remains to be seen whether a similar consensus will arise concerning the idea of defining well-formedness in terms of the outcome of some sort of competition.

B.4 Summary

In this appendix, we have surveyed a number of approaches to grammatical theory that have some relation to the analyses presented in this book. Our survey has been all too brief and has doubtless left out some approaches that certain readers will feel should have been included. Moreover, one thing we can say with certainty about the field of linguistics, at least over the last half century, is that theories of grammar have come and gone quite quickly. And this is likely to continue until the field evolves to a point where the convergent results of diverse kinds of psycholinguistic experiments and computational modeling converge with, and are generally taken to have direct bearing on, the construction of analytic hypotheses. Until that day, any survey of this sort is bound to be both incomplete and rapidly obsolescent.

For comparisons of OT with models that employ inheritance hierarchies with defeasible constraints, see Asudeh in press and Malouf in press.

Syntactic theory: A formal introduction (2nd. ed)

Ivan A. Sag, Thomas Wasow, Emily M. Bender.
Stanford, CA: CSLI Publications. 2003.

References

Abeillé, Anne, and Owen Rambow (Eds.). 2000. *Tree-Adjoining Grammars: Formalisms, Linguistic Analysis and Processing.* Stanford, CA: CSLI Publications.

Akmajian, Adrian, Susan Steele, and Thomas Wasow. 1979. The category AUX in Universal Grammar. *Linguistic Inquiry* 10:1–64.

Akmajian, Adrian, and Thomas Wasow. 1974. The constituent structure of VP and AUX and the position of the verb BE. *Linguistic Analysis* 1:205–245.

Andrews, Avery. 1982. The representation of case in modern Icelandic. In *The Mental Representation of Grammatical Relations* (Bresnan 1982b).

Archangeli, Diana, and D. Terence Langendoen (Eds.). 1997. *Optimality Theory: An overview.* Oxford: Blackwell Publishing.

Arnold, Jennifer, Maria Fagnano, and Michael K. Tanenhaus. 2002. Disfluencies signal, theee, um, new information. *Journal of Psycholinguistic Research.*

Asudeh, Ash. in press. A licensing theory for Finnish. In D. C. Nelson and S. Manninen (Eds.), *Generative Approaches to Finnic and Saami Linguistics.* Stanford: CSLI Publications.

Bach, Emmon. 1979. Control in Montague Grammar. *Linguistic Inquiry* 10:515–31.

Bach, Emmon. 1989. *Informal Lectures on Formal Semantics.* Albany: SUNY Press.

Baker, Mark. 2001. *The Atoms of Language: The Mind's Hidden Rules of Grammar.* New York: Basic Books.

Bar-Hillel, Yehoshua, and Eliyahu Shamir. 1960. Finite-state languages: Formal representations and adequacy problems. *The Bulletin of the Research Council of Israel* 8F:155–166. Reprinted in Y. Bar-Hillel (1964) Language and Information: Selected Essays on Their Theory and Application Reading, MA: Addison-Wesley.

Barbosa, Pilar, Danny Fox, Paul Hagstrom, Martha McGinnis, and David Pesetsky (Eds.). 1998. *Is the Best Good Enough? Optimality and Competition in Syntax.* Cambridge, MA: MIT Press.

Barlow, Michael, and Charles Ferguson (Eds.). 1988. *Agreement in Natural Language: Approaches, Theories, Descriptions.* Stanford.

Barwise, Jon, and John Etchemendy. 1989. Semantics. In *Foundations of Cognitive Science* (Posner 1989).

Bates, Elizabeth, and Brian MacWhinney. 1989. Functionalism and the competition model. In B. MacWhinney and E. Bates (Eds.), *The Cross-linguistic Study of Sentence Processing.* Cambridge: Cambridge University Press.

Baugh, John. 1983. *Black Street Speech: Its History, Structure, and Survival.* Austin: University of Texas Press.

Bear, John. 1981. Gaps as syntactic features. Technical report, Center for Cognitive Science, University of Texas at Austin.

Bender, Emily, and Dan Flickinger. 1999. Peripheral constructions and core phenomena: Agreement in tag questions. In G. Webelhuth, J.-P. Koenig, and A. Kathol (Eds.), *Lexical and Constructional Aspects of Linguistic Explanation*, 199–214. Stanford, CA: CSLI.

Bender, Emily M. 2001. *Syntactic Variation and Linguistic Competence: The Case of AAVE Copula Absence.* PhD thesis, Stanford University.

Bender, Emily M. 2002. Review of R. Martin, D. Michaels and J. Uriagereka (eds.), *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik. Journal of Linguistics* 38:432–439.

Bender, Emily M., Dan Flickinger, and Stephan Oepen. 2002. The grammar matrix: An open-source starter-kit for the rapid development of cross-linguistically consistent broad-coverage precision grammars. In J. Carroll, N. Oostdijk, and R. Sutcliffe (Eds.), *Proceedings of the Workshop on Grammar Engineering and Evaluation at the 19th International Conference on Computational Linguistics*, 8–14.

Bever, Thomas. 1970. The cognitive basis for linguistic structure. In J. R. Hayes (Ed.), *Cognition and the Development of Language.* New York: Wiley.

Bloomfield, Leonard. 1933. *Language.* New York: H. Holt and Company.

