

Accounting for the Differences in the Syntactic Flexibility of Idioms

Sascha Bargmann

in collaboration with Gert Webelhuth, Manfred Sailer, and Christopher Goetze

University of Frankfurt and Goettingen

Laboratoire de Linguistique Formelle – CNRS and University of Paris-Diderot

18 March 2014

A Definition:

“An idiom is an expression larger than a word whose meaning cannot be systematically derived from meanings that the parts have when used independently of each other.”

(Huddleston and Pullum 2002, p. 273)

Prototypical Properties of Idioms:

- **phrasal:** multi-word expression
- **idiomatic:** non-literal and holistic meaning
- **lexically fixed:** none of the words can be replaced
- **syntactically fixed:** the idiom parts cannot be separated

Example of a Prototypical Idiom: *kick the bucket*

- **phrasal:** multi-word expression ✓
- **idiomatic:** non-literal and holistic meaning ✓
- **lexically fixed:** none of the words can be replaced ✓
 - (1) a. * Tom **kicked** the **container**.
 - b. * Tom **threw** the **bucket**.
 - c. * Tom **kicked** a **bucket**.
- **syntactically fixed:** the idiom parts cannot be separated ✓
 - (2) a. * **The bucket** has been **kicked**. (passive)
 - b. * **The bucket** appeared to have been **kicked**. (raising)
 - c. * **The bucket**, Tom has **kicked**. (topicalization)
 - d. * It was **the bucket** that Tom **kicked**. (cleft)
 - e. * **The bucket** Tom **kicked** was astonishing. (bare/contact relative)
 - f. Tom's **kicking the bucket** caused great concern. (verbal gerund)
 - g. * Tom's **kicking** of **the bucket** caused great concern. (nominal gerund)
 - h. * **Which** bucket did Tom **kick**? (*wh*-interrogative)

However, not all idioms are syntactically fixed:¹

- (3) a. Tom spilled the beans.
b. The beans were spilled by Tom.
c. * The beans, Tom has already spilled.
d. * These are the beans that Tom spilled.
- (4) a. Tom pulled some strings to get the job.
b. Strings were pulled every time Tom applied for a promotion.
c. Some strings, Tom has already pulled.
d. The strings that Tom pulled helped Chris get the job.
e. Tom pulled some/a lot of important strings to get you the job.

Wasow et al. (1983) and Nunberg et al. (1994):

There is a strong connection between the syntactic flexibility and the semantic decomposability of idioms.

¹ The unavailability of an idiomatic reading for (3c) and (3d) seems questionable, though.

Our Analysis

We distinguish **3 types of idioms**:

1. **Syntactically Frozen** Idioms (= IPhs) like *kick the bucket*

2. **Mobile** Idioms (= ICEs)
 - (a) **Semantically and syntactically connected** idioms like *spill the beans*
 - (b) **Semantically connected** idioms like *pull strings*

Our General Analytic Strategy

1. A **syntactically frozen** (i.e. *non*-decomposable) idiom is listed as **one single lexical entry** (i.e. as a completely fixed tree) in the **phrasal** lexicon.
2. A **mobile** (i.e. decomposable) idiom is composed of **separate lexical entries** that are listed in the **word or phrasal** lexicon and syntactically combined in the normal way.

The pieces of mobile idioms must be connected in the larger context containing them.

Our Strategy: capture the different degrees of mobility of idiom chunks by imposing different **syntactic and/or semantic connectedness conditions** on them.

Our SEMANTIC REPRESENTATION Account of Mobile Idioms

Central Assumptions concerning Syntax, Semantics, Discourse, and Idioms

Syntax:

HPSG, a constraint-based theory: no transformations, but only conditions of identity or sharing of substructure by different attributes of a common structure.

- **Passivization:**

taken care of in the lexicon by a derivational rule (to be shown and explained later).

- **Topicalization** is done via the interplay of:

1. **the set-valued nonlocal feature SLASH,**

which takes local structures as its values, which contain the local syntactic and semantic information of an expression.

2. a **NONLOCAL FEATURE PRINCIPLE,**

which, in a simplified version, states that the value of each nonlocal feature on a phrasal sign is the union of the values on the daughters.

3. a **COMPLEMENT EXTRACTION LEXICAL RULE**

(to be explained later).

Semantics:

- **Representation language:**
predicate logic with generalized quantifiers and lambda calculus
- **Combinatorics:**
as in Sailer (2003): a version of Flexible Montague Grammar
with lexical type shifting and functional application at phrasal nodes
- **Semantic contribution of a sign (word, phrase, sentence, ...):**
an expression of the semantic representation language
- **SEMANTIC REPRESENTATION (SR):**
part of the overall representation of the sign, i.e. there can be mutual constraints
on the syntactic form and SEMANTIC REPRESENTATION of a sign.

Discourse:

We assume a DRT-like architecture in which ...

- a semantic representation of the preceding discourse is available.
- the SEMANTIC REPRESENTATION of the current sentence is still set apart from that of the preceding discourse.
- anaphoric relations have already been resolved.

Idioms:

- Each part of a mobile idiom makes a unique contribution to the SEMANTIC REPRESENTATION of the larger linguistic context.
- A part of an idiom may require the unique SEMANTIC REPRESENTATION contribution of the other part(s) of the idiom to be present in the SEMANTIC REPRESENTATION of the larger linguistic context containing it.

The SEMANTIC REPRESENTATION Account of Mobile Idioms

Two Case Studies:

Case Study 1: *spill the beans*

Case Study 2: *pull strings*

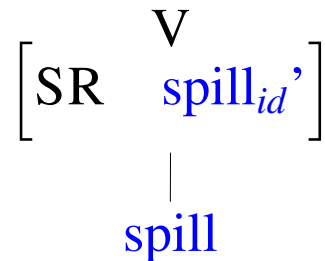
Case Study 1: *spill the beans*

Descriptive generalizations covering the empirical claims in the literature:

1. An NP of exactly the form *the beans* must be present.
2. The NP *the beans* can undergo A-movement but not A'-movement.
3. The NP *the beans* can be pronominalized.
4. Ellipsis of the verb *spill* is possible.

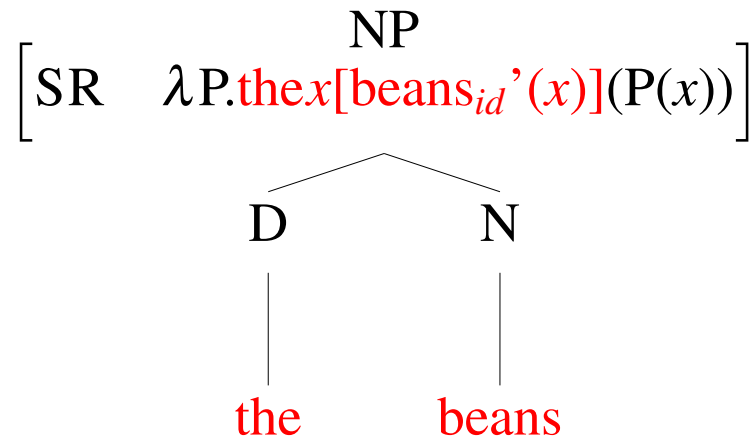
spill the beans – a Semantically and Syntactically Connected Idiom

