What is Linguistics?
Part II

Matilde Marcolli

MAT1509HS: Mathematical and Computational Linguistics

University of Toronto, Winter 2019, T 4-6 and W 4, BA6180

Morphology

- word-forms (sing, sang, sings, singing, . . .)
- lexeme underlying "vocabulary-word", base-form, different word-forms of same lexeme
- morphological rules: two kinds
  - inflection rules (relate different forms of same lexeme): conjugation, declension
  - word formation (combine different lexemes): e.g. dishwasher
- word formation: two kinds
  - derivation: affixing bound-forms (sing-er, slow-ly, ...)
  - compounding: combines complete word forms (dish-washer)

- Some languages extremely rich in compound words, other poor
- Even within same language family huge differences

Curious example: among ancient Indo-European languages, Sanskrit and Ancient Greek are very rich in compound words (Homer's $\rho o \delta o \delta \alpha \kappa \tau \upsilon \lambda o \varsigma$ $H \omega \varsigma$)
but Hittite has no compound word formation at all

- paradigm: the set of all word-forms associated to a given lexeme

Examples:
- conjugation of verbs (tense, aspect, mood);
- declension of nouns (number, gender, case);
- personal pronouns arranged by person, number, gender

Allomorphy

- morpheme: smallest grammatical units, roots and affixes
- allomorphs: different morphemes playing same grammatical role

Example: negation prefixes in English
- *a-*, *an-* (from Greek): anesthesia, anisotropic, acyclic
- *in-*, *im-* (from Latin): impossible, incompressible, invincible
- *un-* (English): unbiased, unaffected, unacceptable

Example: different forms of plural in English
boy $\mapsto$ boys; watch $\mapsto$ watches;
child $\mapsto$ children; woman $\mapsto$ women

Example: strong verbs (sleep/slept)

- phonological allomorphs: regular phonological rules
- suppletive allomorphs: exceptional

Morphological Typology

Grouping languages by morphological structures

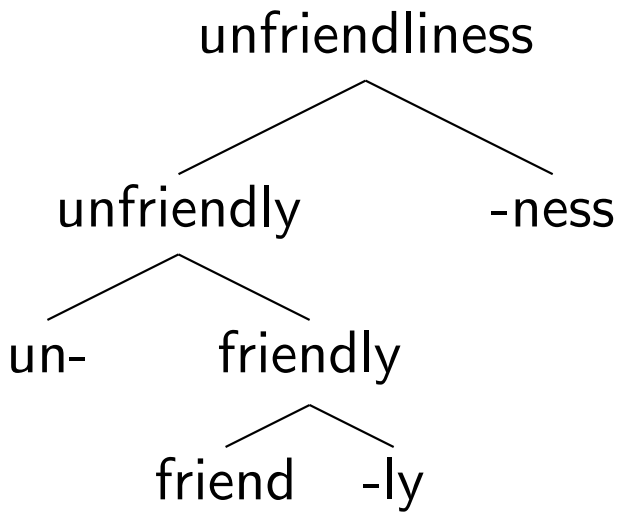
• Analytic: small amount of inflection, replaced by word order and additional word (Mandarin)
• Isolating: few morphemes per word (Vietnamese)
• Synthetic: typically several morphemes can combine in words, high in inflection forms (many Indo-European languages)
Examples:
- German: *Abstimmungsbekanntmachung*
- Russian: Достопримечательность

• Polysynthetic: extremely long compound words with *sentence-words*
(America, Australia, Siberia, Papua New Guinea)
- compositionally polysynthetic, affixally polysynthetic
- incorporating, agglutinating, fusional