Bouma, Gosse, Rob Malouf, and Ivan A. Sag. 2001. Satisfying constraints on extraction and adjunction. *Natural Language and Linguistic Theory* 19:1–65.

Brame, Michael K. 1979. *Essays Toward Realistic Syntax.* Seattle: Noit Amrofer.

Bresnan, Joan. 1978. A realistic transformational grammar. In M. Halle, J. Bresnan, and G. A. Miller (Eds.), *Linguistic Theory and Psychological Reality.* Cambridge, MA: MIT Press.

Bresnan, Joan. 1982a. Control and complementation. In *The Mental Representation of Grammatical Relations* (Bresnan 1982b).

Bresnan, Joan (Ed.). 1982b. *The Mental Representation of Grammatical Relations.* Cambridge, MA: MIT Press.

Bresnan, Joan. 1982c. The passive in lexical theory. In *The Mental Representation of Grammatical Relations* (Bresnan 1982b).

Bresnan, Joan. 1995. Linear order, syntactic rank, and empty categories: On weak crossover. In *Formal Issues in Lexical-Functional Grammar* (Dalrymple et al. 1995).

Bresnan, Joan. 2000. Optimal syntax. In J. Dekkers, F. van der Leeuw, and J. van de Weijer (Eds.), *Optimality Theory: Phonology, Syntax and Acquisition*, 334-385. Oxford: Oxford University Press.

Bresnan, Joan. 2001. *Lexical-Functional Syntax.* Oxford and Cambridge, MA: Blackwell.

Briscoe, Edward, and Ann Copestake. 1999. Lexical rules in constraint-based grammar. *Computational Linguistics* 25(4):487–526.

Briscoe, Edward, Ann Copestake, and Valeria de Paiva (Eds.). 1993. *Inheritance, Defaults, and the Lexicon.* Cambridge: Cambridge University Press.

Brody, Michael. 1995. *Lexico-Logical Form: A Radically Minimalist Theory.* Cambridge, MA: MIT Press.

Burzio, Luigi. 1986. *Italian Syntax.* Dordrecht: Reidel.

Cacciari, Cristina, and Patrizia Tabossi (Eds.). 1993. *Idioms: Processing, Structure, and Interpretation.* Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Cameron, Deborah. 1995. *Verbal Hygiene.* London and New York: Routledge.

Campbell, Lyle. 1985. *The Pipil Language of El Salvador.* New York: Mouton Publishers.

Carpenter, Bob. 1992. *The Logic of Typed Feature Structures: with Applications to Unification Grammars, Logic Programs, and Constraint Resolution.* Cambridge: Cambridge University Press.

Carpenter, Bob. 1997. *Type-Logical Semantics.* Cambridge, MA: MIT Press.

Chierchia, Gennaro, and Sally McConnell-Ginet. 1990. *Meaning and Grammar: An Introduction to Semantics.* Cambridge, MA: MIT Press.

Chomsky, Noam. 1957. *Syntactic Structures.* The Hague: Mouton.

Chomsky, Noam. 1959. Review of B. F. Skinner's *Verbal Behavior. Language* 35:26–58. Reprinted in J. A. Fodor and J. J. Katz (Eds.) (1964), *The Structure of Language: Readings in the Philosophy of Language*, Englewood Cliffs, N.J.: Prentice-Hall.

Chomsky, Noam. 1963. Formal properties of grammars. In R. D. Luce, R. Bush, and E. Galanter (Eds.), *Handbook of Mathematical Psychology*, Vol. Volume II. New York: Wiley.

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax.* Cambridge, MA: MIT Press.

Chomsky, Noam. 1966. *Topics in the Theory of Generative Grammar.* The Hague: Mouton.

Chomsky, Noam. 1970. Remarks on nominalization. In R. A. Jacobs and P. S. Rosenbaum (Eds.), *Readings in English Transformational Grammar.* Waltham, MA: Ginn-Blaisdell.

Chomsky, Noam. 1972. *Language and Mind, enlarged edition.* New York: Harcourt, Brace, Jovanovich.

Chomsky, Noam. 1973. Conditions on transformations. In S. Anderson and P. Kiparsky (Eds.), *A Festschrift for Morris Halle.* New York: Holt, Rinehart and Winston.

Chomsky, Noam. 1975. *The Logical Structure of Linguistic Theory.* Chicago: University of Chicago Press. Written in 1955 and widely circulated in mimeograph form.

Chomsky, Noam. 1977. On wh-movement. In *Formal Syntax* (Culicover et al. 1977).

Chomsky, Noam. 1980. *Rules and Representations.* New York: Columbia University Press.

Chomsky, Noam. 1981. *Lectures on Government and Binding.* Dordrecht: Foris.

Chomsky, Noam. 1986a. *Barriers.* Cambridge, MA: MIT Press.

Chomsky, Noam. 1986b. *Knowledge of Language: Its Nature, Origin, and Use.* New York: Praeger.

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

Chomsky, Noam. 2002. Beyond explanatory adequacy. *MIT Occasional Papers in Linguistics* 20.

Clark, Eve V., and Herbert H. Clark. 1979. When nouns surface as verbs. *Language* 55:767–811.