Lexical Entry



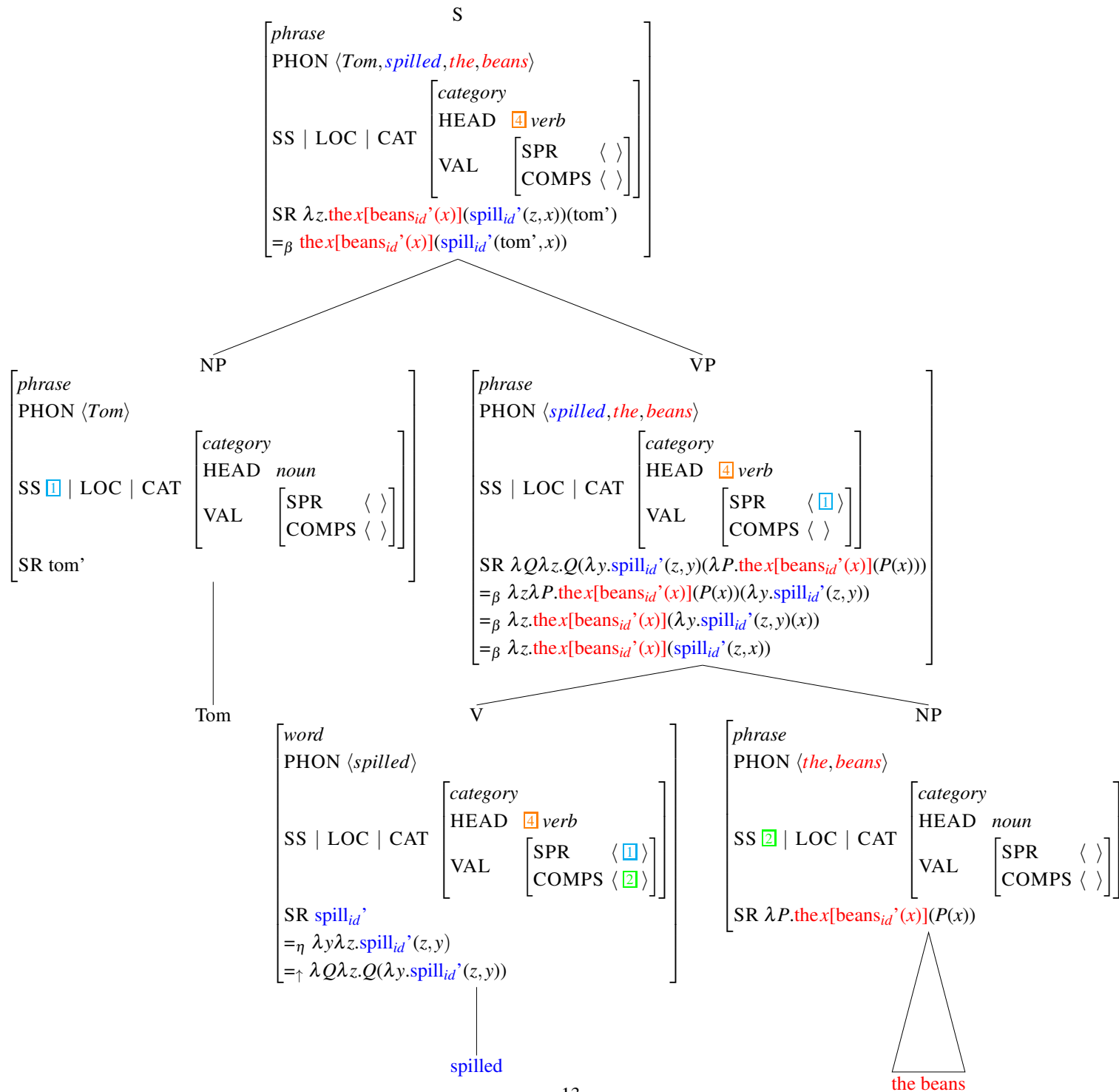
Constraints

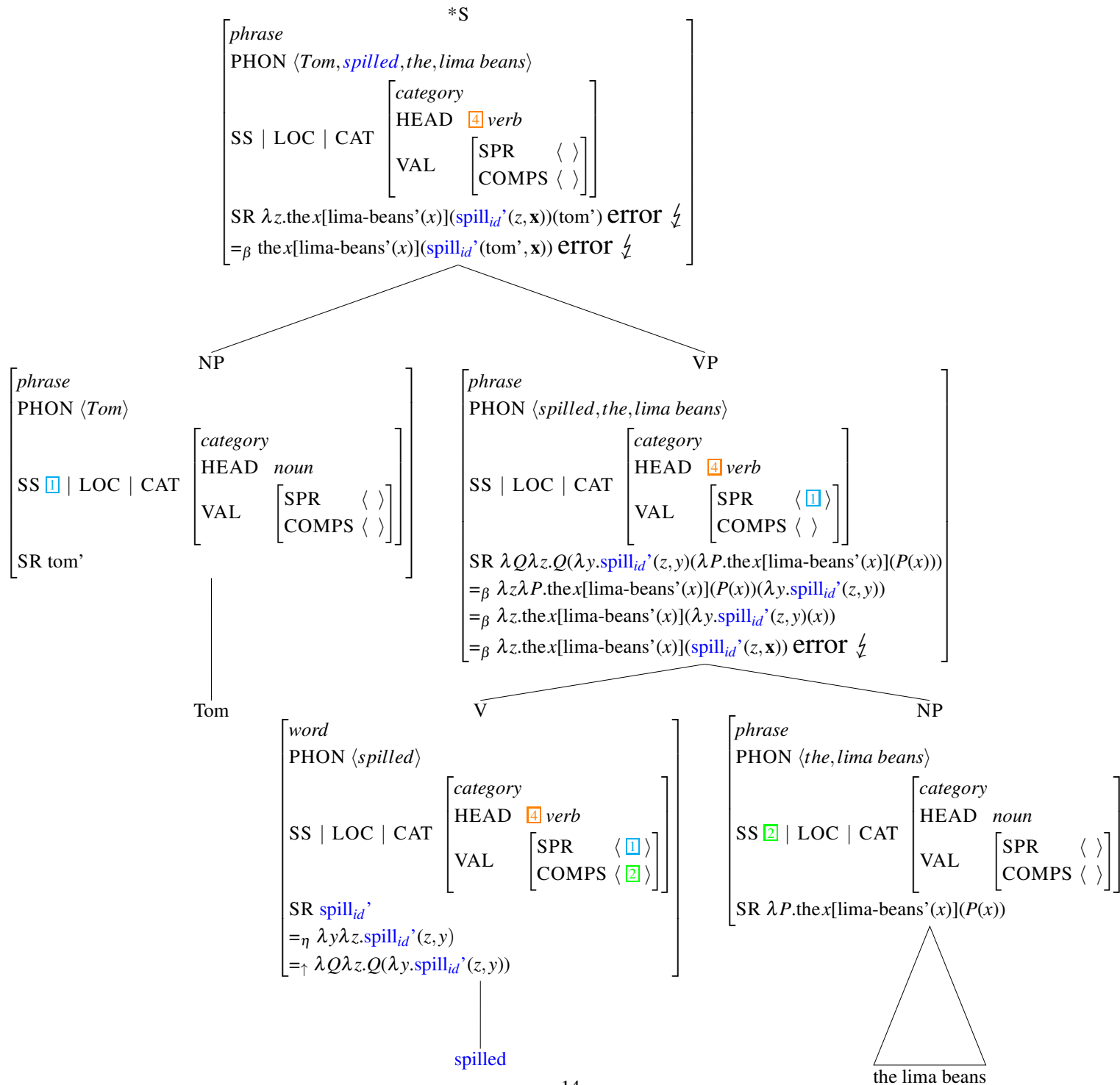
In the SEMANTIC REPRESENTATION, this verb's second argument must be specified by a term of the form $\text{thex}[\text{beans}_{id'}(x)]$ – possibly after anaphor resolution.



