

Honors Thesis Presentation

Chris Callison-Burch

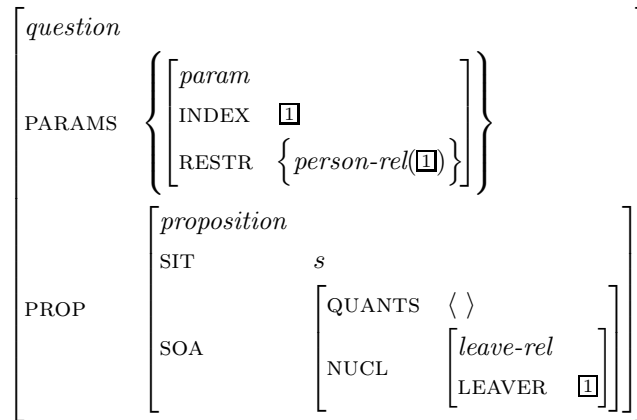
April 4, 2000

1 Some good questions:

- Why are grammar-based Natural Language Processing systems useful?
- What are some of the syntactic phenomena involved with questions?
- How are they modeled in a formal theory or implementation?

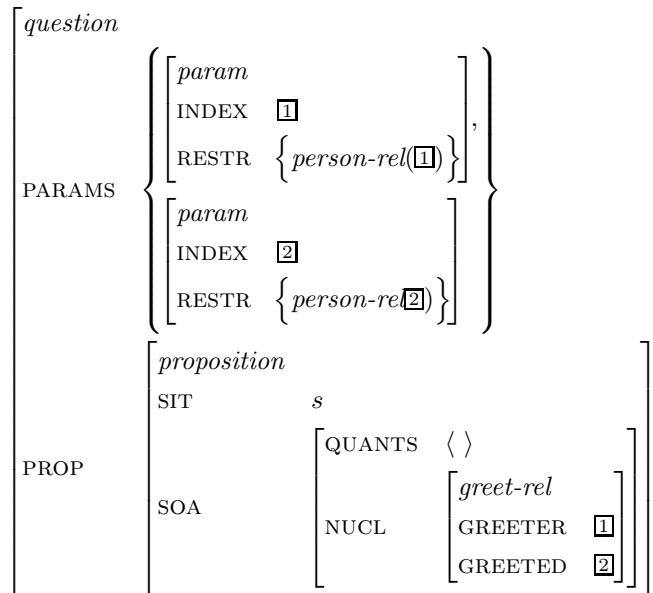
2 An example of the CONTENT of a question:

(1) *Who left?* \mapsto



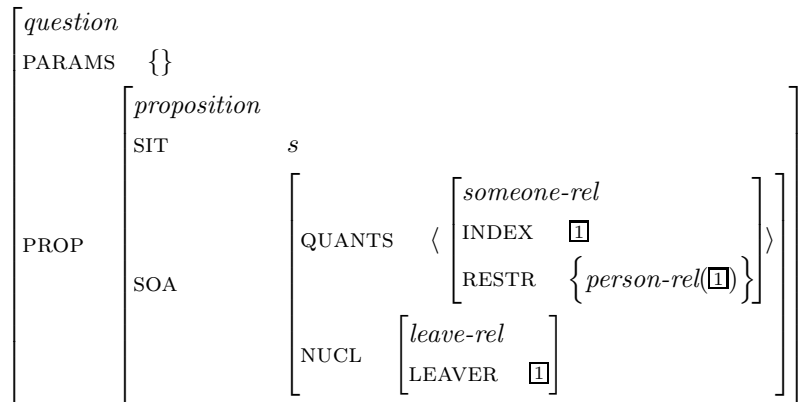
3 Another example of a question's CONTENT:

(2) *Who greeted who?* \mapsto



4 One more example:

(3) *Did someone leave?* \mapsto



5 Types of question constructions:

- Polar interrogative Constructions - ‘yes/no’ questions
“Was Agent Mulder really abducted by aliens?”
- Subject *Wh*-Interrogatives
“Who watched last week’s X-Files episode?”
- Non-subject *Wh*-Interrogatives

“How many alien-human hybrids did Dana Scully unknowingly mother?”

- *In situ* Questions - ‘reprise’ or ‘echo’ questions
“Agent Mulder believes WHAT?”

6 Syntactic phenomena relevant to questions:

- Extraction
“Which government conspiracy did Mulder uncover ___?”
- Inversion
“Did Cancer Man die?” vs. *“Cancer Man did die.”*
- Sensitivity to the presence of Wh-Words
“Whose funds did Mulder waste?”
 * *“The FBI’s funds did Mulder waste?”*

7 Examples of filler-gap dependencies:

The term “long-distance dependency” describes syntactic phenomena in which a constituent is dislocated from the place that it is normally realized, and instead occurs at a potentially unbounded distance from that location.

