Syntactic Theory

Lecture 8b (22.01.2009)

PD Dr.Valia Kordoni

Email: kordoni@coli.uni-sb.de

http://www.coli.uni-saarland.de/courses/syntactic-theory-08/

Head-Driven Phrase Structure Grammar (HPSG)

Introduction – Part IX –

Constraints on Extraction

The CELR and SELR only allow complements and embedded subjects to be extracted. We correctly predict the following structures to be ungrammatical:

(1) Specifier extraction
Kim stole those potatoes. →
   a. * [Those]$_i$, Kim stole ____$_i$ potatoes.
   b. * [Which]$_i$ did Kim steal ____$_i$ potatoes?
   c. [Which potatoes]$_i$ did Kim steal ____$_i$?

(2) Modifier extraction
Sandy has a red car. →
   a. * [Red]$_i$, Sandy has a ____$_i$ car.
   b. * [What color]$_i$ does Sandy have a ____$_i$ car?
   c. [What color car]$_i$ does Sandy have ____$_i$?

(3) Head extraction
Kim just met the queen of Denmark. →
   a. * [The queen]$_i$, Kim just met ____$_i$ of Denmark.
   b. * Who$_i$ did Kim just meet ____$_i$ of Denmark?

slashed subjects

The CELR incorrectly allows some cases of partial extraction (or "subextraction"):

(4) a. Newspaper articles about Kim were very revealing.
→ *Who$_i$ were [newspaper articles about ____$_i$] very revealing?

b. That Sandy stole the potatoes was the prosecution's claim.
→ *[The potatoes]$_i$, [that Sandy stole ____$_i$] was the prosecution's claim.

slashed subjects (cont.)

These ungrammatical sentences can be licensed by applying the CELR to *about* and *stole*. For example, *about* changes into an intransitive preposition:

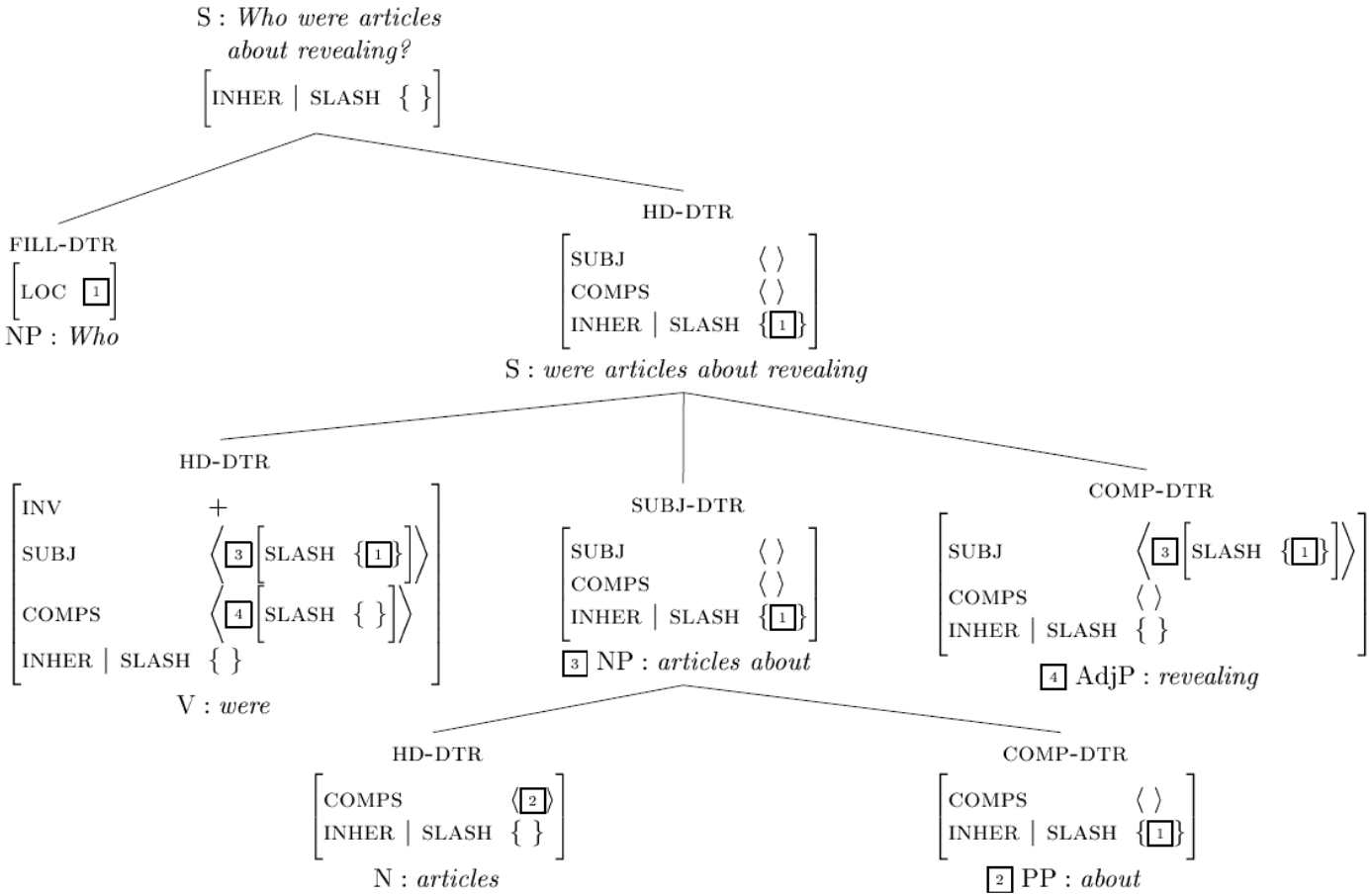
(5)