Clark, Herbert H., and Jean E. Fox Tree. 2002. Using uh and um in spontaneous speaking. *Cognition.*

Clark, Herbert H., and Thomas Wasow. 1998. Repeating words in spontaneous speech. *Cognitive Psychology* 37:201–242.

Copestake, Ann. 1992. *The Representation of Lexical Semantic Information.* PhD thesis, University of Sussex. Published as Cognitive Science Research Paper CSRP 280, 1993.

Copestake, Ann. 2002. *Implementing Typed Feature Structure Grammars.* Stanford: CSLI Publications.

Copestake, Ann, Dan Flickinger, Robert Malouf, Susanne Riehemann, and Ivan A. Sag. 1995. Translation using Minimal Recursion Semantics. In *Proceedings of The Sixth International Conference on Theoretical and Methodological Issues in Machine Translation*, 15–32, Leuven.

Copestake, Ann, Daniel Flickinger, Ivan A. Sag, and Carl Pollard. 1999. Minimal Recursion Semantics: an introduction. Unpublished ms., Stanford University. Available online at: http://www-csli.stanford.edu/~aac/papers/newmrs.ps.

Copestake, Ann, Alex Lascarides, and Daniel Flickinger. 2001. An algebra for semantic construction in constraint-based grammars. In *Proceedings of the 39th Annual Meeting of the Association for Computational Linguistics*, Toulouse, France. Association for Computational Linguistics.

Crain, Steven, and Mark Steedman. 1985. On not being led up the garden path: The use of context by the psychological syntax processor. In D. R. Dowty, L. Karttunen, and A. M. Zwicky (Eds.), *Natural Language Processing.* Cambridge: Cambridge University Press.

Crystal, David. 1985. *A Dictionary of Linguistics and Phonetics.* London: B. Blackwell in association with A. Deutsch.

Culicover, Peter, Adrian Akmajian, and Thomas Wasow (Eds.). 1977. *Formal Syntax.* New York: Academic Press.

Dalrymple, Mary. 1993. *The Syntax of Anaphoric Binding.* Stanford: CSLI Publications.

Dalrymple, Mary. 2001. *Lexical Functional Grammar.* New York: Academic Press. Syntax and Semantics, Volume 34.

Dalrymple, Mary, and Ronald M. Kaplan. 2000. Feature indeterminacy and feature resolution. *Language* 76(4):759–798.

Dalrymple, Mary, Annie Zaenen, John Maxwell III, and Ronald M. Kaplan (Eds.). 1995. *Formal Issues in Lexical-Functional Grammar.* Stanford: CSLI Publications.

Davidson, Donald. 1980. *Essays on Actions and Events.* Oxford: Clarendon Press; New York: Oxford University Press.

Davis, Anthony. 2001. *Linking by Types in the Hierarchical Lexicon.* Stanford: CSLI Publications.

de Swart, Henriëtte. 1998. *Introduction to Natural Language Semantics.* Stanford: CSLI Publications.

Dowty, David, Robert Wall, and Stanley Peters. 1981. *Introduction to Montague Semantics.* Dordrecht: D. Reidel.

Elman, Jeffrey, Elizabeth Bates, Mark Johnson, Annette Karmiloff-Smith, Domenico Parisi, and Kim Plunkett. 1996. *Rethinking Innateness – A Connectionist Perspective on Development.* MIT Press. A Bradford Book.

Emonds, Joseph. 1975. *A Transformational Approach to Syntax.* New York: Academic Press.

Epstein, Samuel, and Daniel Seely (Eds.). 2002. *Derivation and Explanation in the Minimalist Program.* Oxford: Blackwell's.

Everaert, Martin, Erik-Jan Van der Linden, Andre Schenk, and Rob Schreuder (Eds.). 1995. *Idioms: Structural and Psychological Perspectives.* Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Ferguson, Charles. 1971. Absence of copula and the notion of simplicity: A study of normal speech, baby talk, foreigner talk, and pidgins. In D. Hymes (Ed.), *Pidginization and Creolization of Languages.* New York: Cambridge University Press.

Fillmore, Charles J., Paul Kay, Laura Michaelis, and Ivan A. Sag. forthcoming. *Construction Grammar.* Stanford: CSLI Publications.

Fillmore, Charles J., Paul Kay, and Mary Catherine O'Connor. 1988. Regularity and idiomaticity in grammatical constructions: The case of *let alone. Language* 64(3):501–538.

Flickinger, Daniel, Carl Pollard, and Thomas Wasow. 1985. Structure sharing in lexical representation. In *Proceedings of the 23rd Annual Meeting of the Association for Computational Linguistics*, Morristown, N.J. Association for Computational Linguistics.

Fodor, Janet D. 1995. Comprehending sentence structure. In *Language* (Gleitman and Liberman 1995).

Fodor, Jerry A. 1983. *The Modularity of Mind.* Cambridge, MA: MIT Press.

Fodor, Jerry A., Thomas Bever, and Merrill Garrett. 1974. *The Psychology of Language.* New York: McGraw-Hill.

Fraser, Bruce. 1970. Idioms within a transformational grammar. *Foundations of Language* 6:22–42.

Frege, Gottlob. 1892. On sense and reference. *Zeitschrift für Philosophie und philosophische Kritik* 100:25–50. Translation (under the title 'On Sense and Meaning') appears in Geach and Black (1980).