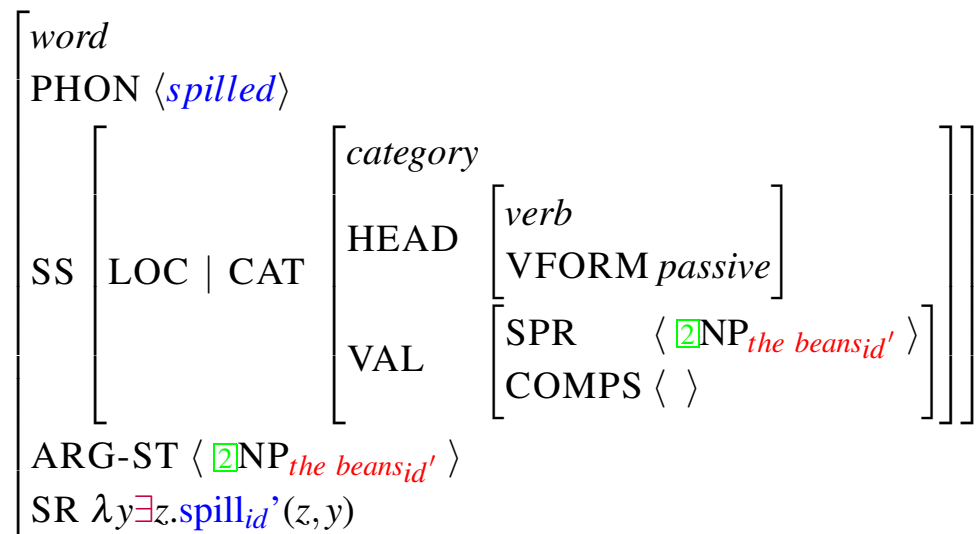
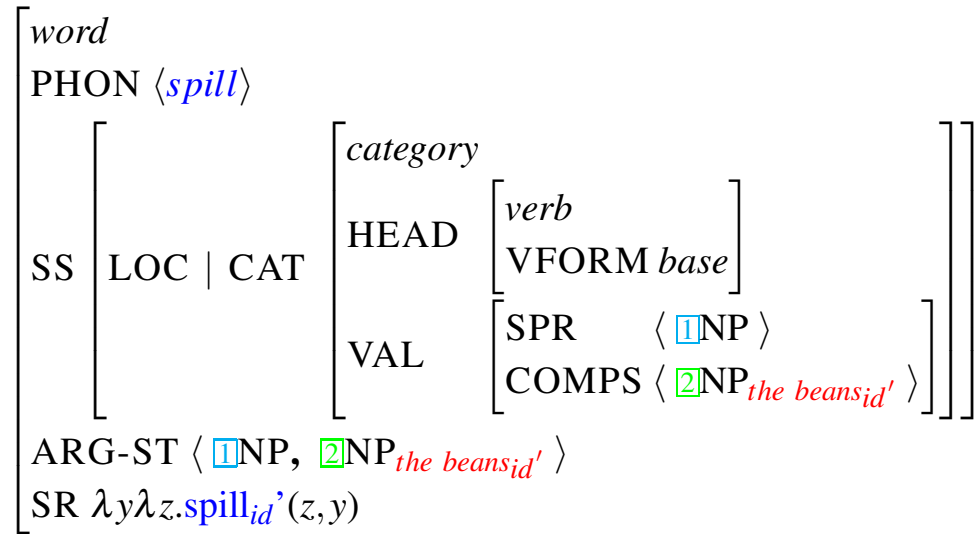
The SYNSEM value of this NP must occur on one of the valence lists of a word with the SEMANTIC REPRESENTATION $\text{spill}_{id'}$

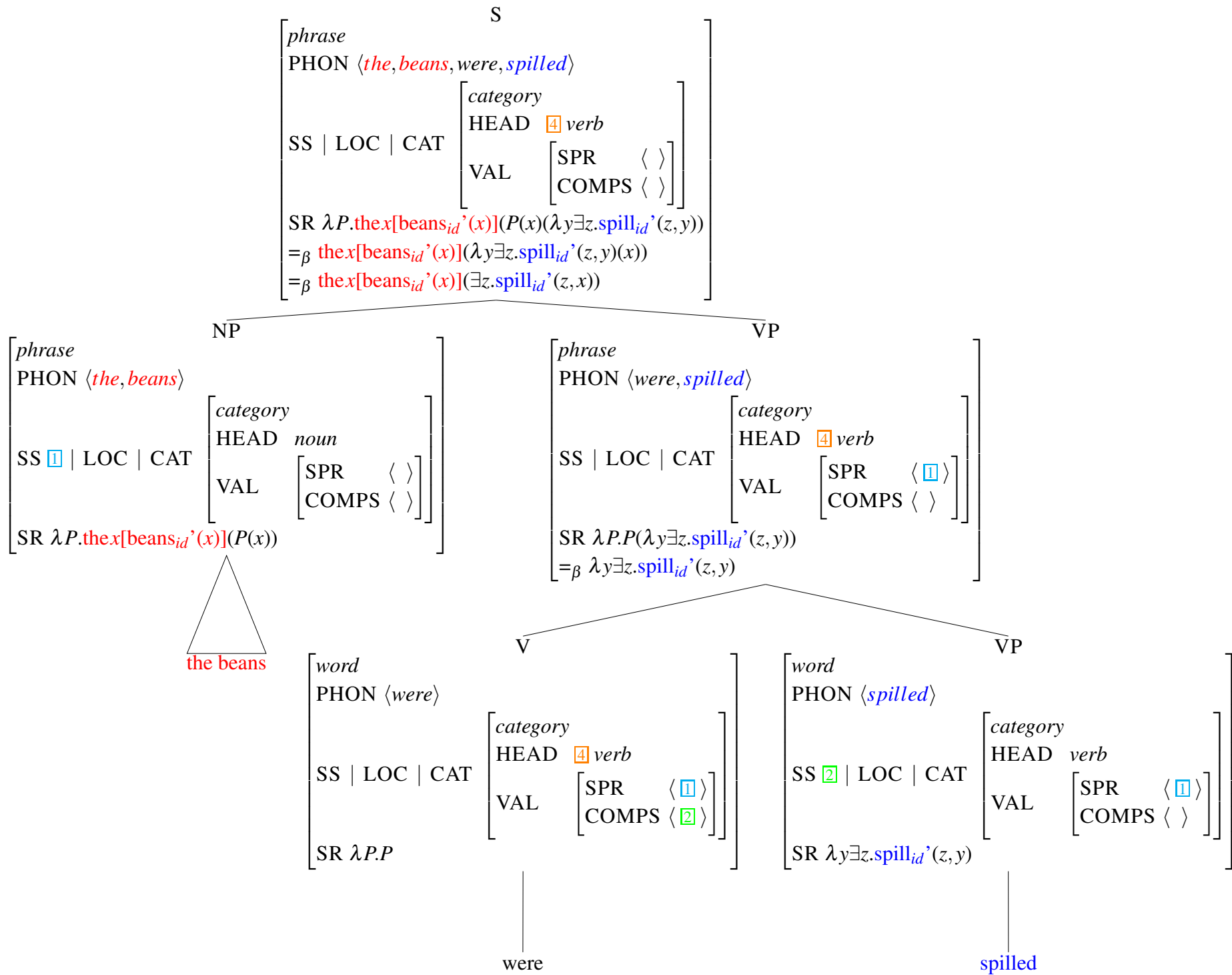
Crucial Assumption: The SEMANTIC REPRESENTATIONS in the lexical entries are contributed solely by the lexical entries themselves or pronouns/ellipsis sites licensed by them.



