

TAU-GU

Day 2

General Introduction to the First Three Talks of Today

There are at least three ways of analyzing relative clauses:

- Raising
- Matching
- Modification

Disadvantage of Raising and Matching:

Total loss of a number of widely agreed-upon empirical generalizations in syntax and morphology

Disadvantage of Modification:

No solution to handling reconstruction effects (yet)

Our Goal:

Provide a solution to handling reconstruction effects under Modification without literal syntactic reconstruction.

There are at least two types of reconstruction that we have to address:

- Binding-theoretic reconstruction (Christopher Götze)
- Idiom reconstruction (Sascha Bargmann)

Idiomatic expressions ...

- represent a key factor in the discussion on the proper analysis of relative clauses.
- are much more pervasive in natural language than has traditionally been assumed.

We therefore need an explicit account of how they actually work.

First Talk:

- **English idiom data**
- **Söhn's Context Of Lexical Licensing (COLL) account** (Söhn 2006)

Second Talk:

- **The Semantic Representation (SR) theory**
- **Comparison of the two accounts** in their capability to capture the English idiom data

About Bear Services and Flabbergasted Building Blocks
– Applying Söhn’s Idiom Theory to English Data

Sascha Bargmann
University of Frankfurt

TAU-GU
29 November 2014

Outline

1. English Idiom Data

- (a) A Definition and Prototypical Properties of Idioms
- (b) The Syntactic (Non-)Flexibility of English VP Idioms

2. Söhn's Context of Lexical Licensing (COLL) Account

- (a) Theoretical Assumptions
- (b) Application to the English Idiom Data

3. Summary and Evaluation

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 2002:273)

Prototypical Properties of Idioms:

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

***kick the bucket* (= die), an almost prototypical Idiom:**

- **phrasal:** multi-word expression ✓
- **idiomatic:** non-literal meaning ✓
- **lexically fixed:** none of the words can be replaced ✓
 - (1) a. Kim **kicked** **the bucket**.
 - b. * Kim **kicked** **the** pail.
 - c. * Kim booted **the bucket**.
 - d. * Kim **kicked** a **bucket**.
- **syntactically fixed:** the idiom parts cannot be separated (✓)
 - (2) a. * **The bucket** appeared to have been **kicked**. (raising + passive)
 - b. * **The bucket**, Kim has **kicked**. (topicalization)
 - c. * **The bucket** (that) Kim **kicked** was astonishing. (relative clause)
 - d. * What **bucket** did Kim **kick**? (*wh*-interrogative)
 - e. * Kim feared **the bucket** would be **kicked**, and it was. (pronom. + V-ellipsis)
 - f. * Kim **kicked** **bucket** after **bucket**. (*N-after-N* Construction)
 - g. * Kim's **kicking** of **the bucket** caused great concern. (nominal gerund)
 - h. Kim's **kicking** **the bucket** caused great concern. (verbal gerund)
 - i. Kim **kicked** **the** proverbial/social **bucket**. (adjectival modification)

spill beans (= reveal secrets), a less prototypical idiom:

- **phrasal:** multi-word expression ✓
- **idiomatic:** non-literal meaning ✓
- **lexically fixed:** none of the words can be replaced ✓

- (3) a. Snowden **spilled** the **beans**.
b. * Snowden **spilled** the kernels/seeds/grain.
c. * Snowden **slopped/scattered/shed** the **beans**.

- **syntactically fixed:** the idiom parts cannot be separated – **not the case!**

- (4) a. Most (of the) **beans** seem to have been **spilled** already. (raising + passive)
b. Most (of the) **beans**, Snowden has already **spilled**. (topicalization)
c. That's all (of) the **beans** that Snowden has **spilled** for now. (relative clause)
d. What **beans** has Snowden **spilled**, though? (*wh*-interrogative)
e. Snowdon hoped **beans** would be **spilled**, but they weren't. (pronom. + V-ellipsis)
f. Snowdon **spilled bean** after **bean**. (N-*after*-N Construction)
g. Snowden's **spilling** of (the) **beans** has been a major topic. (nominal gerund)
h. Snowden's **spilling** (the) **beans** has been a major topic. (verbal gerund)
i. Snowden **spilled** some top-secret **beans**. (det variation + adj modification)

pull strings (= use connections), another less prototypical idiom:

- **phrasal:** multi-word expression ✓
- **idiomatic:** non-literal meaning ✓
- **lexically fixed:** none of the words can be replaced ✓

- (5) a. Kim **pulled** some **strings** to get the job.
b. * Kim **pulled** some threads/filaments to get the job.
c. * Kim some **strings** to get the job.