- (4) a. [These bagels]_i, I like ____i. (topicalization)
 b. [These bagels]_i, they say they like ____i. (topicalization)
 c. [Whose bagels]_i do you like ____i? (*wh*-interrogative)
 d. [From whom]_i did you buy these bagels ____i? (*wh*-interrogative)
 e. [What great bagels]_i they bought ____i! (*wh*-exclamative)

8 The head-filler phrase type:

All the above constructions inherit from the same parent, the head-filler phrase:

- (5) *hd-fill-ph*:

$$\left[\text{SLASH } \boxed{\Sigma_2} \right] \rightarrow \left[\text{LOC } \boxed{\Gamma} \right], \text{ H } \left[\begin{array}{l} \textit{phrase} \\ \text{SUBJ } \langle \rangle \\ \text{HEAD } \textit{verb} \\ \text{SLASH } \{ \boxed{\Gamma} \} \uplus \boxed{\Sigma_2} \end{array} \right]$$

Note that in the version of HPSG presented in the Ginzburg and Sag, as well as my implementation, maximal subtypes of *phrase* are treated as grammar rules. So you can essentially read the above as a context free grammar production.

9 The feature SLASH:

The feature SLASH is used to store all arguments which are not canonically realized by the head-complement constructions.

(6) SLASH Amalgamation Constraint

$$word \Rightarrow \left[\text{SS} \left[\text{LOC|CAT} \left[\text{ARG-ST} \left\langle [\text{SLASH } \Sigma_1], \dots, [\text{SLASH } \Sigma_n] \right\rangle \right] \right] \right]$$

$$\left[\text{SLASH } (\Sigma_1 \cup \dots \cup \Sigma_n) \right]$$

The SLASH Amalgamation Constraint guarantees that a word stores all of its arguments which are not canonically realized, i.e. which participate in filler-gap construction.

10 Non-canonical synsems

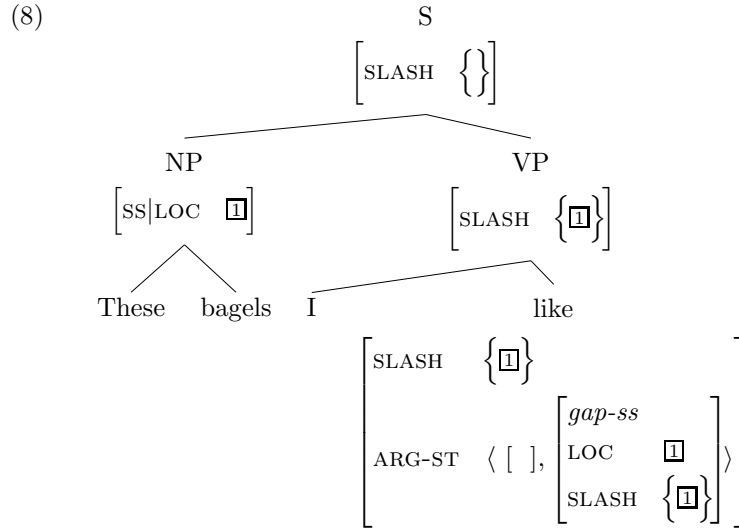
Arguments are put into SLASH by being marked as a type of non-canonical synsem, *gap-ss*, which has the following constraint:

$$(7) \quad gap-ss \Rightarrow \left[\text{LOC} \quad \boxed{\square} \right]$$

$$\left[\text{SLASH} \quad \left\{ \boxed{\square} \right\} \right]$$

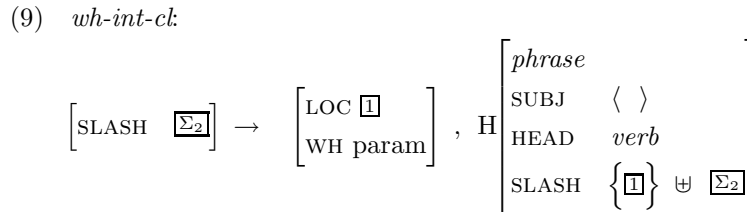
This is what causes an filler-gap constituent to be added to the SLASH list of the word that it is an argument of.

11 An illustration:



12 *Wh*-interrogative clause constructions

The type for *wh*-interrogatives inherits from *hd-fill-ph* and adds the additional constraint that the filler daughter have the value *param* for its WH feature:



This guarantees that the filler daughter of a *wh-int-cl* be a *wh*-word, because all others words are marked [WH *wh-none*].

13 The WH Agreement Principle

Since the *wh*-word can be properly contained within the filler constituent of a *wh*-interrogative construction, the following principle is used to percolate the WH value up to the phrase level:

(10) WH Agreement Principle:

$$word \Rightarrow \left[SS \left[\begin{array}{l} LOC|CAT|SPR \quad \langle [WH \quad \boxed{I}] \rangle \\ WH \quad \boxed{I} \end{array} \right] \right]$$

This correctly predicts the following data:

- (11) a. Whose books did Kim read?
b. Whose pictures of Sandy did Kim like?
c. *Sandy's pictures of whom did Kim like?