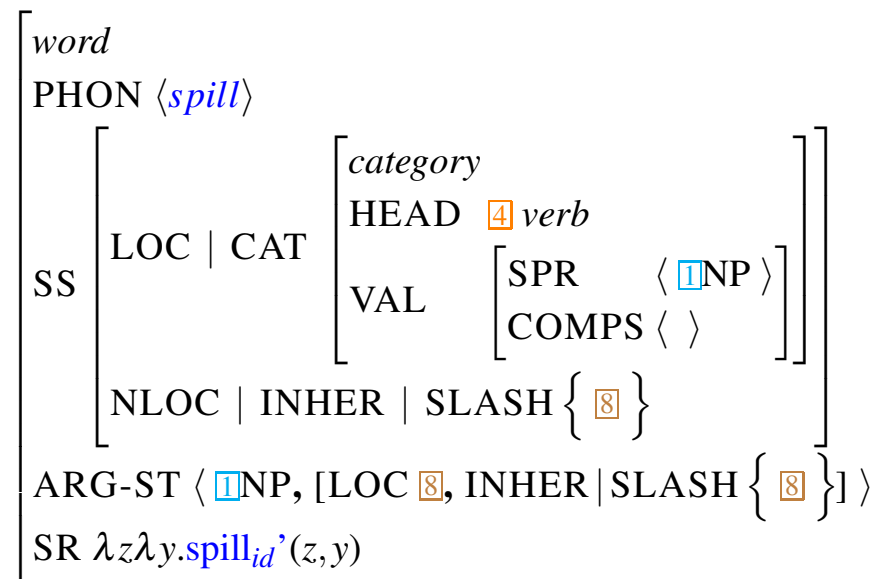
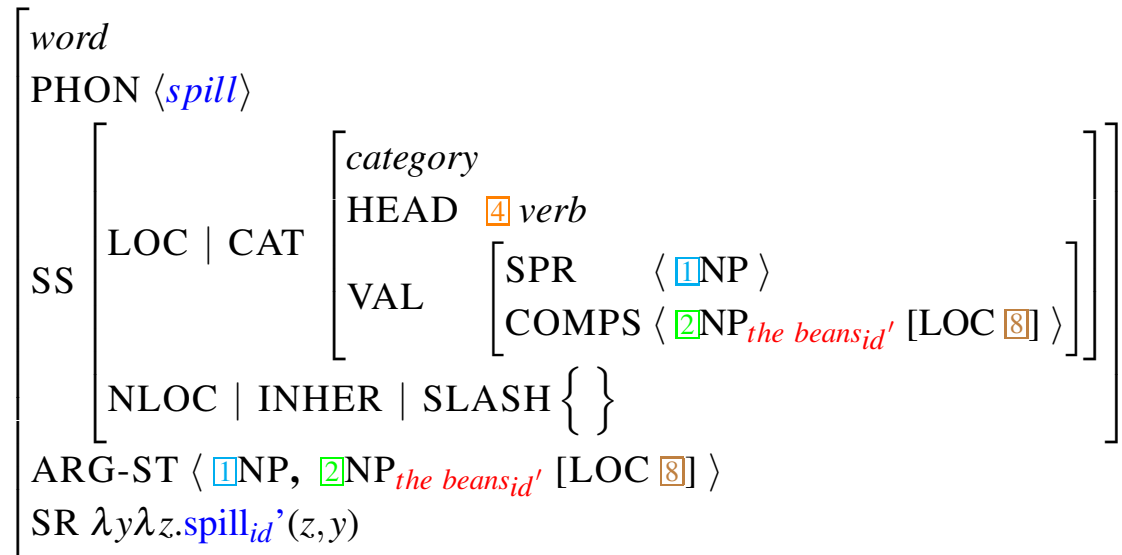


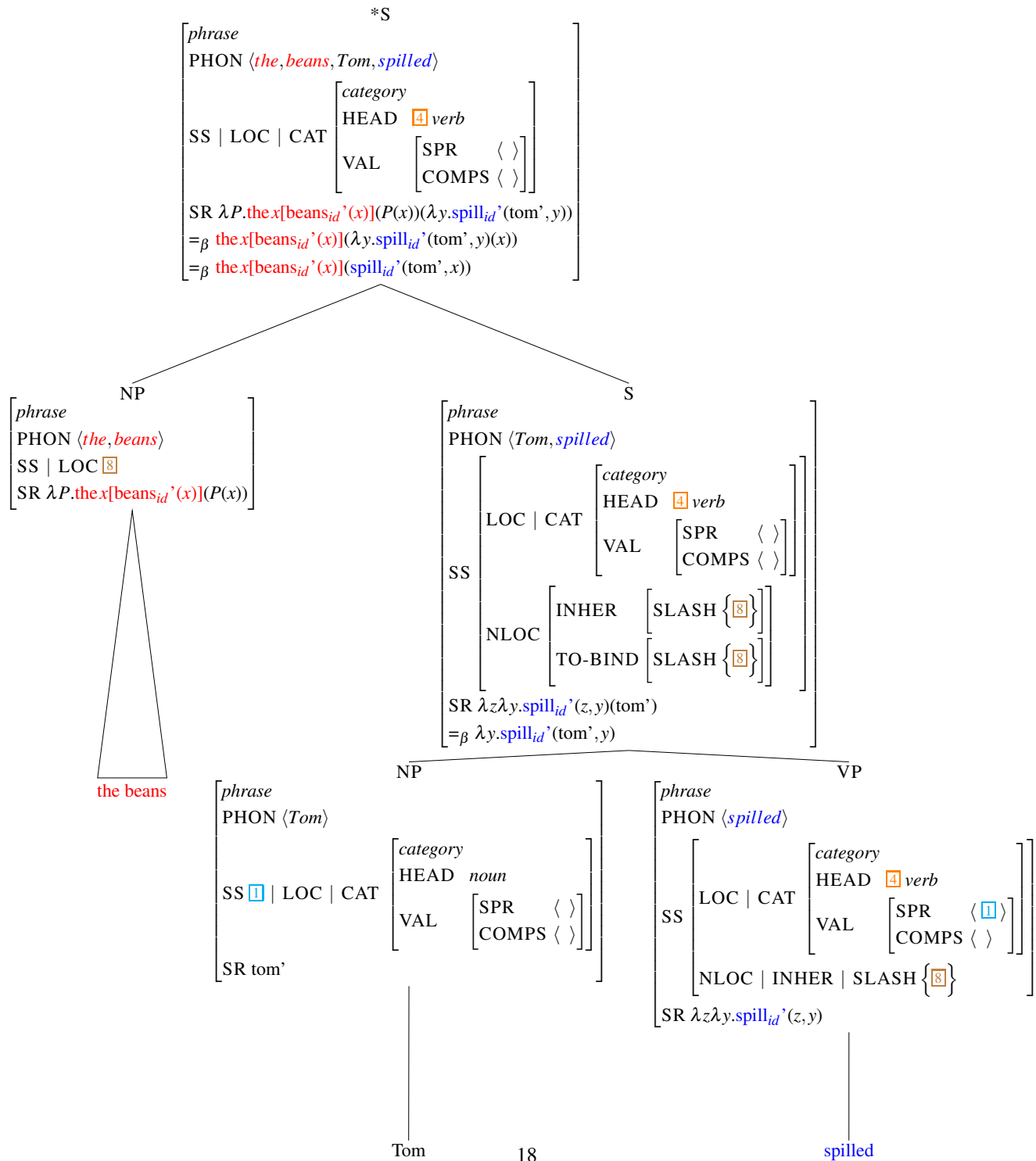
spill undergoing the PASSIVE LEXICAL RULE:





spill undergoing the **COMPLEMENT EXTRACTION LEXICAL RULE**:





Assumptions on Pronouns

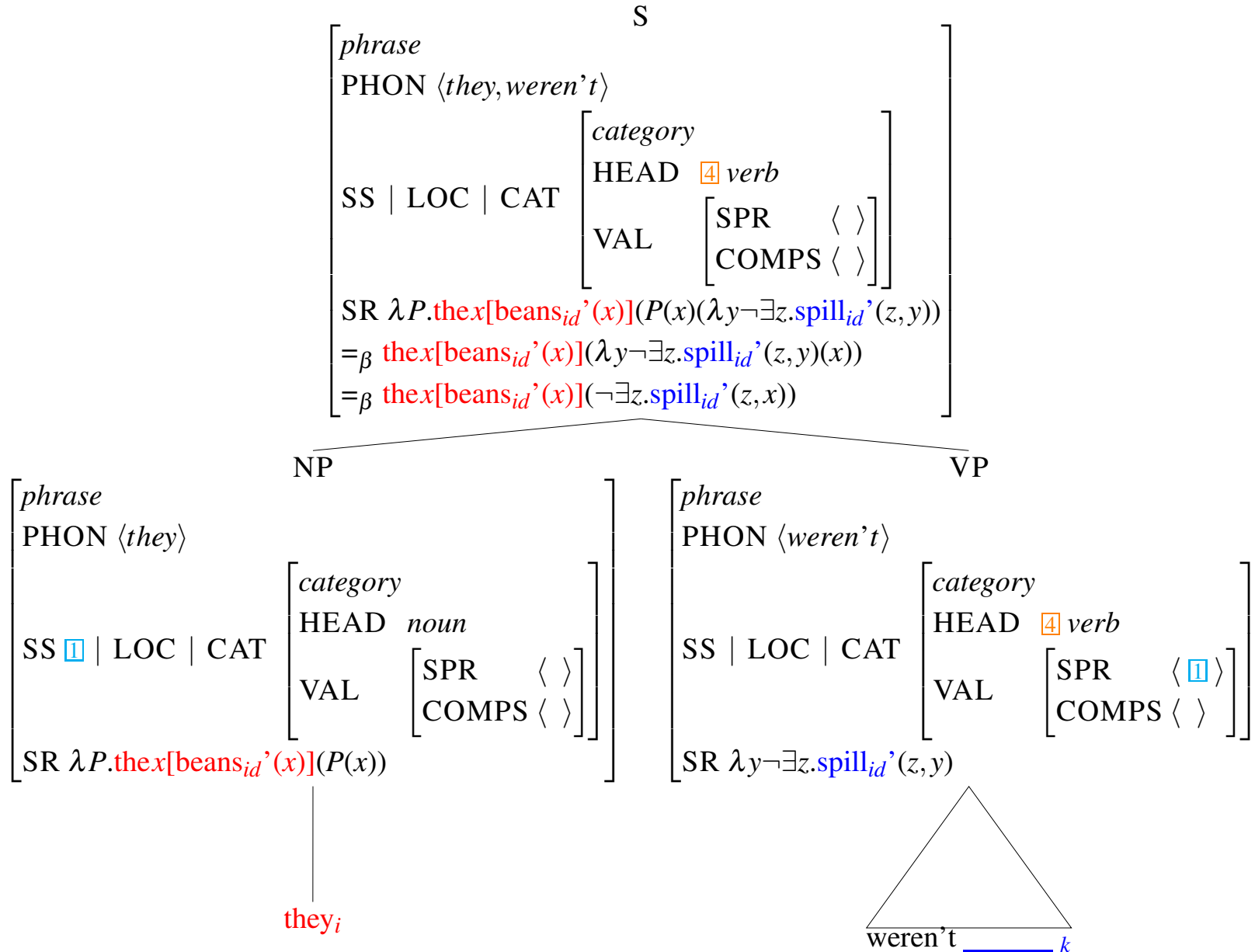
A personal pronoun is interpreted as a definite NP whose restrictor is identical to the restrictor of its antecedent:

- (5) a. Preceding discourse: **A woman_i** entered the room.
Current sentence: **She_i** whistled.
- b. Preceding discourse: $\exists x[\text{woman}'(x)](\text{enter-room}'(x))$
Current sentence: **the_x[woman}'(x)](whistle}'(x))**

I was worried that *the beans_i* might be *spilled_k*,

SR: $\text{the } x[\text{beans}_{id'}(x)](\text{worried}'(\text{speaker}', \text{might}'(\exists z. \text{spill}_{id'}(z, x))))$

but



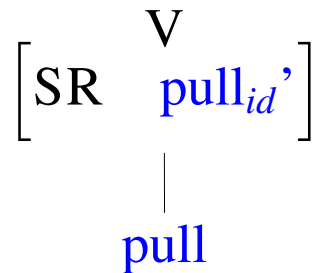
Case Study 2: *pull strings*

pull strings is more flexible than *spill the beans* in several respects:

1. The determiner of the NP headed by *strings* is not frozen.
2. The NP headed by *strings* can undergo both A- and A'-movement.
3. *strings* can occur in a main clause without *pull*, if another occurrence of *strings* did cooccur with *pull* in the preceding discourse.

pull strings – a Semantically Connected Idiom

Lexical Entry



Constraint

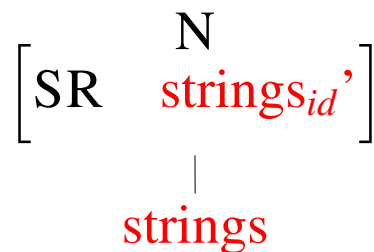
In the SEMANTIC REPRESENTATION, the verb's second argument is specified by a quantifier Qx that is restricted by $\text{strings}_{id}'(x)$ – possibly after anaphor resolution.