$$\begin{bmatrix} \text{PHON} & \langle about \rangle \\ \text{COMPS} & \left\langle \text{NP} \left[\text{LOC} \ \boxed{1} \right] \right\rangle \\ \text{INHER} \mid \text{SLASH} & \{\ \} \end{bmatrix} \mapsto \begin{bmatrix} \text{PHON} & \langle about \rangle \\ \text{COMPS} & \langle\ \rangle \\ \text{INHER} \mid \text{SLASH} & \{\boxed{1}\} \end{bmatrix}$$

The INHER | SLASH specification is passed up the tree as expected.

slashed subjects (cont.)

(6) Structure for sentence (4a):



slashed subjects (cont.)

So far we have no way to block this structure. Notice that the SLASHed NP *articles about* (marked [3] in the tree) is well-formed. In fact, the gap-containing phrases in (4) *can* occur in object position:

(7) a. Sandy wrote articles about Kim.
→ Who$_i$ did Sandy write [articles about ____$_i$]?

b. Kim insisted that Sandy stole the potatoes.
→ [The potatoes]$_i$, Kim insisted [that Sandy stole ____$_i$].

We don't want to block the application of CELR, since it is needed for these instances of subextraction from objects. The ungrammaticality of (4) suggests that there is a constraint against SLASHed phrases appearing as subjects.

Parasitic gaps

A complication:

(8) Articles about [the president]$_i$ called him$_i$ a racist.
→ [Which president]$_i$ did [articles about ____$_i$] call ____$_i$ a racist?

Even though this sentence contains a gapped subject, it is grammatical (or at least less ungrammatical than (4a)). The subject-internal gap is somehow licensed by the co-referential object-internal gap. The subject gap is a **parasitic gap**, because it can only occur in the presence of another (independently grammatical) **host gap**.

(9) Subject Condition:
A lexical head's subject can be SLASHed only if one of its cmoplements is.

Parasitic gaps (cont.)

This can be stated as a constraint on *valence*:

(10)

$$\left[\textit{valence}\ \left[\text{SUBJ}\ \left\langle\left[\text{INHER}\mid\text{SLASH}\ \{[\ ]\}\right]\right\rangle\right]\right] \Rightarrow \left[\begin{array}{l}\textit{valence}\\ \text{SUBJ}\ \left\langle\left[\text{INHER}\mid\text{SLASH}\ \{[\ ]\}\right]\right\rangle\\ \text{COMPS}\ \left\langle\ldots\left[\text{INHER}\mid\text{SLASH}\ \{[\ ]\}\right]\ldots\right\rangle\end{array}\right]$$

Referring back to the tree in (6), we can see that the lexical entries for *were* and *revealing* both fail to meet the Subject Condition: their common subject, articles about, contains a gap, but their COMPS lists contain no licensing host gaps.

Parasitic gaps (cont.)