Gamut, L. T. F. 1991. *Logic, Language, and Meaning.* Chicago: University of Chicago Press.

Garrett, Merrill. 1990. Sentence processing. In D. Osherson and H. Lasnik (Eds.), *Language: An Invitation to Cognitive Science*, Vol. Volume 1, first edition. Cambridge, MA: MIT Press.

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

Gazdar, Gerald. 1982. Phrase structure grammar. In P. Jacobson and G. K. Pullum (Eds.), *The Nature of Syntactic Representation.* Dordrecht: Reidel.

Gazdar, Gerald, Ewan Klein, Geoffrey K. Pullum, and Ivan A. Sag. 1985. *Generalized Phrase Structure Grammar.* Cambridge, MA: Harvard University Press and Oxford: Basil Blackwell.

Gazdar, Gerald, and Geoffrey K. Pullum. 1981. Subcategorization, constituent order, and the notion 'head'. In M. Moortgat, H. van der Hulst, and T. Hoekstra (Eds.), *The Scope of Lexical Rules.* Dordrecht: Foris.

Gazdar, Gerald, Geoffrey K. Pullum, and Ivan A. Sag. 1982. Auxiliaries and related phenomena in a restrictive theory of grammar. *Language* 58:591–638.

Geach, Peter, and Max Black (Eds.). 1980. *Translations from the Philosophical Writings of Gottlob Frege, third edition.* Oxford: Basil Blackwell.

Ginzburg, Jonathan, and Ivan A. Sag. 2000. *Interrogative Investigations: The Form, Meaning and Use of English Interrogatives.* Stanford: CSLI Publications.

Gleitman, Lila R., and Mark Liberman (Eds.). 1995. *Language.* Vol. 1 of *An Invitation to Cognitive Science.* Cambridge, Mass: MIT Press. Second edition.

Goldberg, Adele E. 1995. *A Construction Grammar Approach to Argument Structure.* Chicago: University of Chicago Press.

Green, Lisa. 1993. *Topics in African American English: The Verb System Analysis.* PhD thesis, University of Massachusetts at Amherst.

Green, Lisa J. 2002. *African American English: A Linguistic Introduction.* Cambridge: Cambridge University Press.

Greenbaum, Sidney. 1996. *The Oxford English Grammar.* Oxford: Oxford University Press.

Grice, H. Paul. 1989. *Studies in the Way of Words.* Cambridge, MA: Harvard University Press.

Haegeman, Liliane. 1994. *Introduction to Government and Binding Theory, second edition.* Oxford and Cambridge, MA: Basil Blackwell.

Harman, Gilbert. 1963. Generative grammar without transformation rules: A defense of phrase structure. *Language* 39:597–616.

Harris, Randy Allen. 1993. *The Linguistic Wars.* Oxford: Oxford University Press.

Harris, Zellig S. 1970. *Papers in Structural and Transformational Linguistics.* Dordrecht: Reidel.

Hauser, Marc D., Noam Chomsky, and W. Tecumseh Fitch. 2002. The faculty of language: What is it, who has it, and how did it evolve? *Science* Nov 22 2002:1569–1579.

Hockett, Charles Francis. 1958. *A Course In Modern Linguistics.* New York: Macmillan.

Hopcroft, John E., Rajeev Motwani, and Jeffrey D. Ullman. 2001. *Introduction to Automata Theory, Languages, and Computation.* Reading, MA: Addison-Wesley.

Huck, Geoffrey J., and John A. Goldsmith. 1995. *Ideology and Linguistic Theory.* London and New York: Routledge.

Huddleston, Rodney, and Geoffrey K. Pullum. 2002. *The Cambridge Grammar of the English Language.* Cambridge: Cambridge University Press.

Hudson, Richard. 1984. *Word Grammar.* Oxford: Blackwell.

Hudson, Richard. 1990. *English Word Grammar.* Oxford: Blackwell.

Hudson, Richard. 1998. Word grammar. In V. Agel, et al (Ed.), *Dependency and Valency: An International Handbook of Contemporary Research.* Berlin: Walter de Gruyter.

Jackendoff, Ray. 1972. *Semantic Interpretation in Generative Grammar.* Cambridge, MA: MIT Press.

Jackendoff, Ray. 1975. Morphological and semantic regularities in the lexicon. *Language* 51:639–671.

Jackendoff, Ray. 1994. *Patterns in the Mind.* New York: Basic Books.

Jackendoff, Ray. 2002. *Foundations of Language: Brain, Meaning, Grammar, Evolution.* Oxford: Oxford University Press.

Johnson, David. 1977. On relational constraints on grammars. In P. Cole and J. Sadock (Eds.), *Syntax and Semantics, Volume 8: Grammatical Relations.* New York: Academic Press.

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

Johnson, David, and Paul Postal. 1980. *Arc Pair Grammar.* Princeton: Princeton University Press.

Joshi, Aravind K. 1990. Processing crossed and nested dependencies: An automaton perspective on the psycholinguistic results. *Language and Cognitive Processes* 5:1–27.

Joshi, Aravind K. 2003. Tree-adjoining grammars. In R. Mitkov (Ed.), *Handbook of Computational Linguistics.* Oxford University Press.

