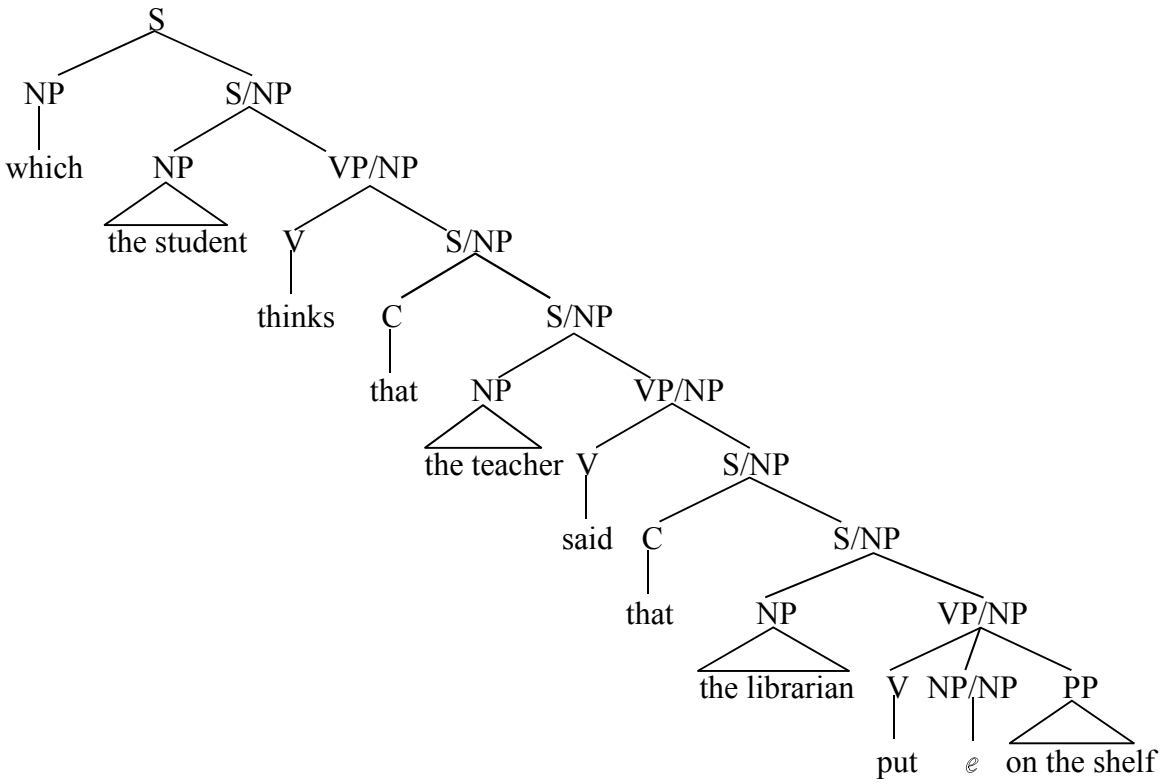
Phrase-Structure Grammar, p. 1

LDD constructions have traditionally been seen as the paradigm case of a construction that cannot be handled by context-free phrase structure rules. The PSG literature has therefore focused on this issue. The earlier version of modern PSG is called Generalized Phrase Structure Grammar (GPSG). In the mid-to-late 1980s GPSG morphed into a theoretical framework which was sufficiently different to get its own name: Head-driven Phrase Structure Grammar (HPSG). For a textbook-level introduction to HPSG, see Sag and Wasow (1999).

Early concept (GPSG: Gazdar 1981; Gazdar, Klein, Pullum, and Sag 1985): The transformational conception in terms of displacement is essentially correct, but it need not be done transformationally. Instead, define a new type of category ("slashed" category), where the category name indicates that something is missing. E.g. "S/NP" means "S with a missing NP".