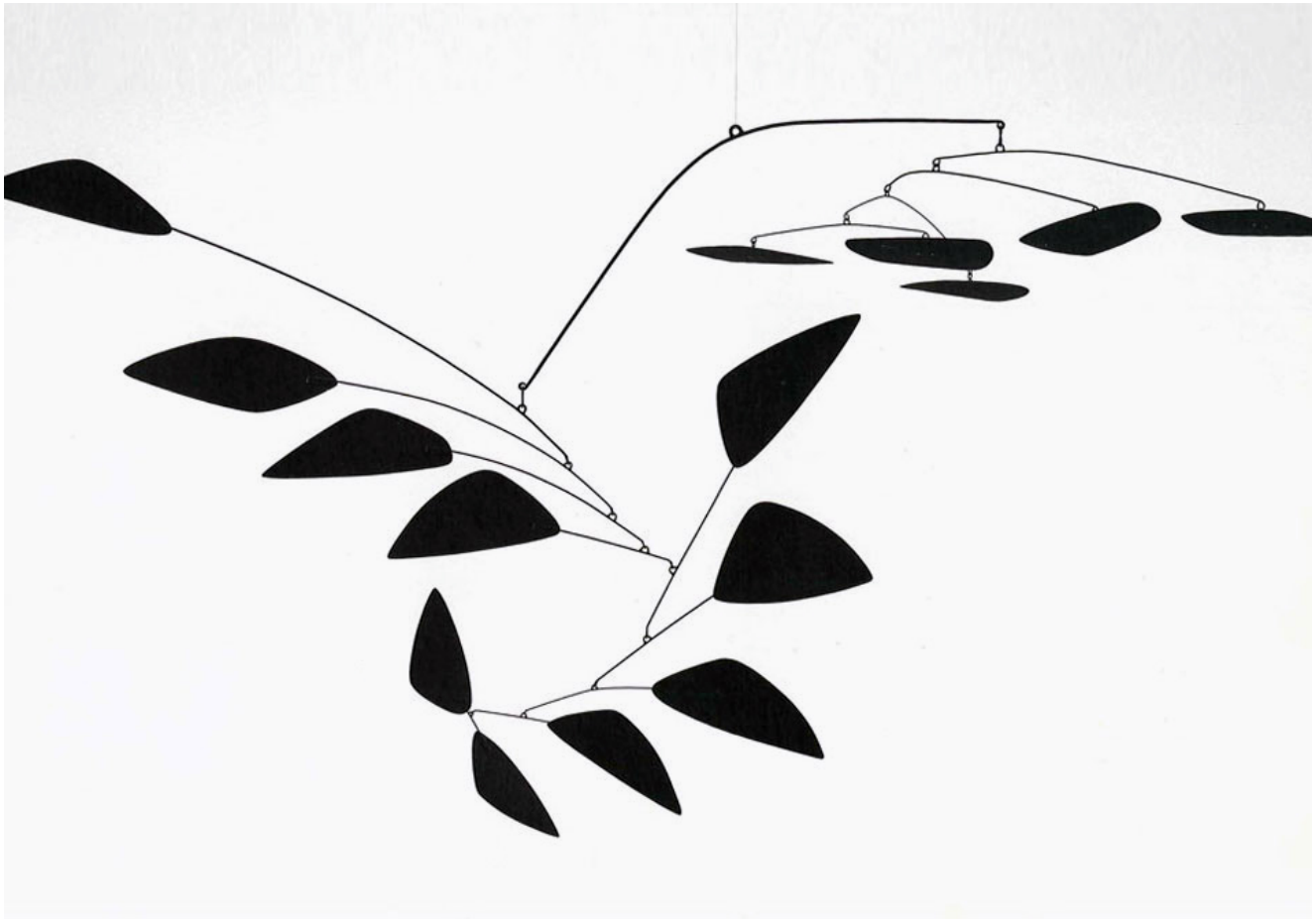
Hierarchical structures

- *bracketing* (as in non-associative algebra)

$$\text{unfriendliness} = (\,(\,\text{un-}\,(\,(\,\text{friend}\,)\,\text{-ly}\,)\,)\,\text{-ness}\,)$$



Linguistics kind of looks like this...