$\text{strings}_{id}'(x)$ restricts a quantifier Qx and either

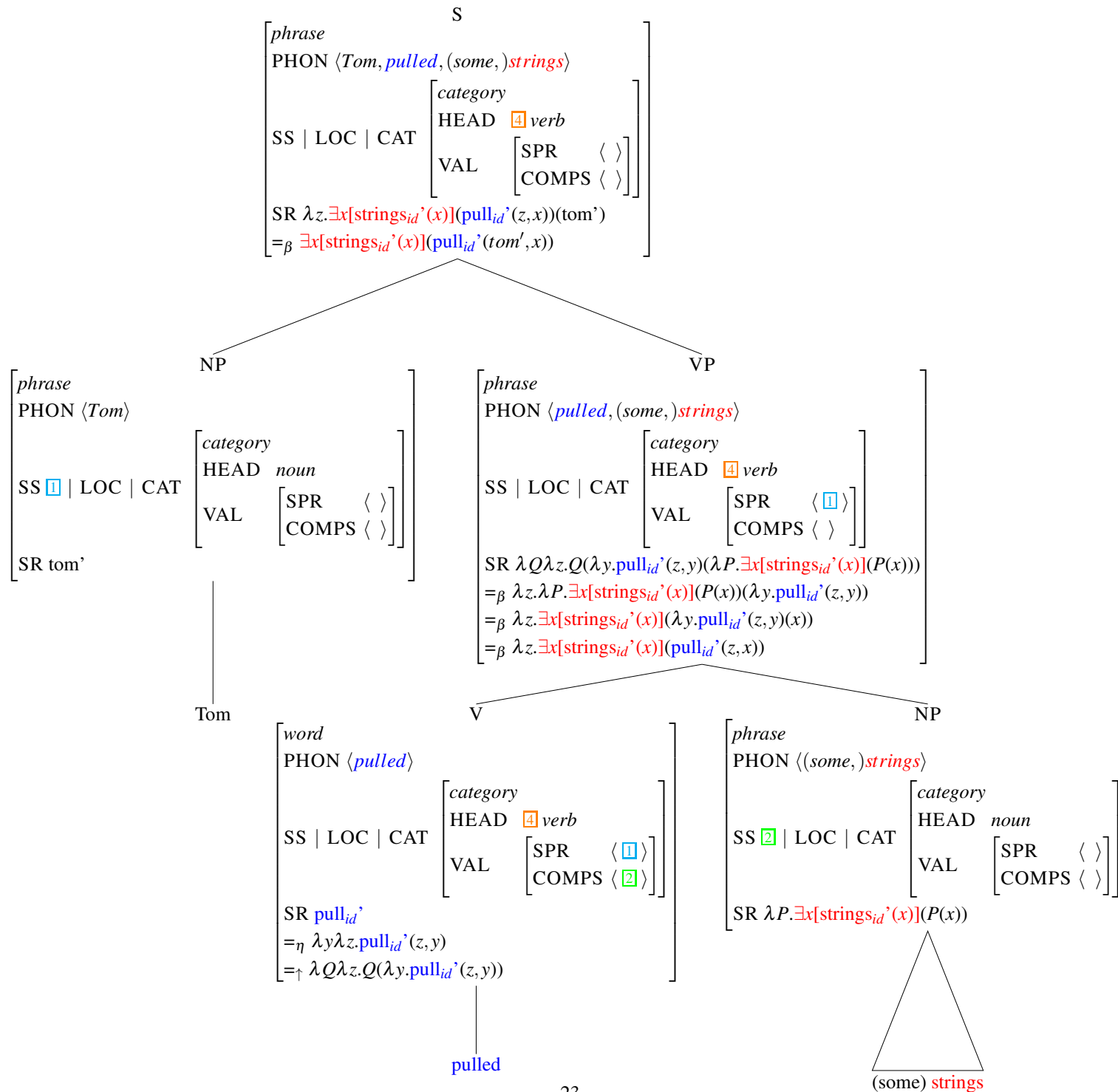
a Qx binds the second argument of pull_{id}'
(possibly after anaphor resolution)