It is standardly assumed that a parasitic gap must be co-referential with its host gap, as it is in (8). The Subject Condition does not enforce this, however, and there is some evidence that a subject gap can be parasitic on a non-coreferential complement gap:

(11) [People that sensitive]$_i$, I never know [which topics]$_j$ jokes about ____$_j$ are likely to offend ____$_i$. (P&S 94, p. 199)

In (11), the VALENCE value of *are* satisfies the Subject Condition:

(12)

$$\begin{bmatrix} \text{SUBJ} & \left\langle \begin{bmatrix} \textit{“jokes”} \\ \text{INHER} \mid \text{SLASH} & \left\{ \left[\textit{“topics”} \right] \right\} \end{bmatrix} \right\rangle \\ \text{COMPS} & \left\langle \text{AdjP} \begin{bmatrix} \textit{“likely”} \\ \text{INHER} \mid \text{SLASH} & \left\{ \left[\textit{“people”} \right] \right\} \end{bmatrix} \right\rangle \end{bmatrix}$$

Since the two gaps contribute two distinct *local* objects, we get two fillers: *people that sensitive* and *which topics*. In contrast, the two gaps in (8) both contribute the same *local* object, and since INHER | SLASH is set-valued, the dependency can only be registered once. One filler—*which president*—binds both gaps.

Parasitic gaps (cont.)

The Subject Condition does not account for the licensing of parasitic gaps in adjuncts:

(13) a. John sinned unforgivably by mocking Mary.
→ *[Who]$_i$ did John sin unforgivably by mocking ____$_i$?

b. John insulted Mary$_i$ by mocking her$_i$.
→ [Who]$_i$ did John insult ____$_i$ by mocking ____$_i$?

In (13b) the presence of a co-referential object gap seems to license the adjunct gap, which is ungrammatical on its own, as in (13a). On the other hand, there are grammatical examples of non-parasitic adjunct gaps:

(14) That's [the symphony]$_i$ that Schubert died without finishing ____$_i$. (P&S 94, p. 191)

More evidence against a general constraint against adjunct subextraction can be found in Hukari and Levine (1995). Examples like the one in (14) can be handled with the extraction rules we have seen already. In particular, the NONLOCAL Feature Principle applies to head-adjunct phrases as usual.

Other island constraints

Ross (1967): certain phrases can be characterized as "islands" from which words and phrases cannot escape.

Complex NP Constraint

A complex NP is an NP that contains a clause:

- Factitive complement

  (15) a. evidence [that aliens have visited Earth]
  b. a suggestion [that everyone welcomed]

- Relative clause

  (16) a. the rock [underneath which we hid the treasure]
  b. someone [who can cook Mexican food]

Complex NP Constraint: Complex NPs are islands. I.e., nothing can be extracted out of the clause in a complex NP.

Complex NP Constraint (cont.)

P&S question the claim for factitive NPs:

(17) a. [Which country]$_i$ did you hear rumors that we had infiltrated ____$_i$? (P&S 94, p. 206)

b. [Which planets]$_i$ have you found evidence that aliens have visited ____$_i$?

These are therefore treated as grammatical in HPSG. Both of these examples can be analyzed via the CELR.

Complex NP Constraint (cont.)

For relative clauses, the restriction is clearer:

(18) a. * [How much treasure]$_i$ did Kim discover the rock underneath which we hid ____$_i$?

b. * [What dishes]$_i$ do we need to hire someone who can cook ____$_i$?

These examples are accounted for through the definitions of the null relativizers that head relative clauses. See P&S 94, Chapter 4 for discussion. But note that the revised analysis of Sag (1997) no longer employs null relativizers, and the issue of the Complex NP Constraint is not mentioned there. The ungrammaticality of the examples in (18) would follow from a constraint stating that the SLASH set is maximally singleton (although this raises other problems...)

Coordinate Structure Constraint

Ross (1967): Coordinated phrases are islands.

A conjunct cannot be extracted:

(19) They met Kim and Sandy.
→ *Who$_i$ did they meet Kim and ____$_i$?
→ *Who$_i$ did they meet ____$_i$ and Sandy?

Since individual conjuncts are not subcategorized for, they do not appear on any COMPS list. Therefore the CELR cannot target a conjunct.

Coordinate Structure Constraint (cont.)

Subextraction out of a conjunct is prohibited:

(20) Kim drank a beer and enjoyed the party. →

a. * What$_i$ did Kim drink ____$_i$ and enjoy(ed) the party?

b. * What$_i$ did Kim drink a beer and enjoy ____$_i$?