The slash notation carries down the information about the filler locally, one node at a time. This is a feature, but not an ordinary feature that goes from phrase to head (a HEAD feature); rather, it is a non-HEAD feature (called a FOOT feature in GPSG.) The semantic interpretation of the trace makes it coreferential with the filler. Slashed categories are introduced and eliminated by phrase structure rules such as:

S → XP S/XP

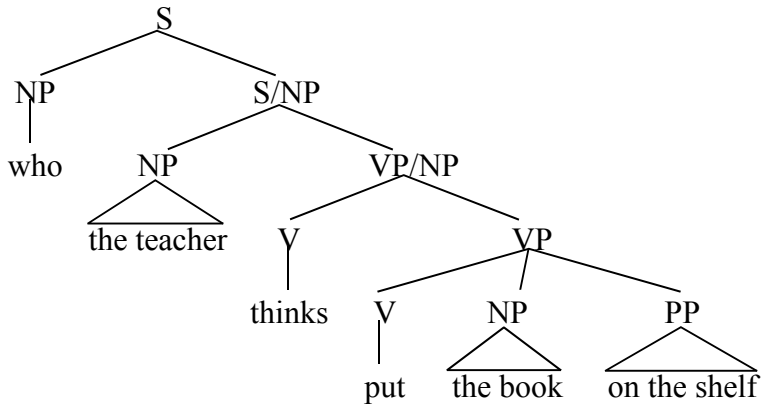
XP/XP → *e*

Phrase-Structure Grammar, p. 2

The introduction of traces into the transformational account made non-transformational versions of *wh* movement possible, since the trace (the XP/XP element in GPSG) encoded the "deep structure" position of the *wh* element in surface syntactic structure. This is ironic, since the existence of traces has become very controversial in non-transformational theories.

Gazdar (1981) argued that subject extraction does not involve a trace, but rather a bare VP.

the librarian [who the teacher thinks put the book on the shelf]



HPSG

In HPSG, the GPSG analysis has been further developed, but the basic idea—a gap whose properties are encoded in a SLASH feature—remains.

Linguistic items are formalized as signs, consisting of feature structures. Lexical entries and syntactic rules are constraints on these feature structures, represented graphically as an attribute-value matrix (AVM). For example, the lexical representation of the word *put* is something like:

Phrase-Structure Grammar, p. 3

$$\begin{bmatrix} \textit{lexeme} \\ \text{PHONOLOGY} \quad \langle \text{put} \rangle \\ \text{SYNSEM} \begin{bmatrix} \textit{synsem} \\ \text{LOC} \begin{bmatrix} \textit{local} \\ \text{CAT} \begin{bmatrix} \textit{cat} \\ \text{HEAD} \quad \textit{verb} \\ \text{SUBJ} \quad \langle \boxed{1} \rangle \\ \text{COMPS} \quad \langle \boxed{2}, \boxed{3} \rangle \end{bmatrix} \\ \text{CONT} \begin{bmatrix} \textit{put-rel} \\ \text{PUTTER} \quad \boxed{1} \\ \text{PUTTEE} \quad \boxed{2} \\ \text{LOCATION} \quad \boxed{3} \end{bmatrix} \end{bmatrix} \end{bmatrix} \\ \text{ARG-ST} \quad \langle \boxed{1}\,\text{NP}, \boxed{2}\,\text{NP}, \boxed{3}\,\text{PP} \rangle \end{bmatrix}$$

Each piece of the AVM is labeled with the type of linguistic object that it models. (More on this later.) This is usually the top line of the AVM, and italicized. (In earlier work in HPSG, it was a left-subscript to the AVM.)

A sign (in the Saussure sense) is an arbitrary sound-meaning pair. In HPSG, it includes the attributes PHONOLOGY and SYNSEM ("syntax-semantics"); in most current work, ARG-ST ("argument structure") is a separate attribute. Within the SYNSEM, the major feature structure is LOC ("local"), which contains the feature structures CAT ("category", including category in the traditional sense plus valence information) and CONT ("content", semantic information); it can also include discourse information (CONTEXT).