Alexander Calder, *Mobile*, 1960

Lexicology (and lexical semantics)

• *diachronic*: changes across time in the use of words and word formation

• *synchronic* (Structuralist): lexical relations (at a given time), syntagmatic lexical relations (culturally determined patterns of association between lexical units)

• various WordNet lexical and semantic databases

• Phraseology: *phrasemes*= multi-word lexical units, includes study of idiomatic expressions (e.g. "it's raining cats and dogs")

• Etymology: origin and history of words, crucial role in historical linguistics: comparative methods, reconstruction of proto-languages

Syntax the large-scale structure of languages

- the basic units of structure at this level are *sentences*
- rules and principles governing sentence structure (within a language, or across languages)
- origin of scientific syntactic theory: 4th century BCE
अष्टाध्यायी (*Aṣṭādhyāyī*) of पाणिनि (Pāṇini)
- origin of "traditional grammar": 2nd century BCE
Διονύσιος ὁ Θρᾷξ Τέχνη γραμματική
- Dionysius Thrax's Techne was a primarily morphological grammar, little emphasis on syntax, while Pāṇini focused on all aspects (phonology, morphology, syntax): basis of modern syntactic theory

Modern Syntactic Theory:

• i-language versus e-language: internal language (mental) as opposed to external (community based records of language use): focus on i-language as object of study

• grammaticality: judgement on whether a sentence is well formed (grammatical) in a given language, i-language gives people the capacity to decide on grammaticality

• generative grammar: produce a set of rules that correctly predict grammaticality of sentences

• universal grammar: ability to learn grammar is built in the human brain, e.g. properties like distinction between nouns and verbs are universal, ... is universal grammar a falsifiable theory?

Modern Syntactic Theory: some of the main models

- Transformational grammar
- Government and Binding (Principles and Parameters)
- Minimalist Program
- Head-driven Phrase Structure Grammar
- Lexical Functional Grammar
- Tree-adjoining Grammar

Transformational Grammar (Chomsky, 1957)

- Sentences have two levels of structure: *deep structure* and *surface structure*
- *deep structure*: closer to semantic level, properties common across languages, mapped to surface structure via *transformations* that operate on parse trees underlying sentences
- *surface structure*: language specific
- a transformational grammar is a system of tree automata
- in more recent theories (minimalist program), deep structure and surface structure replaced by *logical form* and *phonetic form*

...more details later, after discussing formal languages

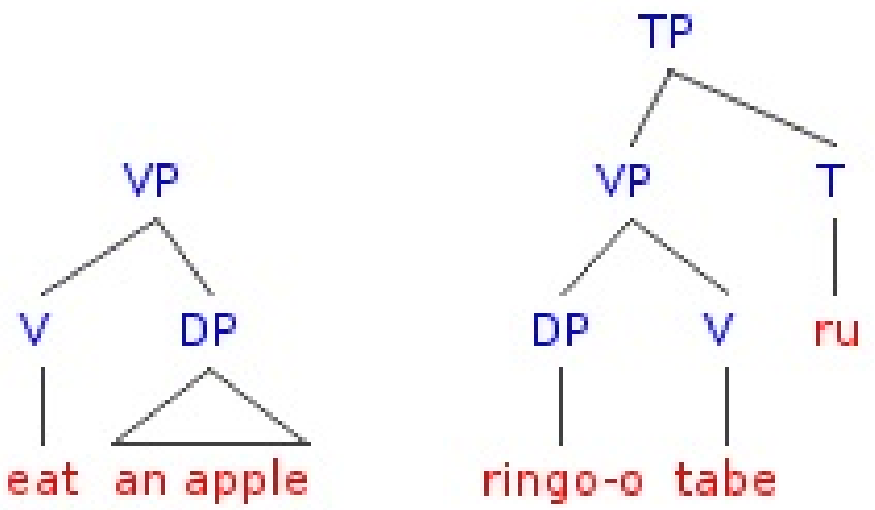
Government and Binding (Principles and Parameters)

(Chomsky, 1981)

- *principles*: general rules of grammar
- *parameters*: binary variables (on/off switches) that distinguish languages in terms of syntactic structures
- Example of parameter: head-directionality

(head-initial versus head-final)