But "across-the-board" subextraction is grammatical—i.e., the same information must be missing from all conjuncts:

(21) a. Sandy drank [a beer]$_i$ and enjoyed it$_i$.

b. → What$_i$ did Sandy drink ____$_i$ and enjoy ____$_i$?

Coordinate Structure Constraint (cont.)

Coordinate structures are analyzed as non-headed phrases in HPSG. The features of a coordinate structure are (in part) determined by the following constraint:

(22) Coordination Principle
The CATEGORY and NONLOCAL features of a coordinate structure subsume those of each conjunct daughter.

In the absence of any other constraints, this principle allows *any* two structures to be coordinated. After all, the mother could simply be specified as [ ], which subsumes everything. In practice, however, the construction in which the coordinate structure appears will instantiate certain features on it, and these features must then appear on every conjunct daughter, by the Coordination Principle.

Coordinate Structure Constraint (cont.)

For example, the VP *drank a beer and enjoyed it* in (21a), by virtue of appearing in that sentence, must have at least the following specifications:

(23)

$$\begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{VFORM} & \textit{fin} \\ \text{INV} & - \end{bmatrix} \\ & \overset{\textit{verb}}{} \\ \text{SUBJ} & \left\langle \begin{bmatrix} \textit{“Sandy”} \end{bmatrix} \right\rangle \\ \text{COMPS} & \langle\ \rangle \end{bmatrix} \\ \text{NONLOC} \mid \text{INHER} \mid \text{SLASH} & \{\ \} \end{bmatrix}$$

This information must be inherited by both conjuncts; i.e., *drank a beer* and *enjoyed it* must both be finite, SUBJ-unsaturated VPs, containing no gap.

Coordinate Structure Constraint (cont.)

In (21b), the coordinate VP *drink and enjoy* has the following specifications:

(24)

$$\left[\begin{array}{ll} \text{CAT} & \left[\begin{array}{ll} \text{HEAD} & \left[\begin{array}{ll} \text{VFORM} & bse \\ \text{INV} & - \end{array}\right] \\ & verb \\ \text{SUBJ} & \left\langle \left[\textit{“Sandy”} \right] \right\rangle \\ \text{COMPS} & \langle\ \rangle \end{array}\right] \\ \text{NONLOC} \mid \text{INHER} \mid \text{SLASH} & \left\{ \left[\textit{“what”} \right] \right\} \end{array}\right]$$

By the Coordination Principle, both conjuncts must contain the same gap. This principle therefore ensures that either all conjuncts are SLASHed (across-the-board extraction) or none of them are.

Extraction without Lexical Rules

Bouma, Malouf, and Sag (1998) present two objections to the HPSG3 theory of extraction we have been considering:

- Inelegant to have separate mechanisms for complement, subject, and adjunct extraction. Linguistic data from a wide variety of languages suggest that extraction is a more unified phenomenon.

  In particular, the analysis of *Who adores John?* as an ordinary head-subject phrase, rather than a head-filler phrase, is difficult to reconcile with evidence that they pattern with other *wh*-questions with respect to many grammatical processes.

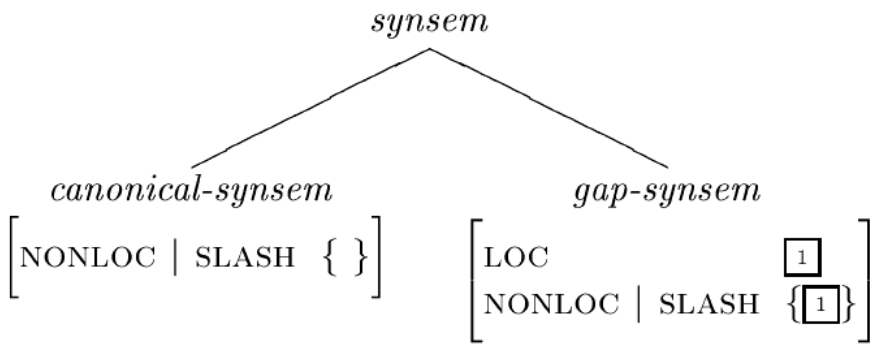
  Similarly, adjunct extraction is seen to behave much like complement extraction cross-linguistically, calling into question the use of separate mechanisms in P&S 94.

- Lexical rules are problematic in general, and seem particularly inappropriate for analyzing extraction.