The same piece of feature structure sometimes occurs in more than one place in the overall structure. For example, the value of the subject of *put* appears as the value of the attribute SYMSEM|LOC|CAT|SUBJ, the value of the attribute SYNSEM|LOC|CONT|PUTTER (CONT is short for CONTENT), and the first argument in the ARG-ST. This is shown by boxed numbers, called tags.

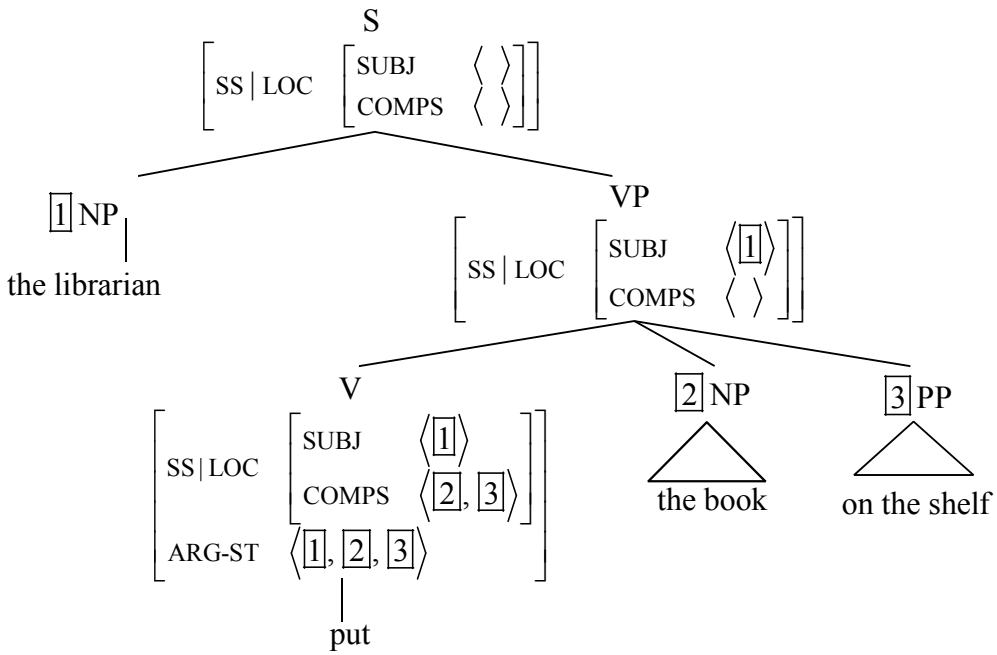
Some of the information in the lexical representation of *put* is redundant. For example, the ARG-ST list is identical to the concatenation of the SUBJ list and the COMPS list:

Argument Realization Principle (ARP)

$$\textit{word} \quad \Rightarrow \quad \begin{bmatrix} \text{SS} \mid \text{LOC} \mid \text{CAT} \begin{bmatrix} \text{SUBJ} \quad \langle \boxed{1} \rangle \\ \text{COMPS} \quad \langle \boxed{2} \rangle \end{bmatrix} \\ \text{ARG-ST} \quad \langle \boxed{1} \oplus \boxed{2} \rangle \end{bmatrix}$$