- **syntactically fixed:** the idiom parts cannot be separated – **not the case!**

- (6) a. **Strings** seem to have been **pulled** to get Alex the job. (raising + passive)
b. Some **strings**, Kim has already **pulled** to get Alex the job. (topicalization)
c. The **strings** (that) Kim **pulled** helped Alex get the job. (relative clause)
d. What **strings** did Kim **pull**, though? (*wh*-interrogative)
e. Kim feared **strings** would be **pulled**, but they weren't. (pronom. + V-ellipsis)
f. Kim **pulled string** after **string** to get the job. (N-*after*-N Construction)
g. Kim's **pulling** of **strings** was the only way to get the job. (nominal gerund)
h. Kim's **pulling strings** was the only way to get the job. (verbal gerund)
i. Kim **pulled** a lot of important **strings** for you. (det variation + adj modification)

Overview over the syntactic flexibility of the three English VP idioms we looked at:

| | <i>kick the bucket</i> | <i>spill beans</i> | <i>pull strings</i> |
|-----------------------------------|------------------------|--------------------|---------------------|
| NP Movement | * | ✓ | ✓ |
| <i>Wh</i> Movement | * | ✓ | ✓ |
| Ellipsis of the Verb | * | ✓ | ✓ |
| Pronominalization of the NP | * | ✓ | ✓ |
| N- <i>after</i> -N Construction | * | ✓ | ✓ |
| Nominal Gerund | * | ✓ | ✓ |
| Determiner Variation in the NP | * | ✓ | ✓ |
| Adjectival Modification in the NP | ✓ | ✓ | ✓ |

Overview over the syntactic flexibility of the three English VP idioms we looked at:

| | <i>kick the bucket</i> | <i>spill beans</i> | <i>pull strings</i> |
|-----------------------------------|------------------------|--------------------|---------------------|
| NP Movement | * | ✓ | ✓ |
| <i>Wh</i> Movement | * | ✓ | ✓ |
| Ellipsis of the Verb | * | ✓ | ✓ |
| Pronominalization of the NP | * | ✓ | ✓ |
| N- <i>after</i> -N Construction | * | ✓ | ✓ |
| Nominal Gerund | * | ✓ | ✓ |
| Determiner Variation in the NP | * | ✓ | ✓ |
| Adjectival Modification in the NP | ✓ | ✓ | ✓ |

Conclusion: There are at least two types of English VP idioms:

1. Syntactically almost entirely frozen idioms like *kick the bucket*
2. Syntactically very flexible idioms like *spill beans* or *pull strings*

In this talk, I will focus on the second type and use *pull strings* as the example.

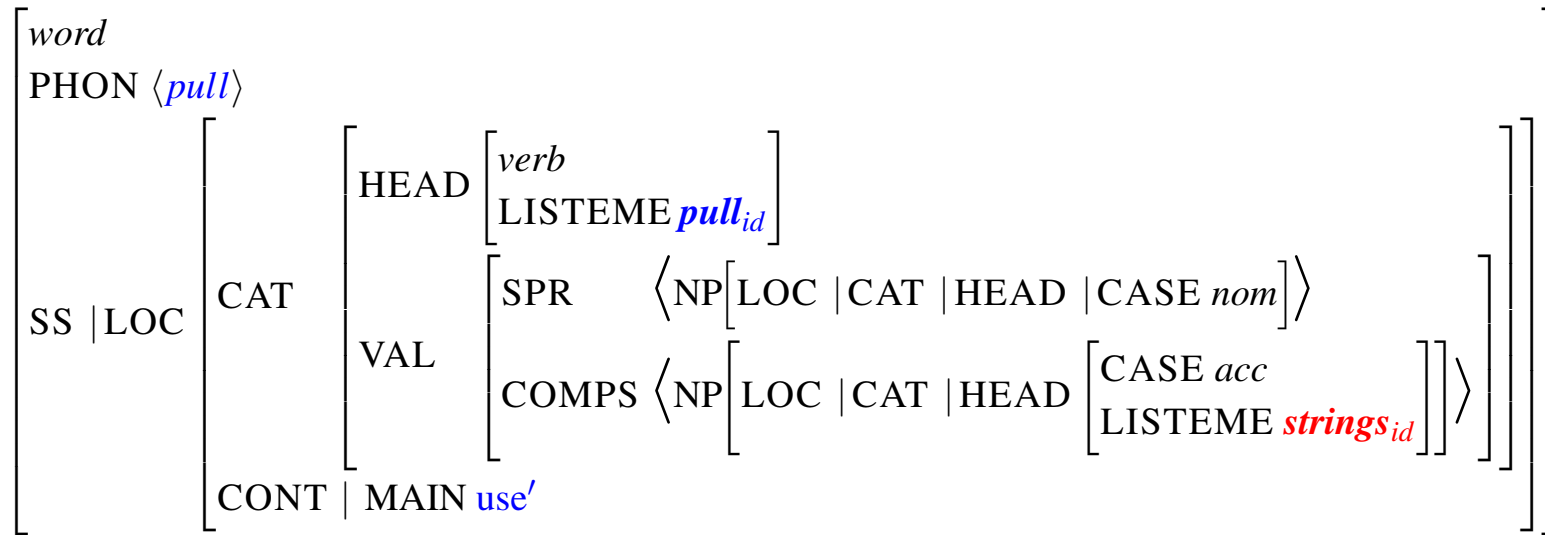
Söhn's Context Of Lexical Licensing (COLL) Account

Theoretical Assumptions on Syntax, Semantics, and Idioms:

- **Syntax:** Head-Driven Phrase Structure Grammar (HPSG)
- **Semantics:** Lexical Resource Semantics (LRS)
- **Idioms:**
 - An idiom like *pull strings* is composed of separate lexical entries (in this particular case *pull* and *strings*) that are syntactically combined in the normal way.
 - These separate lexical entries are subject to idiosyncratic constraints concerning their combinatorial potential: *pull* searches for *strings*, and *strings* searches for *pull*.
 - They do that via the value of the morphosyntactic feature LISTEME: Each lexical entry has a distinct LISTEME value that functions like a genetic code or fingerprint by which the lexical entry can be uniquely identified.
- **Data:** Söhn (2006) is concerned with German VP idioms that have no literal reading.
 - 👉 I will apply his account to a subset of the English idiom data we just looked at.