Constraint-based head-driven extraction

- Two subsorts of *synsem*

(25)



Note the similarity between *gap-ss* and the definition of trace in the analysis of HPSG2. The crucial difference is that there is no lexical entry with a SYNSEM value of type *gap-ss*:

(26) Canonicality

$$sign \Rightarrow \left[\text{SYNSEM} \quad canon\text{-}ss\right]$$

Objects of type *gap-ss* only show up in the valence and ARG-ST lists of other signs, and these items are never realized syntactically.

Constraint-based head-driven extraction (cont.)

- A new feature DEPENDENCY-STRUCTURE is introduced for *words*; it takes a value of type *list*(*synsem*) and it keeps track of all the syntactic dependents of a word. Two relevant constraints:

(27) a. Argument realization[1]

$$word \Rightarrow \begin{bmatrix} \text{ARG-ST} & \boxed{1} \\ \text{DEPS} & \boxed{1} \oplus list(\text{“}adjuncts\text{”}) \end{bmatrix}$$

b. Dependent Realization

$$word \Rightarrow \begin{bmatrix} \text{SUBJ} & \langle \boxed{1}\ synsem \rangle \\ \text{COMPS} & \boxed{2} \ominus list(gap\text{-}ss) \\ \text{DEPS} & \langle \boxed{1} \rangle \oplus \boxed{2} \end{bmatrix}$$

Argument Realization says that the SYNSEM values of the core syntactic dependents of a word (i.e., subject, specifier, complements) appear on the DEPS list. In addition to these, the SYNSEM values of zero or more adjuncts can appear at the end of the DEPS list. Dependent Realization says that the first member of the DEPS list is the subject, while the COMPS list contains the rest of the members of DEPS minus any that happen to be of type *gap-ss*. Note that in combination with Argument Realization, this means that adjuncts will appear on COMPS, as long as they are of type *canon-ss*.

[1] BM&S use $\oplus$ to represent list append. I.e., $\boxed{1} \oplus \boxed{2}$ is equivalent to append($\boxed{1}$, $\boxed{2}$).

Constraint-based head-driven extraction (cont.)

- Any element of a word's DEPS list can be instantiated as type *gap-ss*. This takes the place of all three lexical rules CELR, SELR, AELR. A word amalgamates the SLASH values of all its dependents:

(28) SLASH Amalgamation Constraint (SLAC)

$$word \Rightarrow \begin{bmatrix} \text{DEPS} \left\langle \begin{bmatrix} \text{SLASH} & \boxed{1} \end{bmatrix}, \ldots, \begin{bmatrix} \text{SLASH} & \boxed{n} \end{bmatrix} \right\rangle \\ \text{BIND} \quad \boxed{0} \\ \text{NONLOC} \mid \text{SLASH} \quad (\boxed{1} \cup \ldots \cup \boxed{n}) - \boxed{0} \end{bmatrix}$$

The feature BIND is used by the words that introduce *tough*-constructions. As we have seen, these words subcategorize for SLASHed complements but then they lexically discharge the dependency so it never gives rise to a head-filler phrase. (See section 4.3 for an example.)

Constraint-based head-driven extraction (cont.)

The SLASH specification of a head daughter is passed up to its mother:

(29) SLASH Inheritance Principle (SLIP)

$$(\textit{hd-ph} \wedge \neg \textit{hd-fill-ph}) \Rightarrow \begin{bmatrix} \text{NONLOC} \mid \text{SLASH} & \boxed{1} \\ \text{HD-DTR} \mid \text{NONLOC} \mid \text{SLASH} & \boxed{1} \end{bmatrix}$$

SLAC and SLIP replace the NONLOCAL Feature Principle. Note that the percolation of NONLOCAL information is mediated by the head. This is why BM&S refer to their analysis as "head-driven extraction."

Constraint-based head-driven extraction (cont.)

- We require a new definition for head-filler phrases.

(30) Head-Filler Schema[2]

$$\textit{hd-fill-ph} \Rightarrow \begin{bmatrix} \text{NONLOC} \mid \text{SLASH} & \boxed{2} \uplus \boxed{3} \\ \text{HD-DTR} \mid \text{SLASH} & \boxed{2} \uplus \{\boxed{1}\} \\ \text{FILL-DTR} & \begin{bmatrix} \text{LOC} & \boxed{1} \\ \text{SLASH} & \boxed{3} \end{bmatrix} \end{bmatrix}$$

Unlike the other ID schemas, this one must be stated as a constraint on PHRASE rather than on *con-struc*—a weakness that should be eliminated (e.g., by adding a disjunctive clause to SLIP).