Phrase-Structure Grammar, p. 4

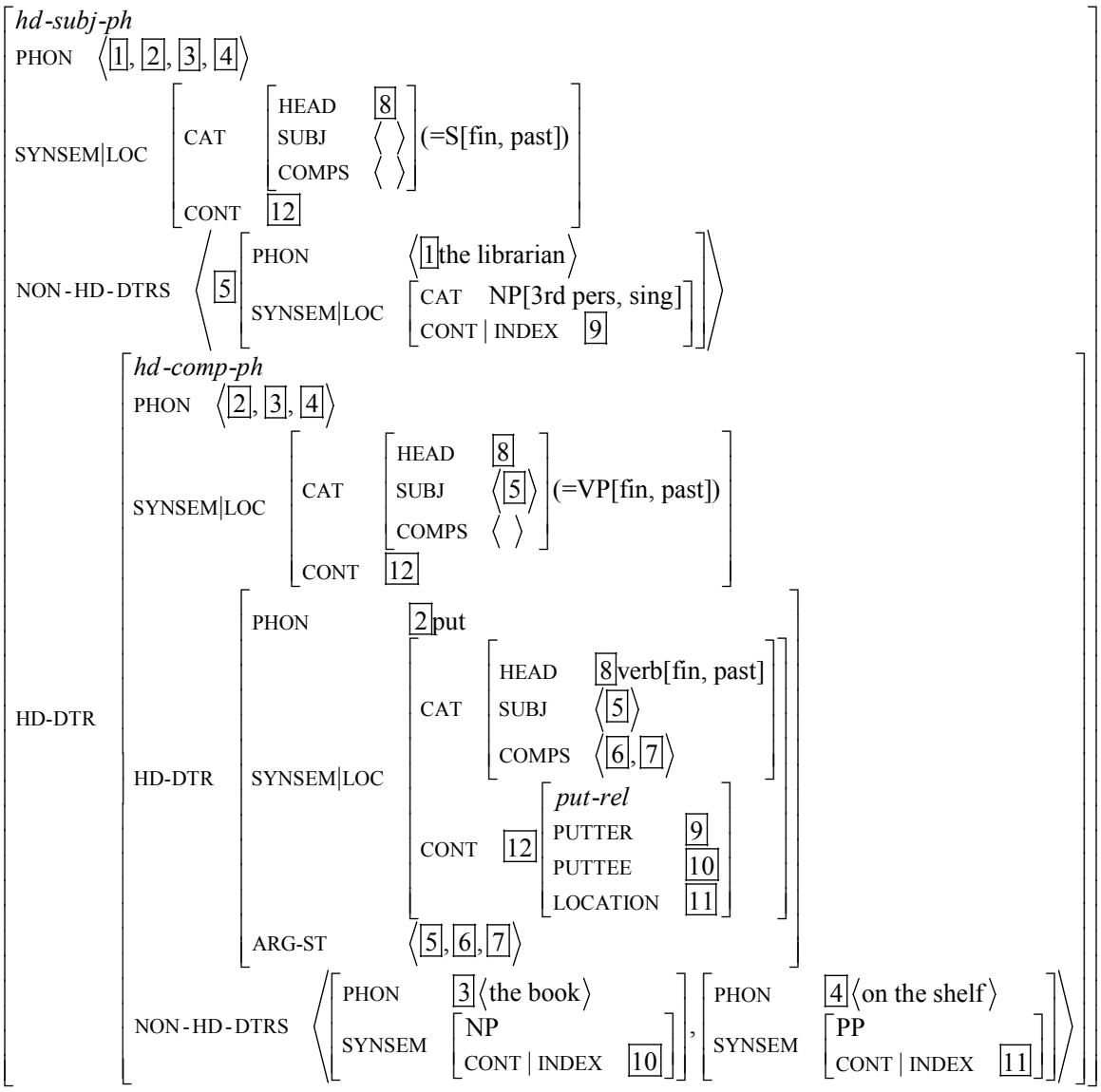
The ARG-ST list is the argument structure (θ grid) of the verb. The COMPS list is what is expressed as a sister of the head, and the SUBJ list is what is expressed as a sister of the VP. (Collectively, SUBJ and COMPS are called the valence lists.) The VP projected from *put* thus doesn't have the COMPS list, and the S doesn't have the SUBJ list either. So the structure of *the librarian put the book on the shelf* is (approximately):



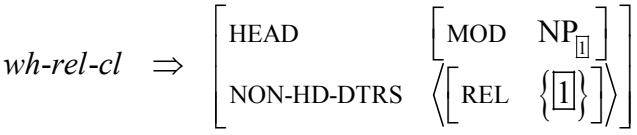
A central concept in HPSG is types of feature structures. These types are arranged in hierarchies. For example, parts of speech are arranged in hierarchies, where "intransitive verb" is a subtype of "verb", for example.