Söhn's Context Of Lexical Licensing (COLL) Account

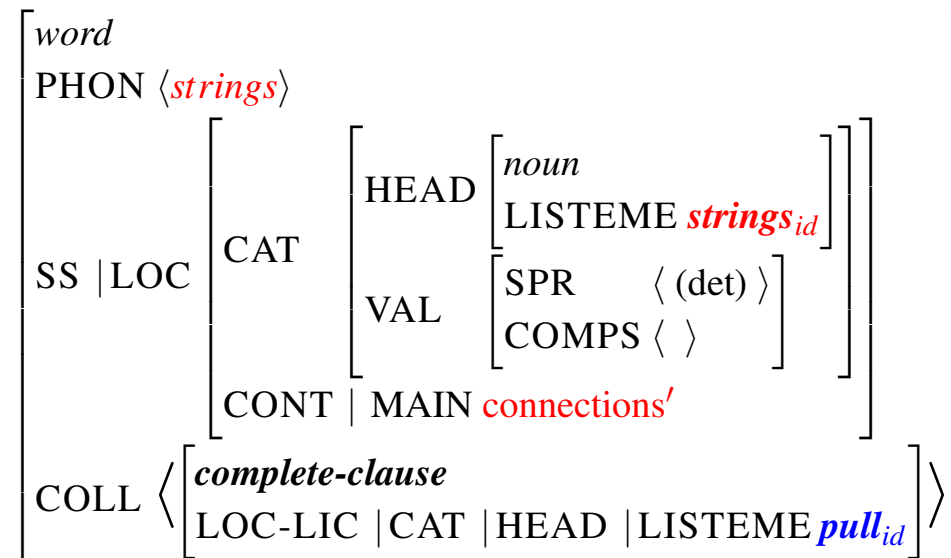
The lexical entry of idiomatic *pull* in Söhn's system:¹



In essence: The verb with the LISTEME value *pull_{id}* wants an NP with the LISTEME value *strings_{id}* as its internal idiomatic argument.

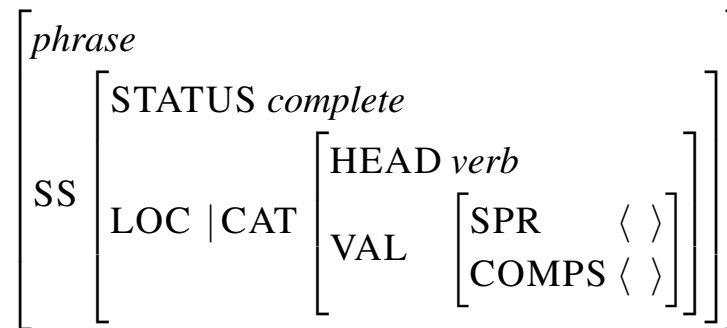
¹VAL=Valence, PHON=Phonology, SS=Syntax+Semantics, LOC=Local, CAT=Category, CONT=Content, SPR=Specifier, COMPS=Complements

The lexical entry of idiomatic *strings* in Söhn's system:²



In essence: The idiomatic noun *strings* with the unique LISTEME value *strings_{id}* wants to occur within a *complete-clause* that has the unique LISTEME value *pull_{id}*.

DEF of *complete-clause*:



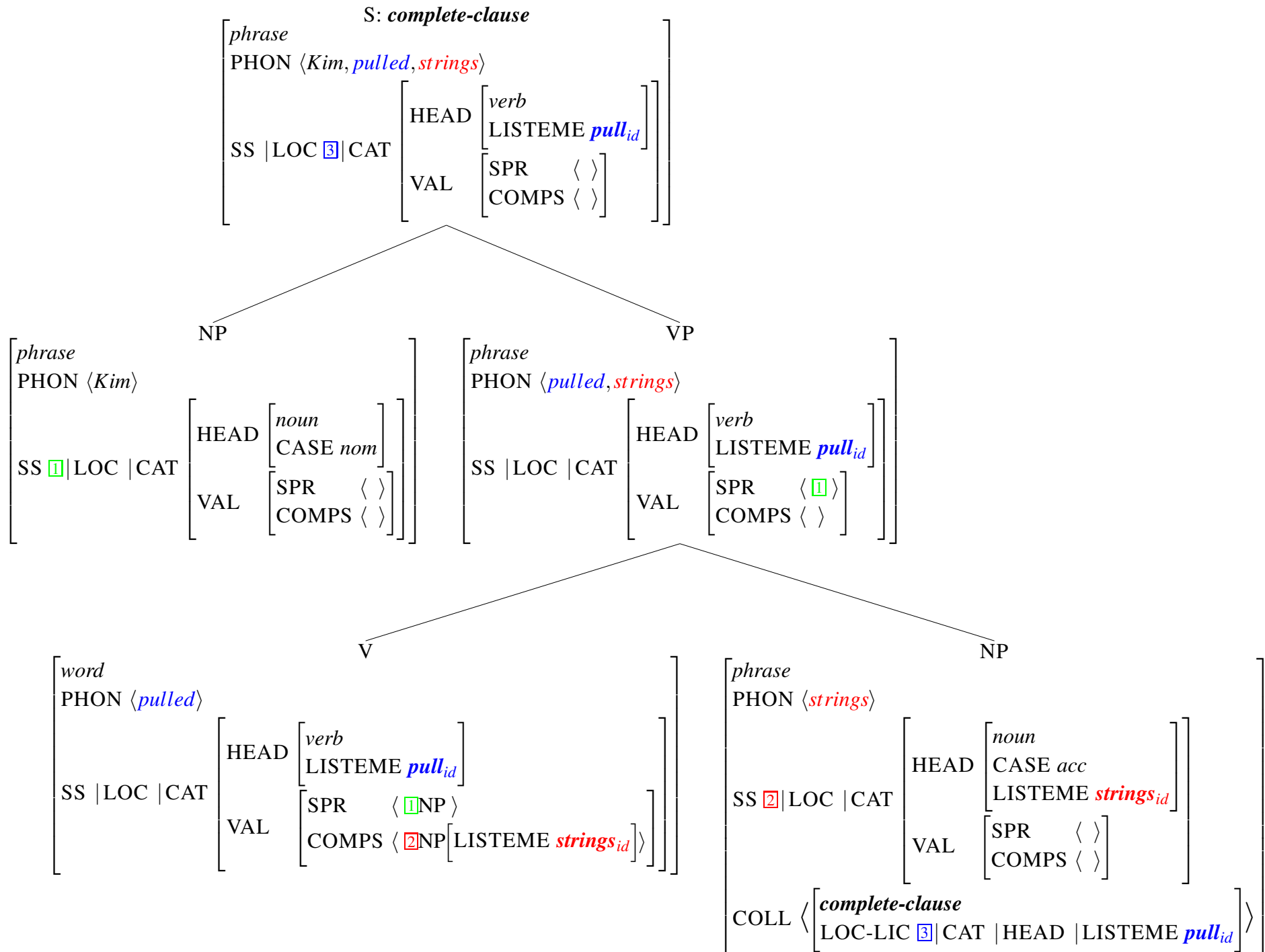
²COLL = Context of Lexical Licensing, LOC-LIC = Local Licenser

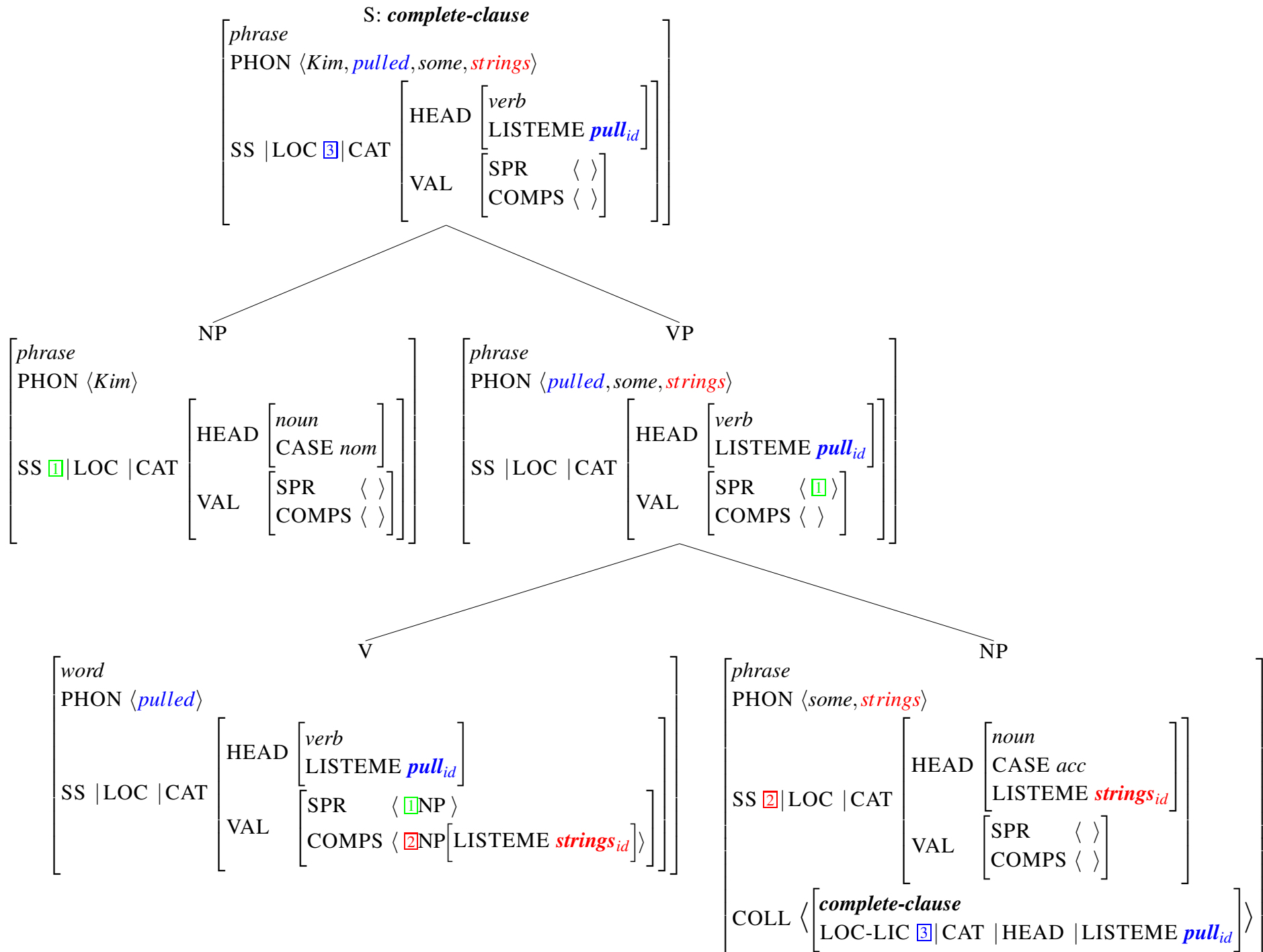
How *pull strings* functions in Söhn's system:

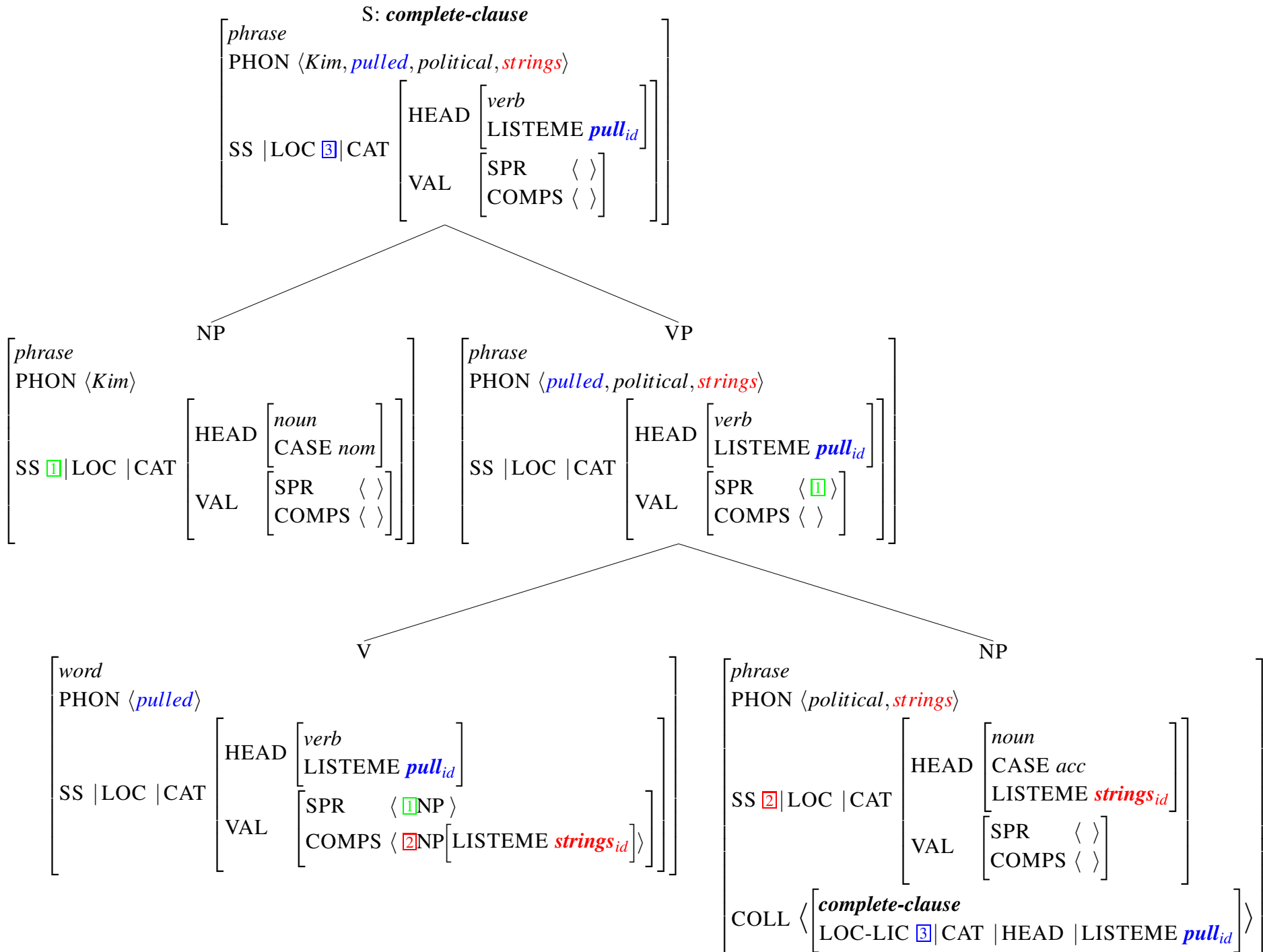
- The pieces of *pull strings* search for each other's LISTEME values.
- The idiom is only licensed if both searches are successful.