Assuming that (1) filler daughters cannot themselves be SLASHed, and (2) the SLASH set is maximally singleton[3], we have the following simplified constraint:

(31)

$$\textit{hd-fill-ph} \Rightarrow \begin{bmatrix} \text{NONLOC} \mid \text{SLASH} & \{\,\} \\ \text{HD-DTR} \mid \text{SLASH} & \{\boxed{1}\} \\ \text{FILL-DTR} & \begin{bmatrix} \text{LOC} & \boxed{1} \\ \text{SLASH} & \{\,\} \end{bmatrix} \end{bmatrix}$$

I.e., the filler daughter's LOCAL value matches the single member of the head daughter's SLASH set. The filler fills the gap.

[2]The symbol $\uplus$ represents disjoint set union, which is the same as set union, but only defined if the arguments are disjoint.

[3]These assumptions hold for the vast majority of (English) examples.

Adjunct extraction

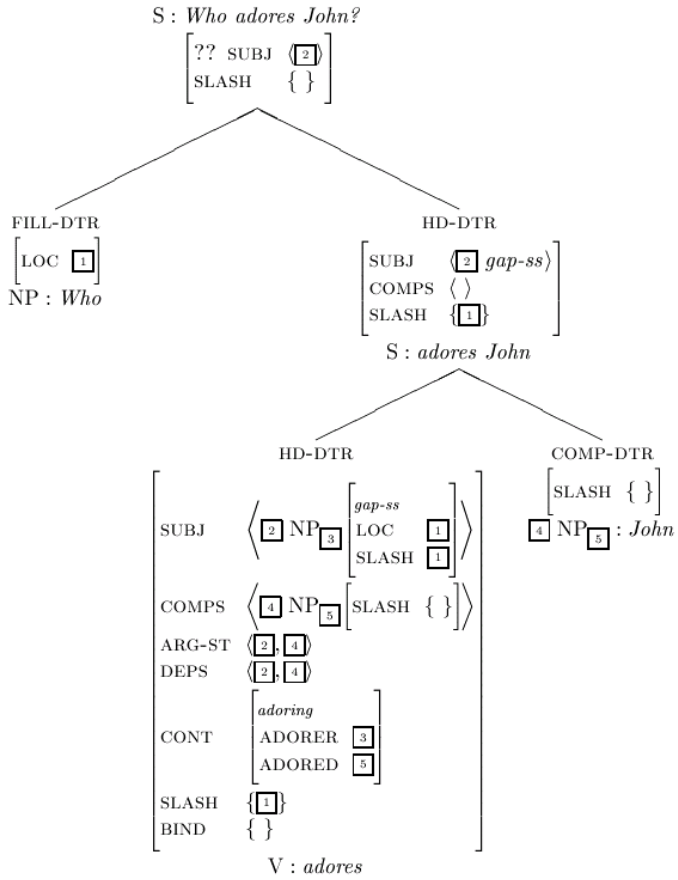
Under this analysis, adjuncts behave exactly like complements; in particular they have the same extraction properties. The only difference between the two types of dependent is that the *synsem* values of complements appear on ARG-ST while those of adjuncts do not. This has consequences for binding, which is defined on ARG-ST.

Subject extraction

Subject extraction has more in common with complement extraction in this analysis. If the first member of the DEPS list happens to be instantiated as an object of type *gap-ss*, then there will be a gap in subject position, which will eventually be associated with a filler (or otherwise bound). The same holds true for the other members of DEPS, except that those correspond to objects and adjuncts.

In particular, extraction of matrix clause subjects is treated as a UDC, in contrast to P&S 94:

(32) Subject extraction: *Who adores John?*



Subject extraction (cont.)

Two problems:

- In the head-filler phrase, the head daughter is still SUBJ-unsaturated. There is no way for this valence requirement to be discharged via the ValP, because there is no sign with the feature [SUBJ : ⟨[2] *gap-ss*⟩] (see (26)). Either the Head-Filler Rule in (30) must be further specified to empty the head daughter's SUBJ list explicitly, or else the entire phrase must simply be allowed to have an unsaturated SUBJ list.