The typing of feature structures applies to phrases as well as words. In fact, trees are a shorthand for a phrasal feature structure. The constituent structure of *the librarian put the book on the shelf* can be represented as the following feature structure.

Phrase-Structure Grammar, p. 5



Phrase types are also arranged in type hierarchies, corresponding to construction types. For example, interrogative, exclamative, and relative clause are subtypes of the filler-gap construction type. This allows general constraints on all filler-gap constructions to be stated, as well as specific constraints on, say, relative clauses. For example, only relative clauses will have the following constraint (from Sag 1997) that the operator (the value of the REL feature) is coreferential with the NP that the relative clause modifies:

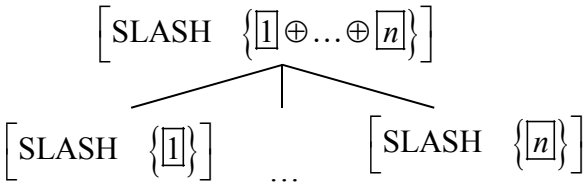


Phrase-Structure Grammar, p. 6

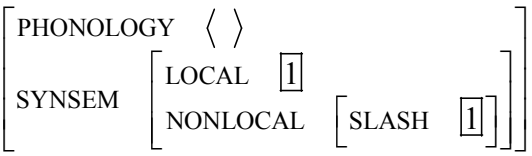
Since feature types, feature hierarchies, and the features themselves are the central tools of linguistic analysis in HPSG, it stands to reason that there will be differences in the exact feature structure in different HPSG analyses.

Now, on to SLASH. (The origin of the name SLASH for the LDD feature is the original GPSG notation. It is somewhat strange outside of its historical context. Sag and Wasow 1999 use the more mnemonic name GAP, but SLASH is standard.)

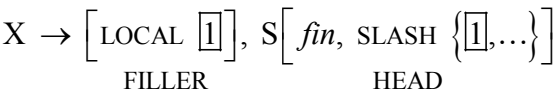
SLASH is also a feature, but a NONLOCAL feature. NONLOCAL features propagate through the tree differently. Sag and Wasow (1999) state the Nonlocal Feature Principle as follows:



In the original HPSG analysis (Pollard and Sag 1994a), as in the GPSG approach, the gap position was marked by a trace, an element with the following feature structure:

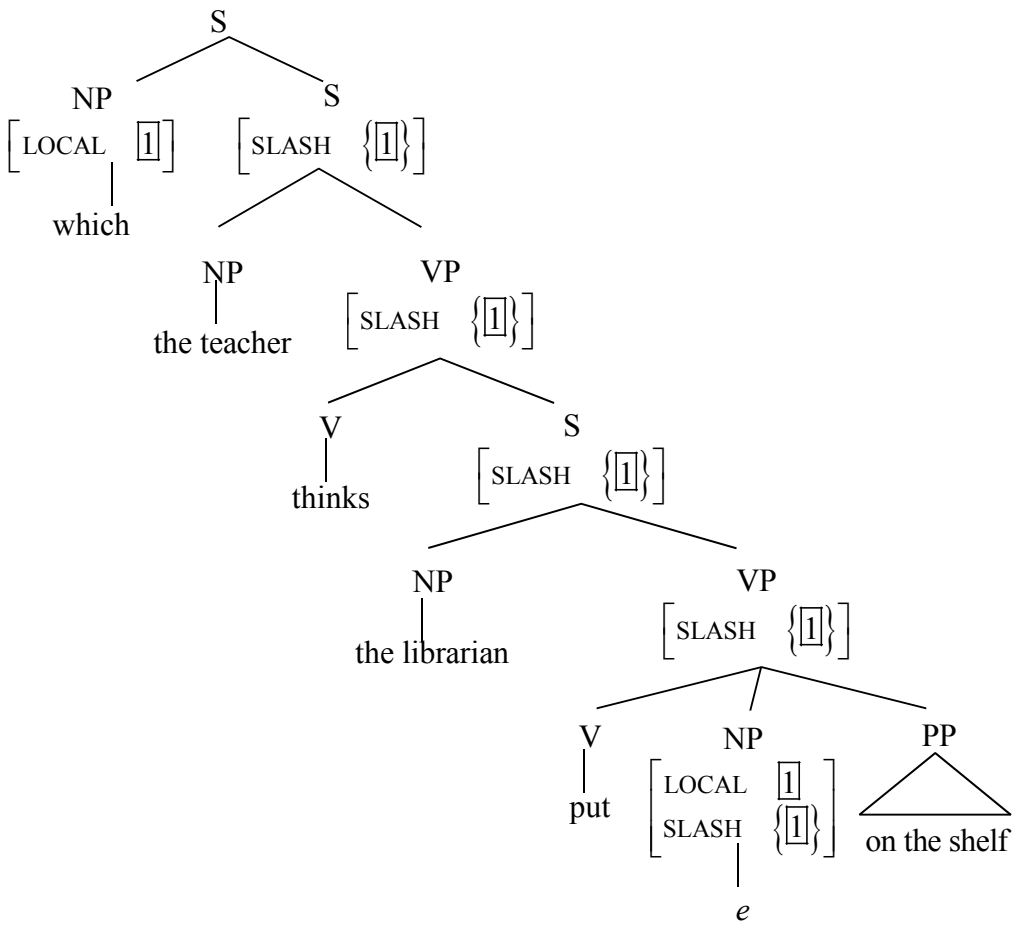


Constructions like topicalization and *wh* fronting are the result of the following phrase structure rule schema:



The SLASH feature is "bound off" by the filler introduced by this phrase structure schema.