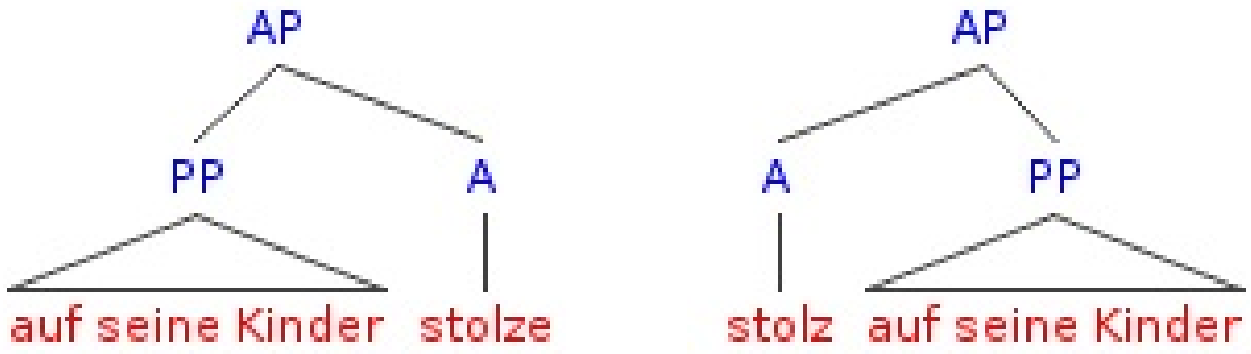
English is head-initial, Japanese is head-final



VP= verb phrase, TP= tense phrase, DP= determiner phrase

...but not always so clear-cut: German can use both structures

auf seine Kinder stolze Vater (head-final) or
er ist stolz auf seine Kinder (head-initial)



AP= adjective phrase, PP= prepositional phrase

• Corpora based statistical analysis of head-directionality (Haitao Liu, 2010): a continuum between head-initial and head-final

Examples of Principles

- *Structure Preservation Principle*: identifies transformations preserving deep structure (e.g. rephrasing in passive form)
- *Projection Principle*: lexical properties preserved when forming new sentences from given ones (phrase structure rules projected from lexical rules)
- *Subjacency Principle*: transformation moves are "local" (don't move elements of phrases across more than one "bounding node" S=sentence and NP=noun phrase)

Examples of Parameters

- *Head-directionality*
- *Subject-side*
- *Pro-drop*
- *Null-subject*

Problems

- Interdependencies between parameters
- Diachronic changes of parameters in language evolution

Word Order and Parameters

- *Subject-side* parameter: positioning of the subject with respect to the head (specifier-head, head-specifier, and subject-initial, subject-medial, subject-final)
- Word Order: SOV, SVO, VSO, VOS, OVS, OSV

Word Orders	Percentage		
SOV	41.03%	Subject-initial	Specifier-Head
SVO	35.44%		
VSO	6.90%	Subject-medial	Head-Specifier
VOS	1.82%	Subject-final	
OVS	0.79%		
OSV	0.29%	Subject-medial	Specifier-Head

Very uneven distribution across world languages

Changes over time in Word Order

• Ancient Greek: switched from Homeric to Classical
- A. Taylor, *The change from SOV to SVO in Ancient Greek*, Language Variation and Change, 6 (1994) 1–37

• Sanskrit: different word orders allowed, but prevalent one in Vedic Sanskrit is SOV
- F.J. Staal, *Word Order in Sanskrit and Universal Grammar*, Springer, 1967

• English: switched from Old English (transitional between SOV and SVO) to Middle English (SVO)
- J. McLaughlin, *Old English Syntax: a handbook*, Walter de Gruyter, 1983.

- Word order distribution: a neuroscience explanation?

- D. Kemmerer, *The cross-linguistic prevalence of SOV and SVO word orders reflects the sequential and hierarchical representation of action in Broca's area*, Language and Linguistics Compass, 6 (2012) N.1, 50–66.

- Internal reasons for diachronic switch?

- F.Antinucci, A.Duranti, L.Gebert, *Relative clause structure, relative clause perception, and the change from SOV to SVO*, Cognition, Vol.7 (1979) N.2 145–176.

Dependent parameters

- null-subject parameter: can drop subject

Example: among Latin languages, Italian and Spanish have null-subject (+), French does not (-)

it rains, piove, llueve, il pleut

- pro-drop parameter: can drop pronouns in sentences

不知道。喜欢吗?

Bù zhīdào. Xǐhuan ma?