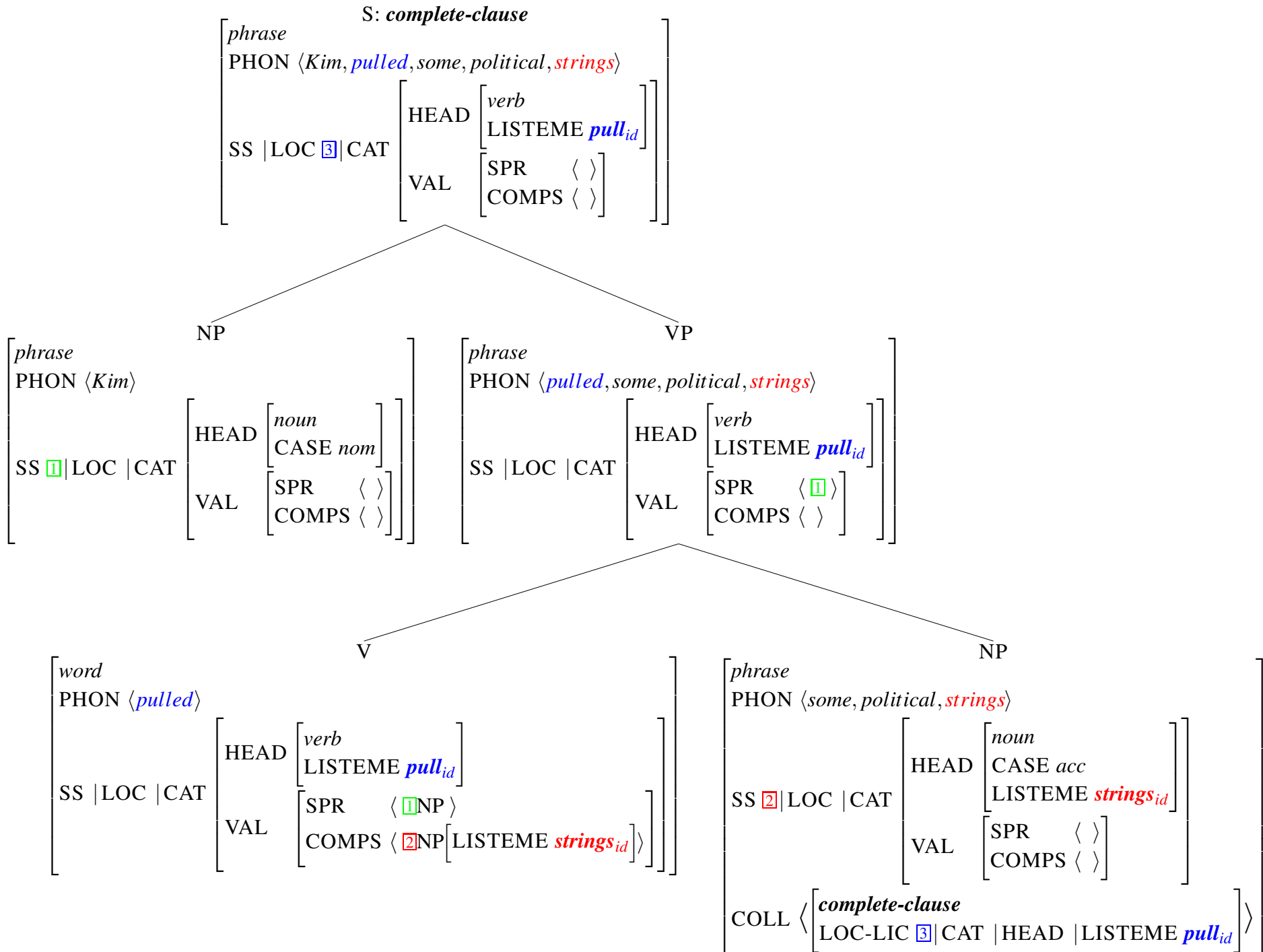
However, there is a difference in how the verb searches for the LISTEME value of the noun and how the noun searches for the LISTEME value of the verb:

- Idiomatic *pull* searches for *strings_{id}* via the standard selection mechanism, that is via its SPR and COMPS list, by specifying the SYNSEM value on the respective list accordingly.
 - It can do that because the Head Feature Principle ensures that the value of the head feature LISTEME is present at the maximal projection of *strings*, so that *pull* can see it and select for it.
- Idiomatic *strings*, on the other hand, canNOT use the standard selection mechanism to find *pull*, as there is no way for *strings* to select *pull_{id}* via its SPR or COMPS list.
 - Here, we need the COLL feature, by whose value *strings* can require to occur within a specific kind of phrase with a specific LISTEME value, namely *pull_{id}*.