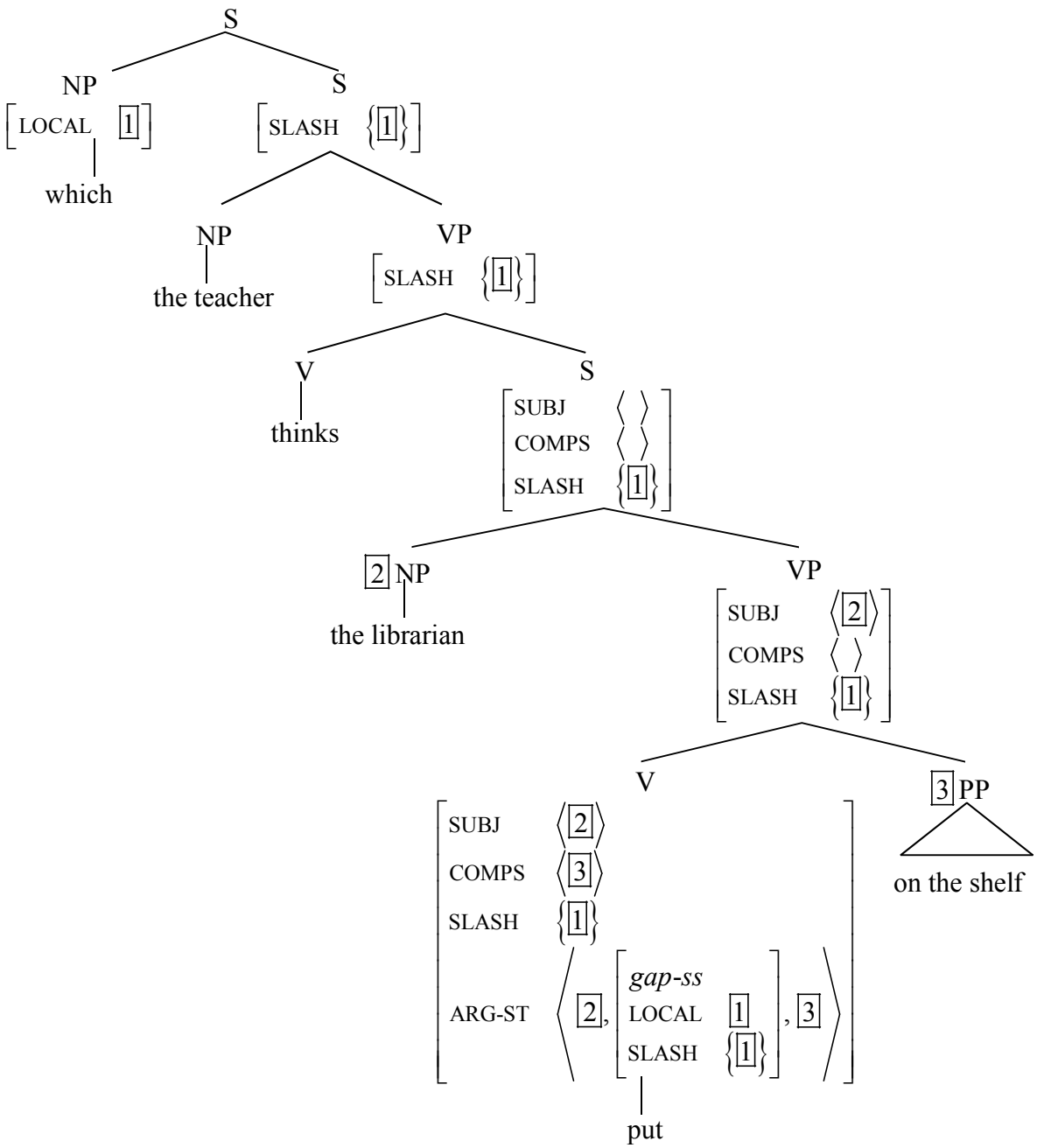
Since the SLASH feature propagates through the tree (by the Nonlocal Feature Principle), and the SLASH feature of the trace is identical to its LOCAL feature, the LOCAL feature of the filler and the LOCAL feature of the trace are (token-)identical.



(To make this work right, Pollard and Sag complicate this picture slightly by including two SLASH features, one as part of NONLOCAL|INHERITED and one as part of NONLOCAL|TO-BIND. In more recent work, the propagation of the SLASH feature is handled differently: the head amalgamates all the SLASH values of its dependents, and then the feature percolates to the phrasal node like all features of the head. See Sag 1997; Bouma, Malouf, and Sag 2001; Ginzburg and Sag 2000. We will use the older, more prevalent approach here, as in Sag and Wasow 1999, but without the INHERITED/TO-BIND distinction.)

Since Pollard and Sag 1994b, however, the standard HPSG approach to LDDs has been traceless. On such an approach, a gap does not appear in the tree. If the extracted element is a non-subject argument of the verb, this means that it is not a complement of the verb in the LDD sentence, and therefore does not appear in the COMPS list.

Showing relevant features:



The tree shows the SYNSEM of the extracted element as being of type *gap-ss*, as proposed by Bouma, Malouf, and Sag (2001) and Ginzburg and Sag (2000). Ordinary signs have SYMSEMs of type *canonical-ss*.