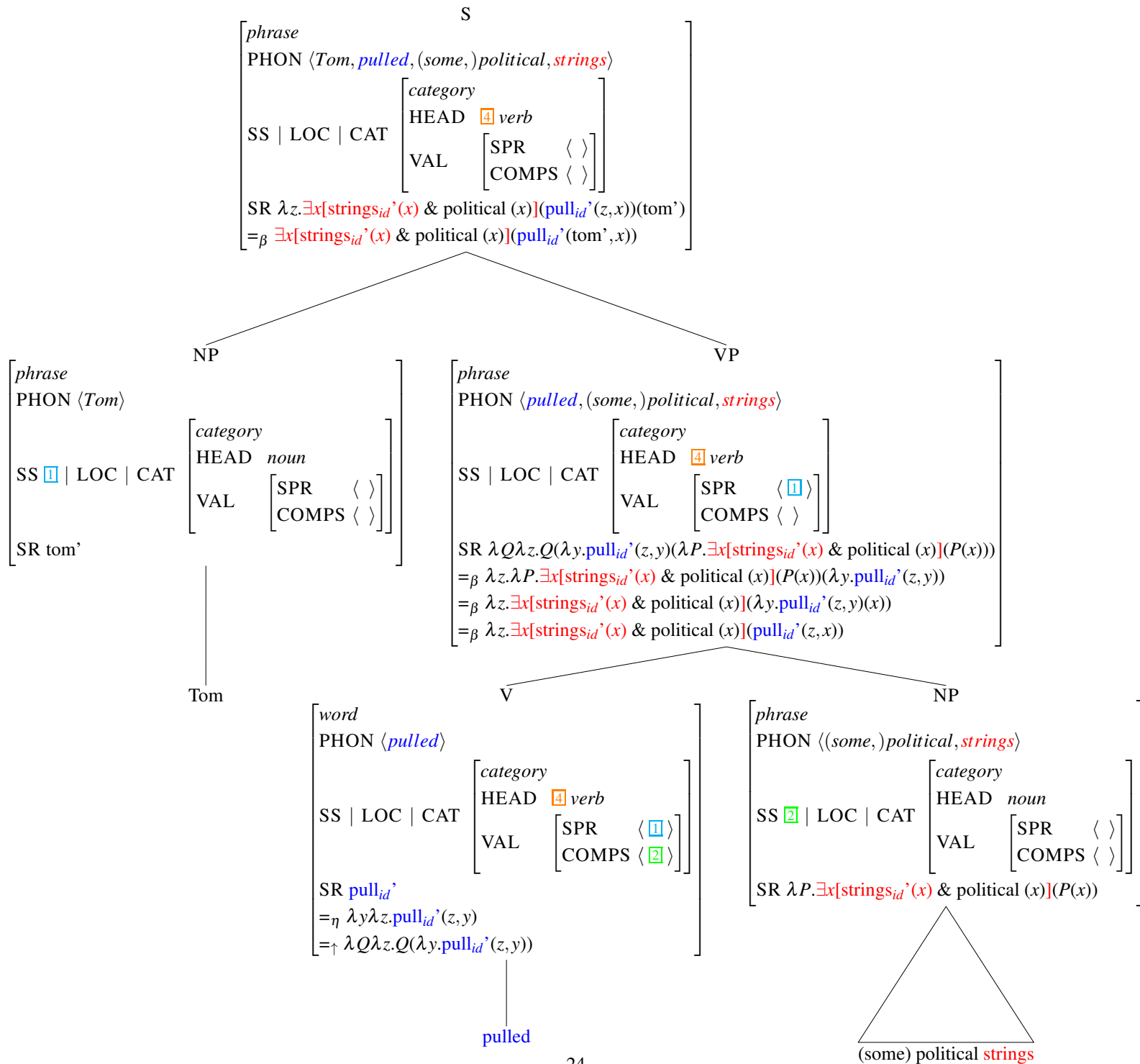
or

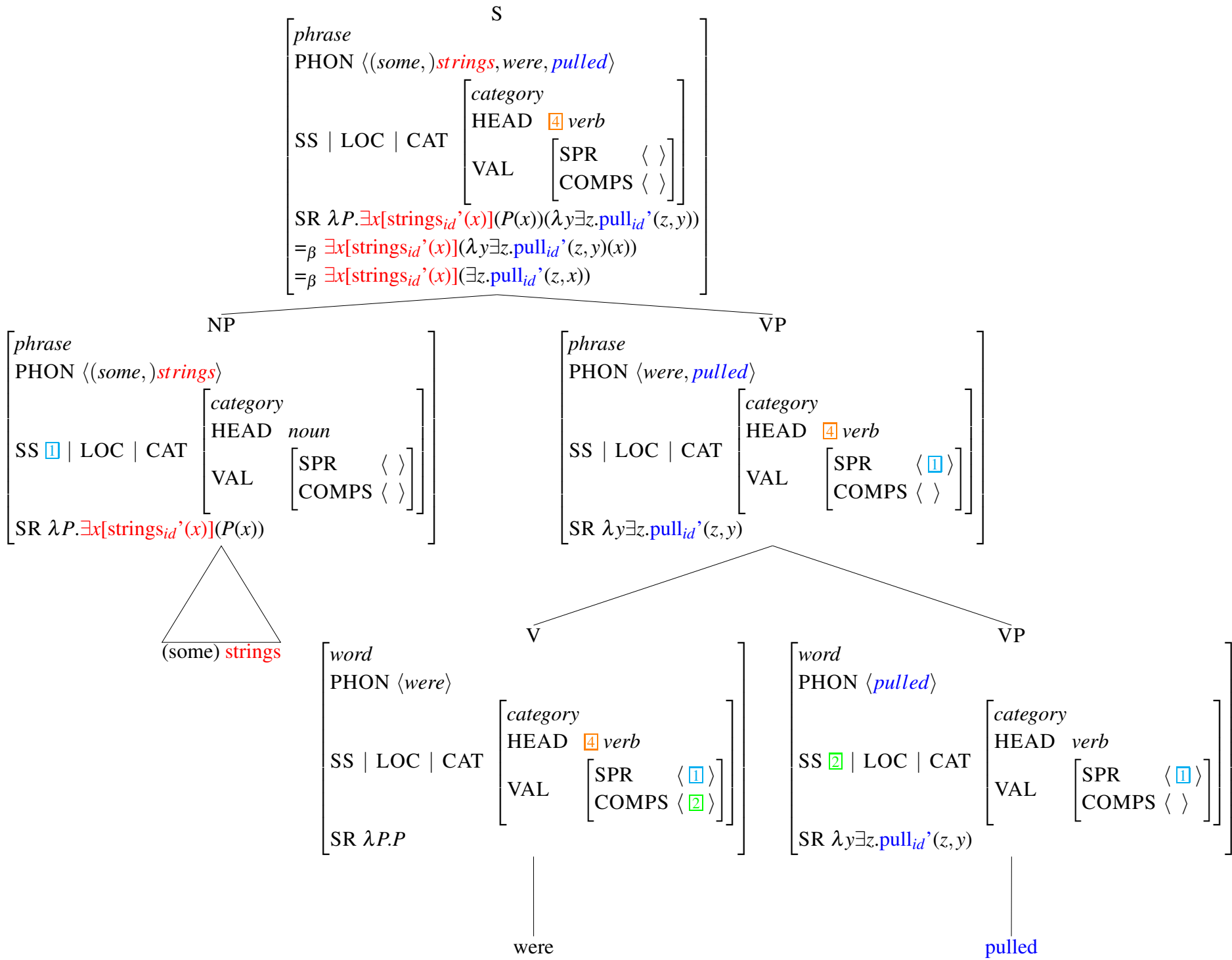
b $\text{strings}_{id}'$ is salient in the present discourse.

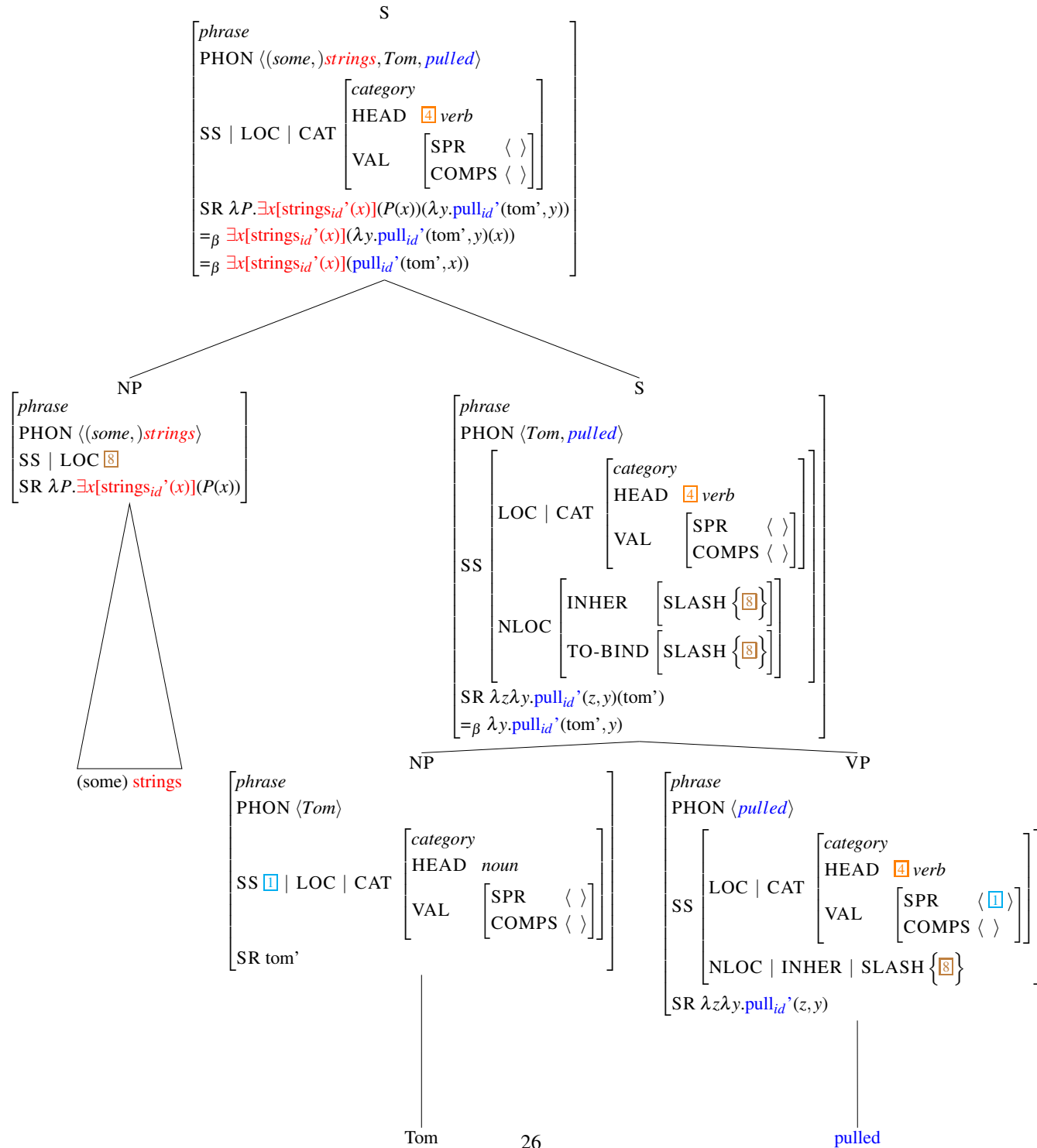


Crucial Assumption: The SEMANTIC REPRESENTATIONS in the lexical entries are contributed solely by the lexical entries themselves or pronouns/ellipsis sites licensed by them.