Joshi, Aravind K., Tilman Becker, and Owen Rambow. 2000. Complexity of scrambling: A new twist to the competence-performance distinction. In A. Abeillé and O. Rambow (Eds.), *Tree-Adjoining Grammars: Formalisms, Linguistic Analysis and Processing.* Stanford: CSLI Publications.

Joshi, Aravind K., Leon S. Levy, and Masako Takahashi. 1975. Tree adjunct grammars. *Journal of Computer and System Sciences* 10(1).

Jurafsky, Daniel, and James H. Martin. 2000. *Speech and Language Processing – An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition.* Upper Saddle River, New Jersey: Prentice Hall.

Kager, René. 1999. *Optimality Theory.* Cambridge: Cambridge University Press.

Kaplan, Ronald M. 1975. *Transient Processing Load in Relative Clauses.* PhD thesis, Harvard University.

Kaplan, Ronald M., and Annie Zaenen. 1989. Long-distance dependencies, constituent structure and functional uncertainty. In M. R. Baltin and A. S. Kroch (Eds.), *Alternative Conceptions of Phrase Structure,* 17–42. University of Chicago Press. Reprinted in Mary Dalrymple, Ronald Kaplan, John Maxwell, and Annie Zaenen (Eds.) (1995), *Formal Issues in Lexical-Functional Grammar,* Stanford: CSLI Publications, Pages 137–165.

Kasper, Robert, Bernard Kiefer, Klaus Netter, and K. Vijay-Shanker. 1995. Compilation of HPSG to TAG. In *Proceedings of the Association for Computational Linguistics (ACL '95),* 92–99.

Katz, Jerrold J., and Paul M. Postal. 1964. *An Integrated Theory of Linguistic Descriptions.* Cambridge, MA: MIT Press.

Katz, Jerrold J., and Paul M. Postal. 1991. Realism versus conceptualism in linguistics. *Linguistics and Philosophy* 14:515–554.

Kay, Martin. 1979. Functional grammar. In C. Chiarello (Ed.), *Proceedings of the Fifth Annual Meeting of the Berkeley Linguistic Society.*

Kay, Paul. 1995. Construction grammar. In J. Verschueren, J.-O. Ostman, and J. Blommaert (Eds.), *Handbook of Pragmatics.* Amsterdam and Philadelphia: John Benjamins.

Kay, Paul. 2002. An informal sketch of a formal architecture for construction grammar. *Grammars* 5:1–19.

Kay, Paul, and Charles J. Fillmore. 1999. Grammatical constructions and linguistic generalizations: The what?s x doing y? construction. *Language* 75.1:1–33.

Keenan, Edward, and Bernard Comrie. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8:63–99.

Keenan, Edward, and Dag Westerståhl. 1997. Generalized quantifiers in linguistics and logic. In *Handbook of Logic and Language,* 837–893, Amsterdam and Cambridge, MA. North-Holland and MIT Press.

Kim, Jong-Bok. 2000. *The Grammar of Negation: A Constraint-Based Approach.* Stanford: CSLI Publications.

Kim, Jong-Bok, and Ivan A. Sag. 2002. French and English negation without head-movement. *Natural Language and Linguistic Theory* 20(2):339–412.

King, Paul J. 1989. *A Logical Formalism for Head-Driven Phrase Structure Grammar.* PhD thesis, University of Manchester.

Koster, Jan. 1987. *Domains and Dynasties, the Radical Autonomy of Syntax.* Dordrecht: Foris.

Kurtzman, Howard S., and Maryellen C. MacDonald. 1993. Resolution of quantifier scope ambiguities. *Cognition* 48:243–279.

Labov, William. 1969. Contraction, deletion, and inherent variability of the English copula. *Language* 45:715–762.

Labov, William. 1972. *Language in the Inner City: Studies in the Black English Vernacular.* Philadelphia: University of Pennsylvania Press.

Labov, William. 1995. The case of the missing copula: The interpretation of zeroes in African-Amerircan English. In *Language* (Gleitman and Liberman 1995), 25–54.

Labov, William, Paul Cohen Cohen, Clarence Robins, and John Lewis. 1968. A study of the nonstandard English of Negro and Puerto Rican speakers in New York City. Technical Report Final Report, Cooperative Research Project No. 3288, United States Office of Education.

Lakoff, George. 1987. *Women Fire and Dangerous Things*. Chicago and London: University of Chicago Press.

Langacker, Ronald. 1987. *Foundations of Cognitive Grammar (vol 1)*. Stanford, CA: Stanford University Press.

Lappin, Shalom, Robert Levine, and David Johnson. 2000a. The revolution confused: A reply to our critics. *Natural Language and Linguistic Theory* 18:873–890.

Lappin, Shalom, Robert Levine, and David Johnson. 2000b. The structure of unscientific revolutions. *Natural Language and Linguistic Theory* 18:665–671.

Lappin, Shalom, Robert Levine, and David Johnson. 2001. The revolution maximally confused. *Natural Language and Linguistic Theory* 19:901–919.

Larson, Richard. 1995. Semantics. In *Language* (Gleitman and Liberman 1995), 361–380.

Lascarides, Alex, Edward Briscoe, Nicholas Asher, and Ann Copestake. 1996. Order independent and persistent typed default unification. *Linguistics and Philosophy* 19(1):1–89.

Lascarides, Alex, and Ann Copestake. 1999. Default representation in constraint-based frameworks. *Computational Linguistics* 25(1):55–105.