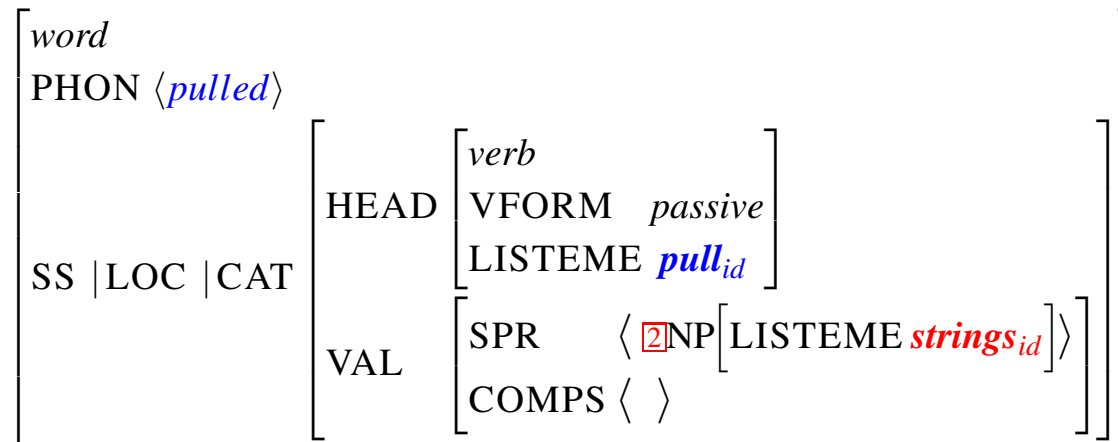
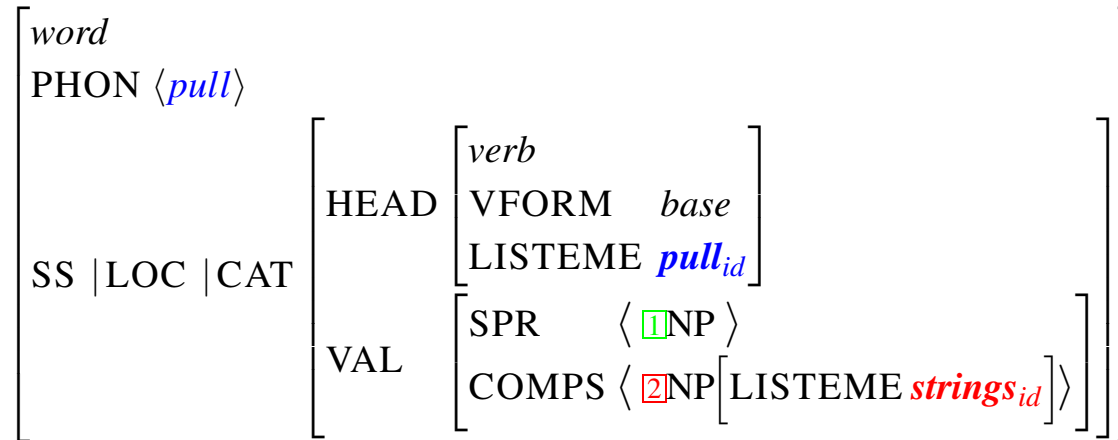