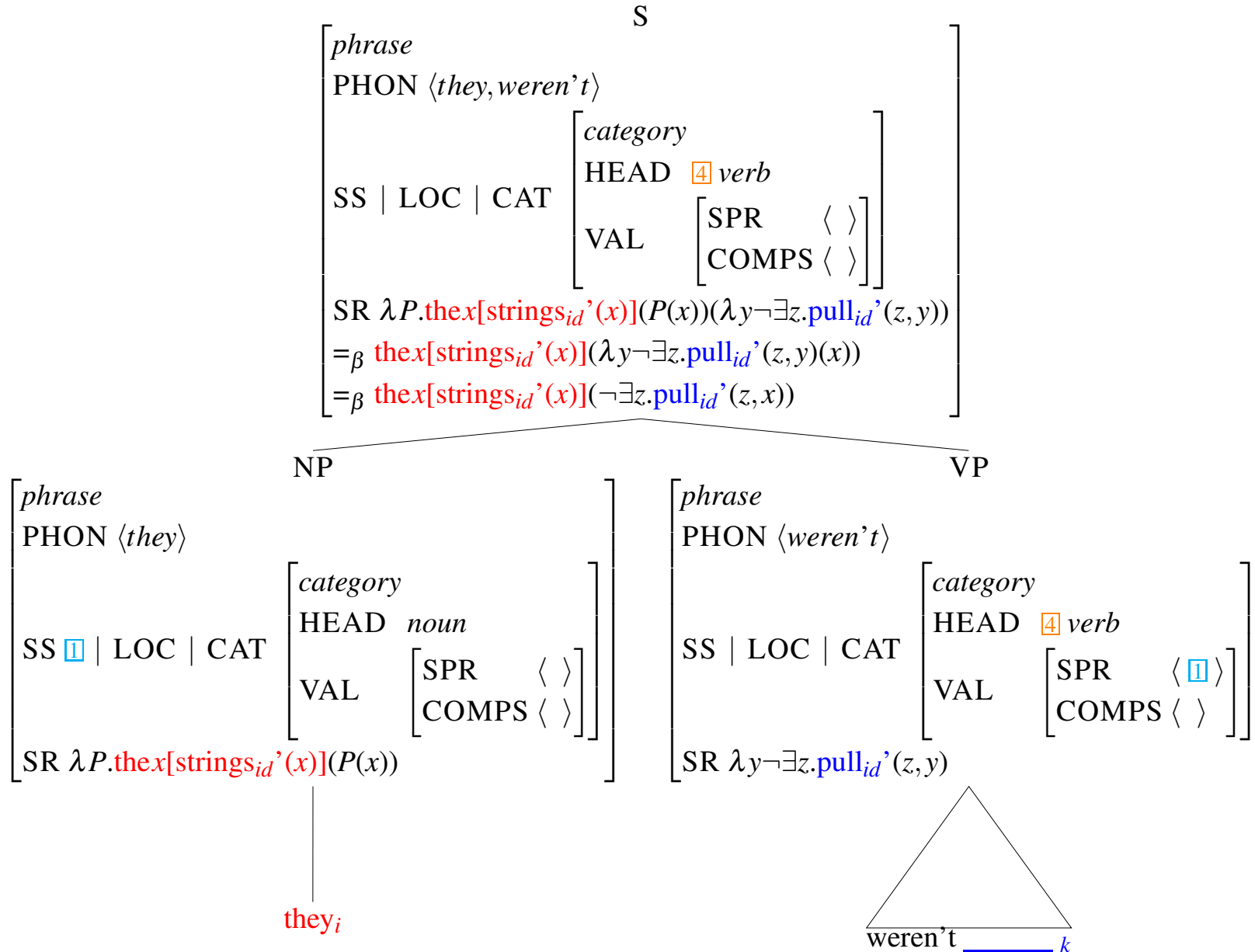




I was worried that *strings_i* might be *pulled_k*,

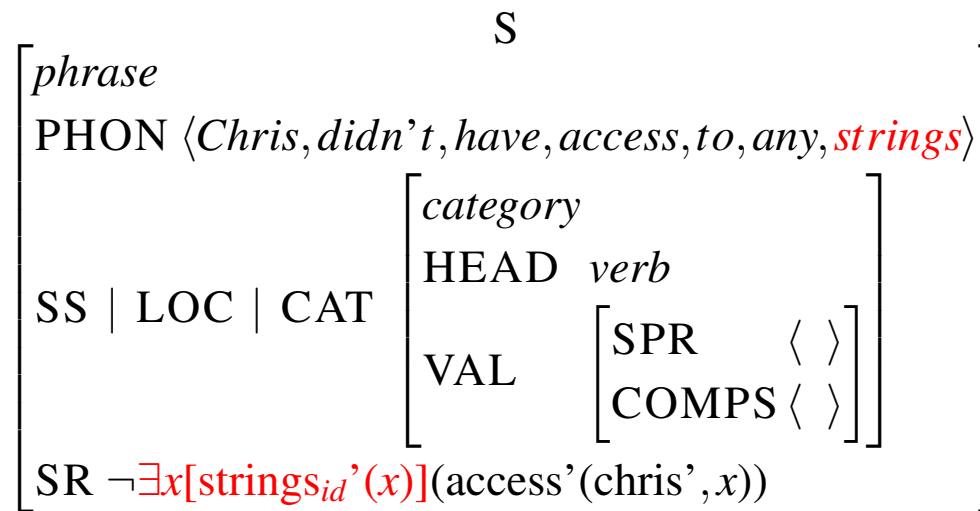
SR: worried'(speaker', might'($\exists x[\text{strings}_{id}'(x)](\exists z.\text{pull}_{id}'(z,x))$))

but



Tom and Chris graduated from law school together with roughly equal records. Tom's uncle is a state senator, and he *pulled strings* to get Tom a clerkship with a state supreme court justice.

Context LF: ... *strings_{id}'* ...



Chris didn't have access to any *strings*

Summary

1. Idioms differ from each other in their degree of syntactic flexibility.
2. Syntactically frozen idioms are treated as surface entries in the phrasal lexicon.
- 3. The parts of a mobile idiom each have their own lexical entry, which makes reference to some syntactic and/or semantic property of the other part(s) of the idiom.**
- 4. Mobile idioms differ from each other in how their parts are linguistically connected.**
- 5. We predict a hierarchy of idioms in terms of the syntactic mobility of their parts: phrasal lexical entry < syntactically connected < semantically connected.**

References

- Rodney Huddleston and Geoffrey K. Pullum. *The Cambridge Grammar of the English Language*. Cambridge University Press, Cambridge, 2002.
- Geoffrey Nunberg, Ivan A. Sag, and Thomas Wasow. Idioms. *Language*, 70:491–538, 1994.
- Manfred Sailer. Combinatorial Semantics and Idiomatic Expressions in Head-Driven Phrase Structure Grammar. Phil. Dissertation (2000). Arbeitspapiere des SFB 340. 161, Universität Tübingen, 2003.
- Thomas Wasow, Ivan A. Sag, and Geoffrey Nunberg. Idioms: an interim report. In S. Hattori and K. Inoue, editors, *Proceedings of the XIIIth International Congress of Linguistics*, pages 102–115, 1983.