Lasnik, Howard. 1976. Remarks on coreference. *Linguistic Analysis* 2:1–22. Reprinted in Lasnik (1989), *Essays on Anaphora*, Dordrecht: Kluwer.

Lasnik, Howard. 1995. The forms of sentences. In *Language* (Gleitman and Liberman 1995).

Lasnik, Howard, Marcela Depiante, and Arthur Stepanov. 2000. *Syntactic Structures Revisited: Contemporary Lectures on Classic Transformational Theory*. Cambridge, MA: MIT Press.

Legendre, Géraldine, Jane Grimshaw, and Sten Vikner (Eds.). 2001. *Optimality-Theoretic Syntax*. Cambridge: MIT Press.

Levin, Beth. 1993. *English Verb Classes and Alternations: A Preliminary Investigation*. Chicago: University of Chicago Press.

Levine, Robert D., and Ivan A. Sag. 2003. *WH*-nonmovement. *Gengo Kenky: Journal of the Linguistic Society of Japan* 123.

MacDonald, Maryellen C., Neal J. Pearlmutter, and Mark S. Seidenberg. 1994. The lexical nature of syntactic ambiguity resolution. *Psychological Review* 101(4).

Malouf, Robert. in press. Cooperating constructions. In E. Francis and L. Michaelis (Eds.), *Linguistic Mismatch: Scope and Theory*. Stanford: CSLI Publications.

Marcus, Gary F. 2001. *The Algebraic Mind: Integrating Connectionism and Cognitive Science*. MIT Press.

Marcus, Gary F. 2004. *The Birth of the Mind*. Basic Books.

Marslen-Wilson, William D., and Lorraine K. Tyler. 1987. Against modularity. In J. Garfield (Ed.), *Modularity in Knowledge Representation and Natural Language Understanding*. MIT Press.

Matthews, Peter. 1993. *Grammatical Theory in the United States: from Bloomfield to Chomsky*. Cambridge: Cambridge University Press.

McCawley, James D. 1971. Tense and time reference in English. In C. J. Fillmore and D. T. Langendoen (Eds.), *Studies in Linguistic Semantics*. New York: Holt, Rinehart, and Winston.

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

Meurers, Walt Detmar. 1999. *Lexical Generalizations in the Syntax of German Non-Finite Constructions*. PhD thesis, Seminar für Sprachwissenschaft, Universität Tübingen, Tübingen, Germany. published 2000 as Volume 145 in Arbeitspapiere des SFB 340, ISSN 0947-6954/00.

Meurers, Walt Detmar. 2001. On expressing lexical generalizations in HPSG. *Nordic Journal of Linguistics* 24(2):161–217. Special issue on 'The Lexicon in Linguistic Theory'.

Michaelis, Laura, and Knud Lambrecht. 1996. Toward a construction-based theory of language function: The case of nominal extraposition. *Language* 72:215–248.

Milsark, Gary. 1977. Towards an explanation of certain peculiarities of the existential construction in English. *Linguistic Analysis* 3:1–31.

Montague, Richard. 1970. Universal grammar. *Theoria* 36:373–398. Reprinted in Richmond Thomason, (Ed.) (1974), *Formal Philosophy*, New Haven: Yale University Press.

Moortgat, Michael. 1997. Categorial type logics. In J. van Benthem and A. ter Meulen (Eds.), *Handbook of Logic and Language*, 93–177. Amsterdam: Elsevier and Cambridge, MA: MIT Press.

Morrill, Glynn. 1994. *Type Logical Grammar*. Dordrecht: Kluwer.

Mufwene, Salikoko S., Guy Bailey, John R. Rickford, and John Baugh (Eds.). 1998. *African-American English: Structure, History, and Use*. London: Routledge.

Nevin, Bruce (Ed.). 2003. *The Legacy of Zellig Harris – Language and Information into the 21st Century, Volume 1*. Amsterdam and Philadelphia: John Benjamins.

Nevin, Bruce, and Stephen Johnson (Eds.). 2003. *The Legacy of Zellig Harris – Language and Information into the 21st Century, Volume 2*. Amsterdam and Philadelphia: John Benjamins.

Newmeyer, Frederick J. 1986. *Linguistic Theory in America, Second Edition*. London: Academic Press.

Nunberg, Geoffrey. 1983. The grammar wars. *The Atlantic Monthly* 256(6):31–58.

Nunberg, Geoffrey, Ivan A. Sag, and Thomas Wasow. 1994. Idioms. *Language* 70:491–538.

Partee, Barbara H. 1995. Lexical semantics and compositionality. In *Language* (Gleitman and Liberman 1995).

Partee, Barbara H., Alice ter Meulen, and Robert Wall. 1990. *Mathematical Methods in Linguistics*. Dordrecht: Kluwer.

Pearlmutter, Neal J., and Maryellen MacDonald. 1992. Plausibility and syntactic ambiguity resolution. In *Proceedings of the 14th Annual Conference on Cognitive Science*, 498–503, Hillsdale, N.J. Erlbaum.

Pedersen, Holger. 1959. *The Discovery of Language: Linguistic Science in the Nineteenth Century.* Bloomington: Indiana University Press. Translated by John Webster Spargo.