- Pro-drop controls Null-subject

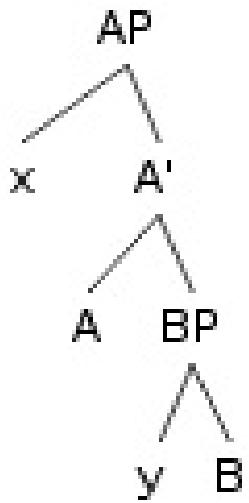
How many independent parameters?

Government and Binding

- based on Principles and Parameters model of language
- if A and B are two nodes in a syntactic tree,

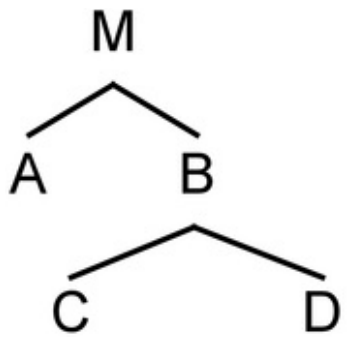
A *m-commands* B iff

- neither node dominates the other
- the maximal projection AP of A dominates B



A m-commands B, but B does not m-command A

• if X and Y are two nodes in a syntactic tree,
X *c-commands* Y (constituent command) iff
- neither node dominates the other
- the first node that dominates X also dominates Y

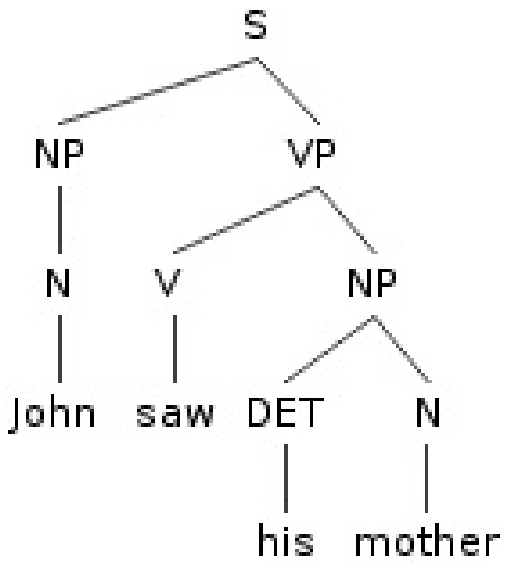


A and B c-command each other, C and D also, A also c-commands C and D

• A node X m-commands all nodes it c-commands, but also the nodes in XP

- node A governs B iff
  - A is a governor (head of lexical category V,N,A,...)
  - A m-commands B
  - no barrier between A and B
- A barrier (between A and B) is a node X in a syntactic tree
  - X c-commands B
  - X does not c-command A

- Binding A binds B iff
- A c-commands B
- A and B are coreferential (refer to the same person)



in this sentence, "John" binds "his"

- These rules are used to test grammaticality of sentences

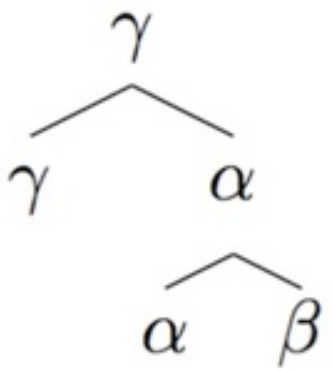
Minimalist Program (Chomsky, 1993)

- a program not a theory: guiding conceptual framework
- minimalist questions, whose answer may be framed within the context of different theories
- within Principles and Parameters setting
- postulates the existence of an underlying *simple computational structure* responsible for linguistic capability in the human mind (related to the idea of Universal Grammar)
- some minimality assumptions: *economy of representation* (sentence structure no more complicated than minimally required to satisfy constraints imposed by grammaticality); *economy of derivation* (transformations only occur if they make parts of sentence interpretable, e.g. disambiguation produced by inflection)

- Bare phrase structure: an explicitly derivational model of sentence building (as opposed to representational)
- Basic operations: merge and move
- Merge: $(\alpha, \beta) \mapsto \{\alpha, \{\alpha, \beta\}\}$ or $\{\beta, \{\alpha, \beta\}\}$

Example: (drink, water) $\mapsto$ { drink , { drink , water }}
Example: (cold, water) $\mapsto$ { water , { cold , water }}
The first merged "drink water" can be inserted in a sentence in place of "drink"; the second merge "cold water" can be inserted in place of "water"

- iterations: $(\gamma, \{\alpha, \{\alpha, \beta\}\}) \mapsto \{\gamma, \{\gamma, \{\alpha, \{\alpha, \beta\}\}\}\}$



- Move: moving parts of a sentence within the sentence

Example:
You are looking for someone
Whom are you looking for?