This traceless analysis, which is universally accepted in HPSG, thus analyzes the gap as an element which is present in the verb's argument structure but not its valence. The ARP, which defines the relationship between argument structure and valence, needs to be restated:

$$\textit{word} \Rightarrow \begin{bmatrix} \text{SS} \mid \text{LOC} \mid \text{CAT} & \begin{bmatrix} \text{LOC} & \begin{bmatrix} \text{SUBJ} & \langle \boxed{1} \rangle \\ \text{COMPS} & \langle \boxed{2} \ominus \boxed{3} \rangle \end{bmatrix} \\ \text{SLASH} & \{ \boxed{3} \} \end{bmatrix} \\ \text{ARG-ST} & \langle \boxed{1} \oplus \boxed{2} \rangle \end{bmatrix}$$

This does not account for subject extraction, for which Sag and Wasow (1999), following Pollard and Sag (1994b), state a lexical rule:

$$\begin{bmatrix} \textit{word} & \\ \text{SS} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{verb} \\ \text{FORM} \quad \text{fin} \end{bmatrix} \\ \text{SUBJ} & \langle [\ ] \rangle \end{bmatrix} \end{bmatrix} \Rightarrow \begin{bmatrix} \text{SS} & \begin{bmatrix} \text{SUBJ} & \langle\ \rangle \\ \text{SLASH} & \{ \boxed{1} \} \end{bmatrix} \\ \text{ARG-ST} & \langle \boxed{1}, \ldots \rangle \end{bmatrix}$$

The treatment of subject extraction has been controversial in the HPSG literature.

Tough movement, where there is no overt filler, is handled by having the *tough* predicate subcategorized for a slashed complement.

$$\begin{bmatrix} \text{ARG-ST} & \langle \text{NP}_{\boxed{1}}, (\text{PP}[\textit{for}]), \text{VP}[\textit{inf}, \text{SLASH} \{ \boxed{2} \text{NP}[\textit{acc}]_{\boxed{1}}, \ldots \}] \rangle \\ \text{NONLOCAL} \mid \text{SLASH} & \{ \boxed{2} \} \end{bmatrix}$$

The subscripted tag here means that just CONTENT|INDEX is shared, not the entire LOCAL. This is to account for the lack of connectivity.

Similarly, non-*wh* relatives involve the following constraint (Sag 1997; it modifies N′ rather than NP to account for the fact that non-*wh* relatives have to precede *wh* relatives):

$$\textit{non-wh-rel-cl} \Rightarrow \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{MOD} & \text{N}'_{\boxed{1}} \end{bmatrix} \\ \text{HD-DTR} & \left\langle \begin{bmatrix} \text{SLASH} & \{ \text{NP}_{\boxed{1}} \} \end{bmatrix} \right\rangle \end{bmatrix}$$

(For more on relative clauses, see Sag 1997. For more on interrogatives, see Ginzburg and Sag 2000.)

Pied piping involves the propagation of another NONLOCAL feature: QUE (or WH) in the case of interrogatives and REL in the case of relatives. This feature propagates through the fronted phrase.

References

Bouma, Gosse, Robert Malouf, and Ivan A. Sag (2001) "Satisfying Constraints on Extraction and Adjunction." *Natural Language and Linguistic Theory* 19: 1–65.

Gazdar, Gerald (1981) "Unbounded Dependencies and Coordinate Structure." *Linguistic Inquiry* 12: 155–184.

Gazdar, Gerald, Ewan Klein, Geoffrey K. Pullum, and Ivan A. Sag (1985) *Generalized Phrase Structure Grammar*. Oxford: Blackwell. **P 151 G45**

Ginzburg, Jonathan, and Ivan A. Sag (2000) *Interrogative Investigations: The Form, Meaning, and Use of English Interrogatives*. Stanford, Calif.: CSLI Publications. **PE 1395 G56**

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

Pollard, Carl, and Ivan A. Sag (1994a) *Head-Driven Phrase Structure Grammar*. Chicago: University of Chicago Press. [Chapters 1–8]. **P 158.4 P65**

Pollard, Carl, and Ivan A. Sag (1994b) "Reflections and Revisions." Chapter 9 of Pollard and Sag 1994a. **P 158.4 P65**

Sag, Ivan A. (1997) "English Relative Clause Constructions." *Journal of Linguistics* 33: 431–484.

Sag, Ivan A., and Thomas Wasow (1999) *Syntactic Theory: A Formal Introduction*. Stanford, Calif.: CSLI Publications.