Penn, Gerald. 2000. *The Algebraic Structure of Attributed Type Signatures.* PhD thesis, Carnegie Mellon University (Computer Science). Document available online at http://www.cs.toronto.edu/~gpenn/publications.html.

Perlmutter, David (Ed.). 1983. *Studies in Relational Grammar 1.* Chicago: University of Chicago Press.

Perlmutter, David, and Paul Postal. 1977. Toward a universal characterization of passivization. In *Proceedings of the 3rd Annual Meeting of the Berkeley Linguistics Society*, Berkeley. University of California, Berkeley. Reprinted in Perlmutter (1983).

Perlmutter, David, and Scott Soames. 1979. *Syntactic Argumentation and the Structure of English.* Berkeley: University of California Press.

Pinker, Steven. 1994. *The Language Instinct.* New York: Morrow.

Pollard, Carl, and Ivan A. Sag. 1987. *Information-Based Syntax and Semantics, Volume 1: Fundamentals.* Stanford: CSLI Publications.

Pollard, Carl, and Ivan A. Sag. 1992. Anaphors in English and the scope of binding theory. *Linguistic Inquiry* 23:261–303.

Pollard, Carl, and Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar.* Chicago: University of Chicago Press.

Pollock, Jean-Yves. 1989. Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20:365–424.

Posner, Michael I. (Ed.). 1989. *Foundations of Cognitive Science.* Cambridge, MA: MIT Press.

Postal, Paul. 1964. *Constituent Structure: A Study of Contemporary Models of Syntactic Description.* Bloomington: Research Center for the Language Sciences, Indiana University.

Postal, Paul. 1974. *On Raising.* Cambridge, MA: MIT Press.

Postal, Paul. 1986. *Studies of Passive Clauses.* Albany: SUNY Press.

Postal, Paul, and Brian Joseph (Eds.). 1990. *Studies in Relational Grammar 3.* Chicago: University of Chicago Press.

Postal, Paul, and Geoffrey K. Pullum. 1988. Expletive noun phrases in subcategorized positions. *Linguistic Inquiry* 19:635–670.

Prince, Alan, and Paul Smolensky. 1993. *Optimality Theory: Constraint Interaction in Generative Grammar.* Tech Report RuCCS-TR-2. ROA-537: Rutgers University Center for Cognitive Science.

Quirk, Randoph, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik. 1972. *A Grammar of Contemporary English.* London and New York: Longman.

Quirk, Randoph, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language.* London and New York: Longman.

Radford, Andrew. 1997. *Syntactic Theory and the Structure of English: A Minimalist Approach.* New York and Cambridge: Cambridge University Press.

Richter, Frank. 1999. RSRL for HPSG. In V. Kordoni (Ed.), *Tübingen Studies in Head-Driven Phrase Structure Grammar*, no. 132 in Arbeitsberichte des SFB 340, 74–115. Tübingen: Universität Tübingen.

Richter, Frank. 2000. *A Mathematical Formalism for Linguistic Theories with an Application in Head-Driven Phrase Structure Grammar.* PhD thesis, Universität Tübingen.

Rickford, John R., and Russell J. Rickford. 2000. *Spoken Soul: The Story of Black English.* Hoboken, NJ: John Wiley & Sons.

Riehemann, Susanne. 2001. *A Constructional Approach to Idioms and Word Formation.* PhD thesis, Stanford University.

Robins, R. H. 1967. *A Short History of Linguistics.* Bloomington: University of Indiana Press.

Rosenbaum, Peter. 1967. *The Grammar of English Predicate Complement Constructions.* Cambridge, MA: MIT Press.

Ross, John R. 1967. *Constraints on Variables in Syntax.* PhD thesis, MIT. Published as Infinite Syntax! Norwood, N.J.: Ablex, 1986.

Ross, John R. 1969. Auxiliaries as main verbs. In W. Todd (Ed.), *Studies in Philosophical Linguistics 1.* Evanston, Ill.: Great Expectations Press.

Ruwet, Nicolas. 1991. *Syntax and Human Experience.* Chicago: University of Chicago Press. Edited and translated by J. Goldsmith.

Sag, Ivan A. 1997. English relative clause constructions. *Journal of Linguistics* 33(2):431–484.

Sag, Ivan A. 2003. Coordination and underspecification. In *Proceedings of the 9th International Conference on Head-Driven Phrase Structure Grammar.* Available online at http://cslipublications.stanford.edu/HPSG/3/hpsg02-toc.html.

Sag, Ivan A. to appear. Rules and exceptions in the English auxiliary system. *Journal of Linguistics.*

Sag, Ivan A., Timothy Baldwin, Francis Bond, Ann Copestake, and Dan Flickinger. 2002. Multiword expressions: A pain in the neck for nlp. In A. Gelbukh (Ed.), *Proceedings of CICLING-2002.* Springer Verlag. Also appeared as LinGO Working Paper No. 2001-03. See http://lingo.stanford.edu/.

Sag, Ivan A., and Janet D. Fodor. 1994. Extraction without traces. In *Proceedings of the Thirteenth Annual Meeting of the West Coast Conference on Formal Linguistics*, Stanford. CSLI Publications.

Sag, Ivan A., and Carl Pollard. 1991. An integrated theory of complement control. *Language* 67:63–113.