- Placeholder symbol (trace) for the moved element

(Whom) are you looking for (t)?
"Whom" and its trace symbol "(t)" are a *chain*,
similarly for other elements that change position

- Currently different approaches on how best to formulate the "Move" operations

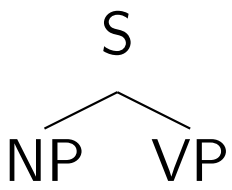
Head-driven phrase structure grammar (HPSG)
(Carl Pollard, Ivan Sag 1987)

- used in natural language processing (parsing)
- lexical data with entries marked by *types* hierarchy
- *signs*: words, or phrases; with location in type hierarchy and with internal *feature structure*
- these properties represented by *attribute value matrices* (AVM)

$$\begin{bmatrix} \text{category} & \textit{noun phrase} \\ \text{agreement} & \begin{bmatrix} \text{number} & \textit{singular} \\ \text{person} & \textit{third} \end{bmatrix} \end{bmatrix}$$

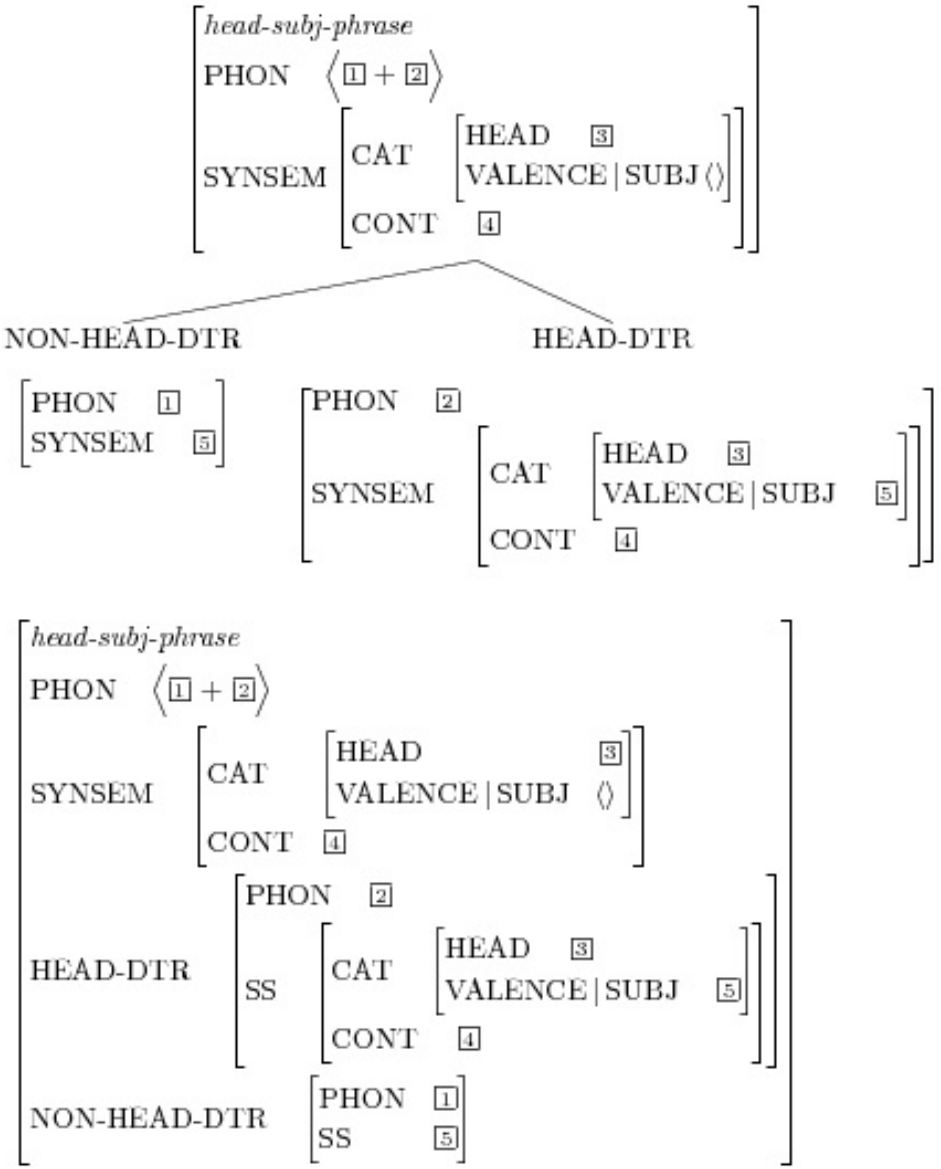
- constraints (ID/LP)
- *immediate dominance* (ID) and *linear precedence* (LP)