- BM&S propose no way to block the "subject in situ," non-UDC analysis assumed by P&S for this sentence. In principle, it is still available, giving rise to a spurious ambiguity for all subject *wh*-questions. The same point can be made for sentence initial adjuncts that modify matrix clauses; these are now ambiguous between a head-filler analysis or a head-adjunct analysis.

Subject extraction (cont.)

That-trace violations in embedded clauses are blocked simply by assuming that the complementizer *that* can combine only with a SUBJ-saturated sentence:

(33)

$$\begin{bmatrix} \text{PHON} & \langle \textit{that} \rangle \\ \text{SPEC} & \text{S}\begin{bmatrix} \text{SUBJ} & \langle\ \rangle \end{bmatrix} \end{bmatrix}$$

As we saw in (32) above, the SLASHed sentence *adores John* is SUBJ-unsaturated. It carries the feature [SUBJ : ⟨*gap-ss*⟩] and therefore cannot combine with the complementizer:

(34) * Who does Kim think $\begin{bmatrix} \text{SPEC} & \boxed{x}\begin{bmatrix} \text{SUBJ} & \langle\ \rangle \end{bmatrix} \end{bmatrix}$ that $\boxed{y}\begin{bmatrix} \text{SUBJ} & \langle \textit{gap-ss} \rangle \end{bmatrix}$ adores John ?

The sentence is ungrammatical because $\boxed{x}$ and $\boxed{y}$ do not unify.

Extraction out of subjects is still assumed to conform to some version of P&S's Subject Condition stated in (9). The treatment of parasitic gaps therefore remains the same, although the modified view of subject extraction is shown to result in a better account of some parasitic gap data.

Tough constructions

The analysis of *easy*-type adjectives is directly imported into the BM&S analysis.

(35) a. Lexical entry for *impossible*, P&S style

$$\begin{bmatrix} \text{PHON} & \langle impossible \rangle \\ \text{HEAD} & adj \\ \text{SUBJ} & \left\langle \text{NP}_{\boxed{1}} \right\rangle \\ \text{COMPS} & \left\langle (\text{PP}[for]), \text{VP} \begin{bmatrix} \text{VFORM} & inf \\ \text{SUBJ} & \langle \text{NP} \rangle \\ \text{INHER} \mid \text{SLASH} & \{\boxed{2}\} \end{bmatrix} \right\rangle \\ \text{TO-BIND} \mid \text{SLASH} & \left\{ \boxed{2}\ \text{NP}[acc] : ppro_{\boxed{1}} \right\} \end{bmatrix}$$

Tough constructions (cont.)

b. Lexical entry for *impossible*, BM&S style

$$\begin{bmatrix}
\text{PHON} & \langle impossible \rangle \\
\text{HEAD} & adj \\
\text{SUBJ} & \left\langle \boxed{1}\ \text{NP}_{\boxed{2}} \right\rangle \\
\text{COMPS} & \left\langle (\boxed{3}\ \text{PP}[for]), \boxed{5}\ \text{VP} \begin{bmatrix} \text{VFORM} & inf \\ \text{SUBJ} & \langle \text{NP} \rangle \\ \text{SLASH} & \{\boxed{4}\} \end{bmatrix} \right\rangle \\
\text{ARG-ST} & \langle \boxed{1}, (\boxed{3}), \boxed{5} \rangle \\
\text{DEPS} & \langle \boxed{1}, (\boxed{3}), \boxed{5} \rangle \\
\text{SLASH} & \{\boxed{4}\} - \{\boxed{4}\} = \{\ \} \\
\text{BIND} & \left\{ \boxed{4}\ \text{NP}[acc] : ppro_{\boxed{2}} \right\}
\end{bmatrix}$$

References

Bouma, G., Malouf, R., & Sag, I. (1998). A unified theory of complement, adjunct, and subject extraction. In G. Bouma, G.-J. M. Kruijff, & R. T. Oehrle (Eds.), *Proceedings of the joint conference on Formal Grammar, Head-Driven Phrase Structure Grammar, and Categorial Grammar* (pp. 83–97).

Hukari, T. E., & Levine, R. D. (1995). Adjunct extraction. *Journal of Linguistics*, *31*, 195–226.

Ross, J. R. (1967). *Constraints on variables in syntax.* Unpublished doctoral dissertation, MIT. (Published as *Infinite Syntax!* Norwood, NJ: ABLEX, 1986)

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