Savitch, Walter J., Emmon Bach, William Marsh, and Gila Safran-Naveh. 1987. *The Formal Complexity of Natural Language.* Dordrecht: D. Reidel.

Schütze, Carson T. 1996. *The Empirical Base of Linguistics.* Chicago: University of Chicago Press.

Sells, Peter. 1985. *Lectures on Contemporary Syntactic Theories.* Stanford: CSLI Publications.

Sells, Peter (Ed.). 2001. *Formal and Empirical Issues in Optimality Theoretic Syntax.* Stanford: CSLI Publications.

Shieber, Stuart. 1986. *An Introduction to Unification-Based Approaches to Grammar.* Stanford: CSLI Publications.

Skinner, B. F. 1957. *Verbal Behavior.* New York: Appleton-Century-Crofts.

Smith, Jeffrey D. 1999. English number names in HPSG. In G. Webelhuth, J.-P. Koenig, and A. Kathol (Eds.), *Lexical and Constructional Aspects of Linguistic Explanation*, 145–160. Stanford: CSLI Publications.

Steedman, Mark. 1996. *Surface Structure and Interpretation.* Cambridge, MA: MIT Press.

Steedman, Mark. 2000. *The Syntactic Process.* Cambridge, MA: MIT Press/Bradford Books.

Steele, Susan. 1981. *An Encyclopedia of AUX.* Cambridge, MA: MIT Press.

Tabossi, Patrizia., Michael J. Spivey-Knowlton, Ken McRae, and Michael K. Tanenhaus. 1994. Semantic effects on syntactic ambiguity resolution: Evidence for a constraint-based resolution process. In C. Umilta and M. Moscovitch (Eds.), *Attention and Performance XV.* Hillsdale, N.J.: Erlbaum.

Tanenhaus, Michael K., Michael J. Spivey-Knowlton, Kathleen M. Eberhard, and July C. Sedivy. 1995. Integration of visual and linguistic information in spoken language comprehension. *Science* 268:1632–1634.

Tanenhaus, Michael K., Michael J. Spivey-Knowlton, Kathleen M. Eberhard, and July C. Sedivy. 1996. Using eye movements to study spoken language comprehension: Evidence for visually mediated incremental interpretation. In T. Inui and J. L. McClelland (Eds.), *Attention and Performance XVI: Information Integration in Perception and Communication*, 457–478. Cambridge, MA: MIT Press.

Tanenhaus, Michael K., and John C. Trueswell. 1995. Sentence comprehension. In J. Miller and P. Eimas (Eds.), *Handbook of Perception and Cognition*, Vol. 11. London: Academic Press.

Tesnière, Lucien. 1959. *Elements de Syntaxe Structurale.* Paris: C. Klincksieck.

Thomason, Richmond (Ed.). 1974. *Formal Philosophy: Selected Papers of Richard Montague.* New Haven: Yale University Press.

Tomasello, Michael. 1992. The social bases of language acquisition. *Social Development* 1:67–87.

Trask, Robert Lawrence. 1993. *A Dictionary of Grammatical Terms in Linguistics.* London and New York: Routledge.

Trask, Robert Lawrence. 2002. Review of Baker 2001. *The Human Nature Review* 2:77–81. Available online at http://human-nature.com/nibbs/02/trask.html.

Trueswell, John C., Michael K. Tanenhaus, and Susan M. Garnsey. 1992. Semantic influences on parsing: Use of thematic role information in syntactic disambiguation. *Journal of Memory and Language* 33.

Trueswell, John C., Michael K. Tanenhaus, and Christopher Kello. 1993. Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 19:528–553.

Vijay-Shanker, K. 1987. *A Study of Tree-Adjoining Grammars.* PhD thesis, University of Pennsylvania.

Villavicencio, Aline, and Ann Copestake. 2002. On the nature of idioms. *LinGO Working Paper* No. 2002-04. Available online at: http://lingo.stanford.edu/.

Warner, Anthony. 1993. *English Auxiliaries: Structure and History.* Cambridge and New York: Cambridge University Press.

Warner, Anthony R. 2000. English auxiliaries without lexical rules. In R. Borsley (Ed.), *Syntax and Semantics Volume 32: The Nature and Function of Syntactic Categories*, 167–220, San Diego and London. Academic Press.

Wasow, Thomas. 1977. Transformations and the lexicon. In *Formal Syntax* (Culicover et al. 1977).

Wasow, Thomas. 1989. Grammatical theory. In *Foundations of Cognitive Science* (Posner 1989).

Webelhuth, Gert (Ed.). 1995. *Government and Binding Theory and the Minimalist Program.* Oxford: Basil Blackwell.

Weir, David. 1987. *Characterizing Mildly Context-Sensitive Grammar Formalisms.* PhD thesis, University of Pennsylvania.

Wilkes-Gibbs, Deanna. 1986. *Collaborative Proccesses of Language Use in Conversation.* PhD thesis, Stanford University.

Wood, Mary. 1993. *Categorial Grammars.* London and New York: Routledge.

XTAG, The Research Group. 2000. A lexicalized tree-adjoining grammar for English. Technical Report 98-18, Institute for Research in Cognitive Science, Philadelphia.

Zwicky, Arnold, and Geoffrey K. Pullum. 1983. Cliticiziation vs. inflection: English *n't*. *Language* 59:502–13.