Example:

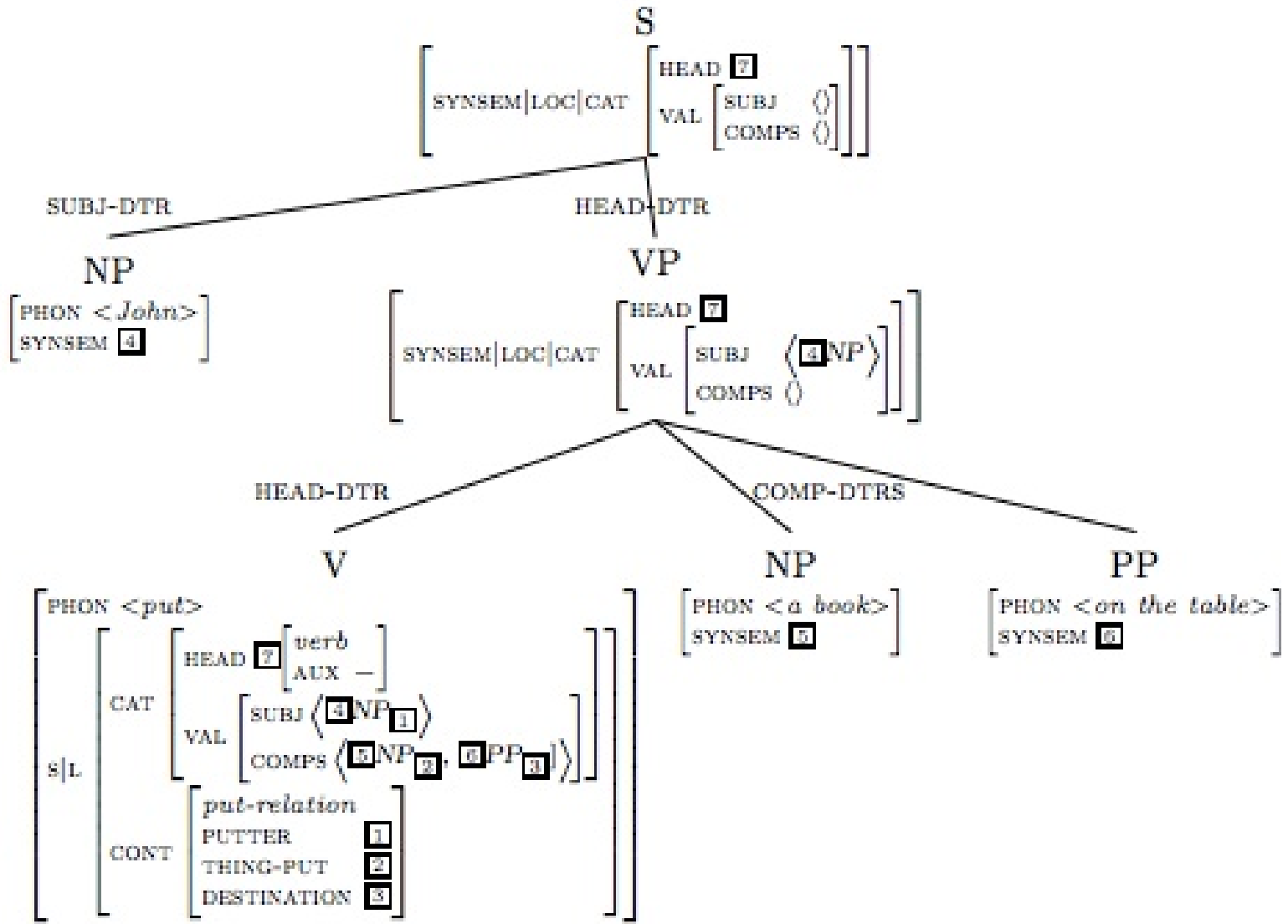


S-node dominates NP and VP nodes: ID relation
and NP precedes VP: LP relation [NP VP]

Example: immediate dominance rule for a head-subj-phrase

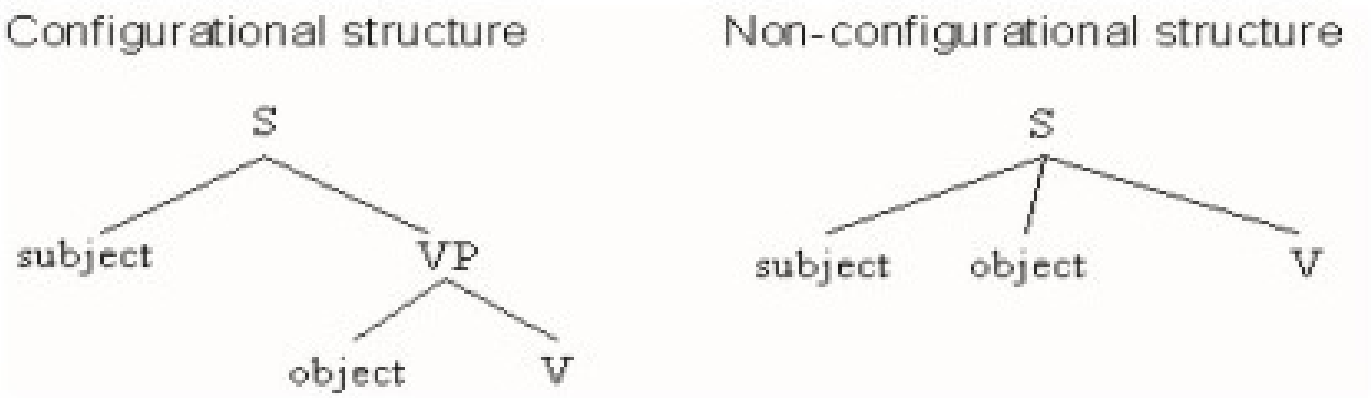


- HPSG generates strings by combining signs



Lexical Functional Grammar (LFG)
(Joan Bresnan and Ronald Kaplan, 1982)

- some languages appear to challenge rigid phrase structure, and have free word order
- *non-configurational languages* (Mohawk, Nahuatl, Warlpiri, ...)



- Principles and Parameters approach: not really non-configurational
- Lexical Functional Grammar: alternative approach that works for non-configurational

- Feature structures (as with HPSG): f-structure
- Syntactic constituents (trees, ID/LP): c-structure
- other levels of structure: s-structure (semantic), m-structure (morphological), p-structure (phonological)
- operations interpreted lexically instead of acting on trees (e.g. passivization: both active and passive forms in lexicon)
- no deep structure/surface structure
- includes a theory of *syntactic-semantic interface*: glue semantics

Tree Adjoining Grammar (Aravind Joshi, 1969)

- developed as formal languages (as a generalization of context-free grammars)
- rooted trees with a marked foot leaf node (a word); *basic trees* and *auxiliary trees* (these have same symbol labeling root and foot)
- two operations: *substitution* (leaf/root grafting) and *adjunction* (insertion of an auxiliary tree at an internal node labelled by auxiliary root/foot label)
- main idea: these two operations should suffice to describe all syntactic dependencies
- LTAG: lexicalized tree-adjoining grammar: each elementary tree associated with an item in a lexical database (XTAG project, LTAG parser)

... more details after discussing formal languages