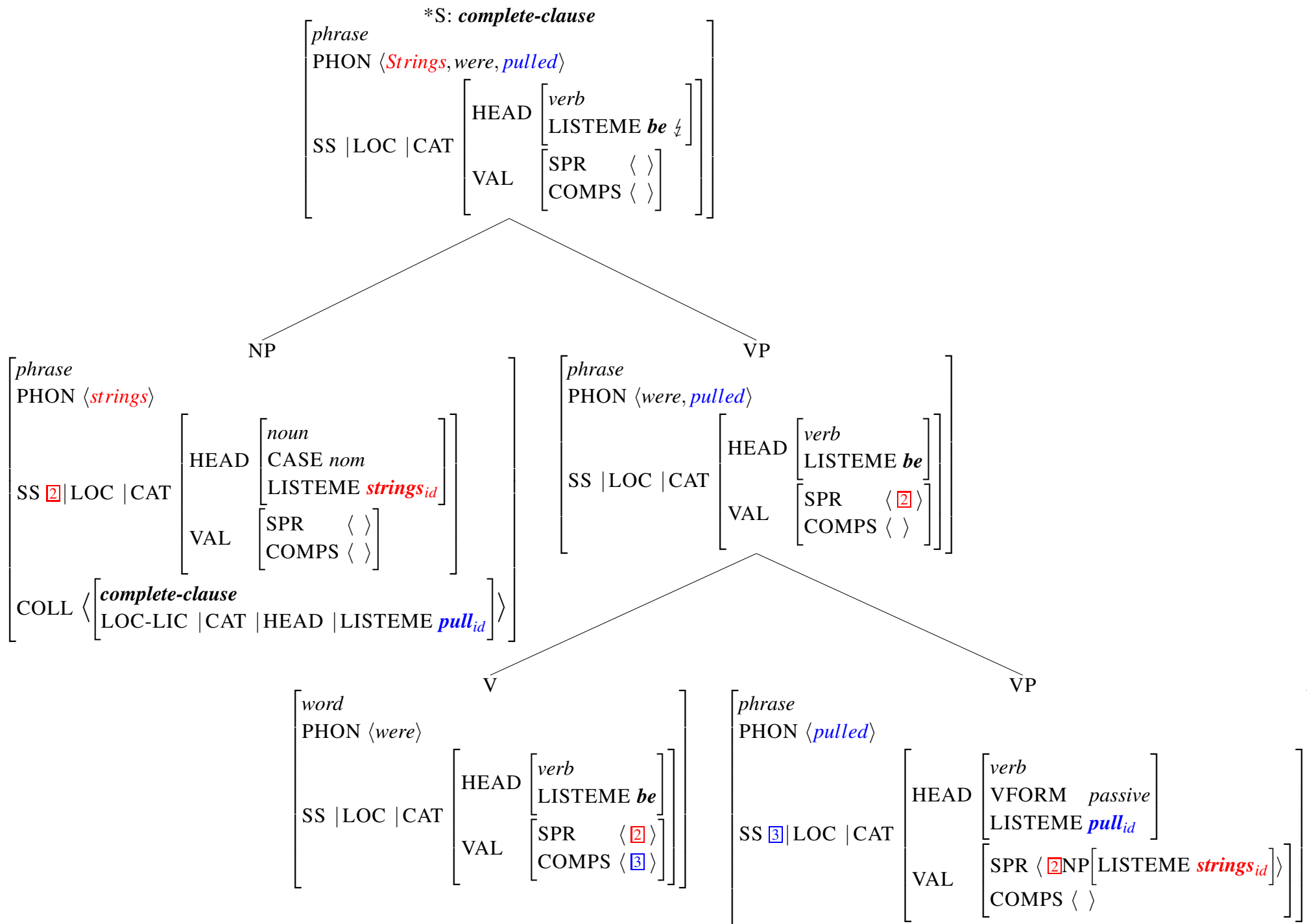




Passivization

In HPSG, passivization is taken care of in the lexicon by the **PASSIVE LEXICAL RULE**:





Basic Idea:

As LISTEME is a head feature, the HEAD FEATURE PRINCIPLE ensures that it is always present at the maximal projection of the word it originates from.

Problem:

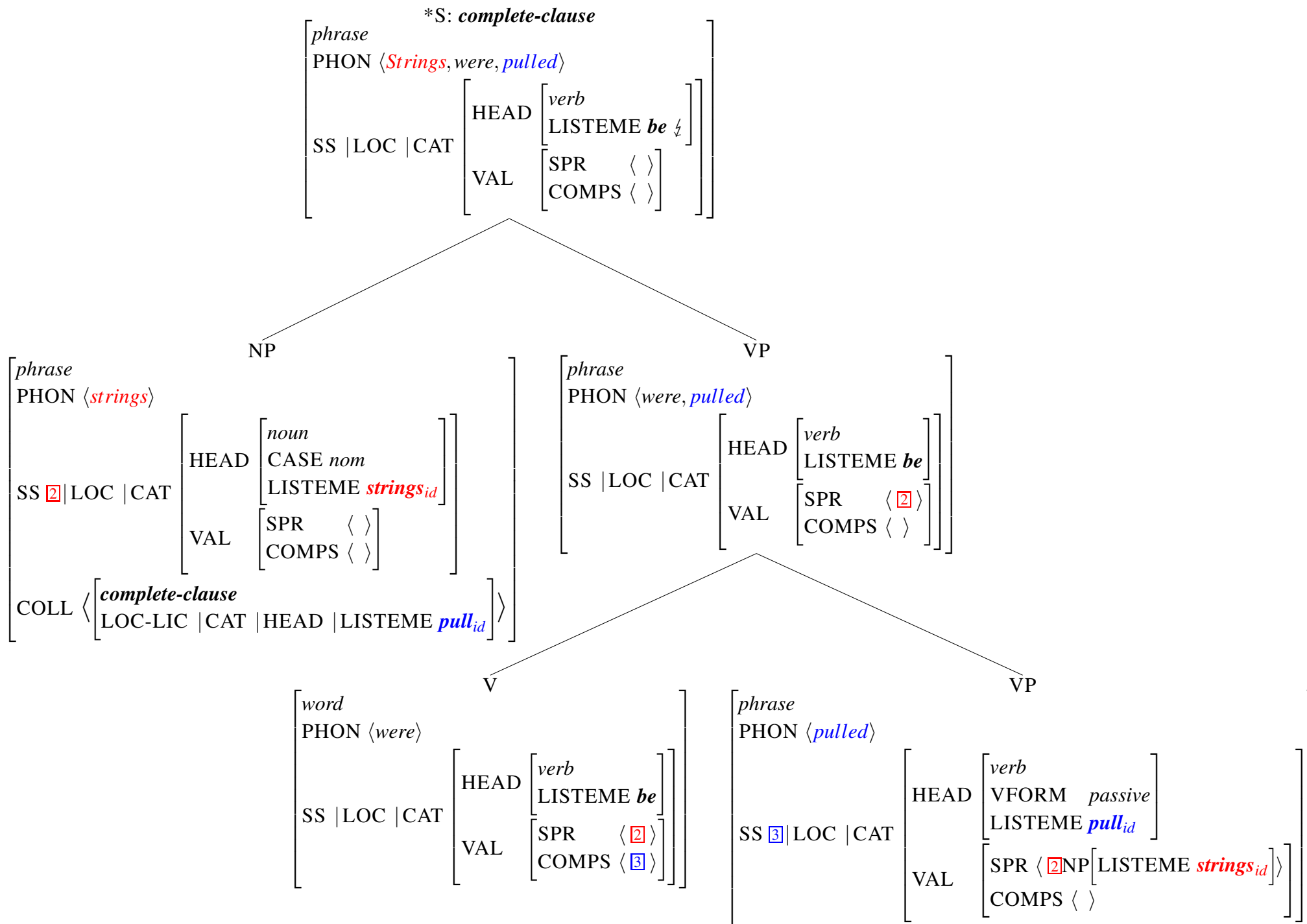
In a passive sentence like *Strings were pulled*, however, *pull* is not the head of the VP that takes *strings* as its specifier.

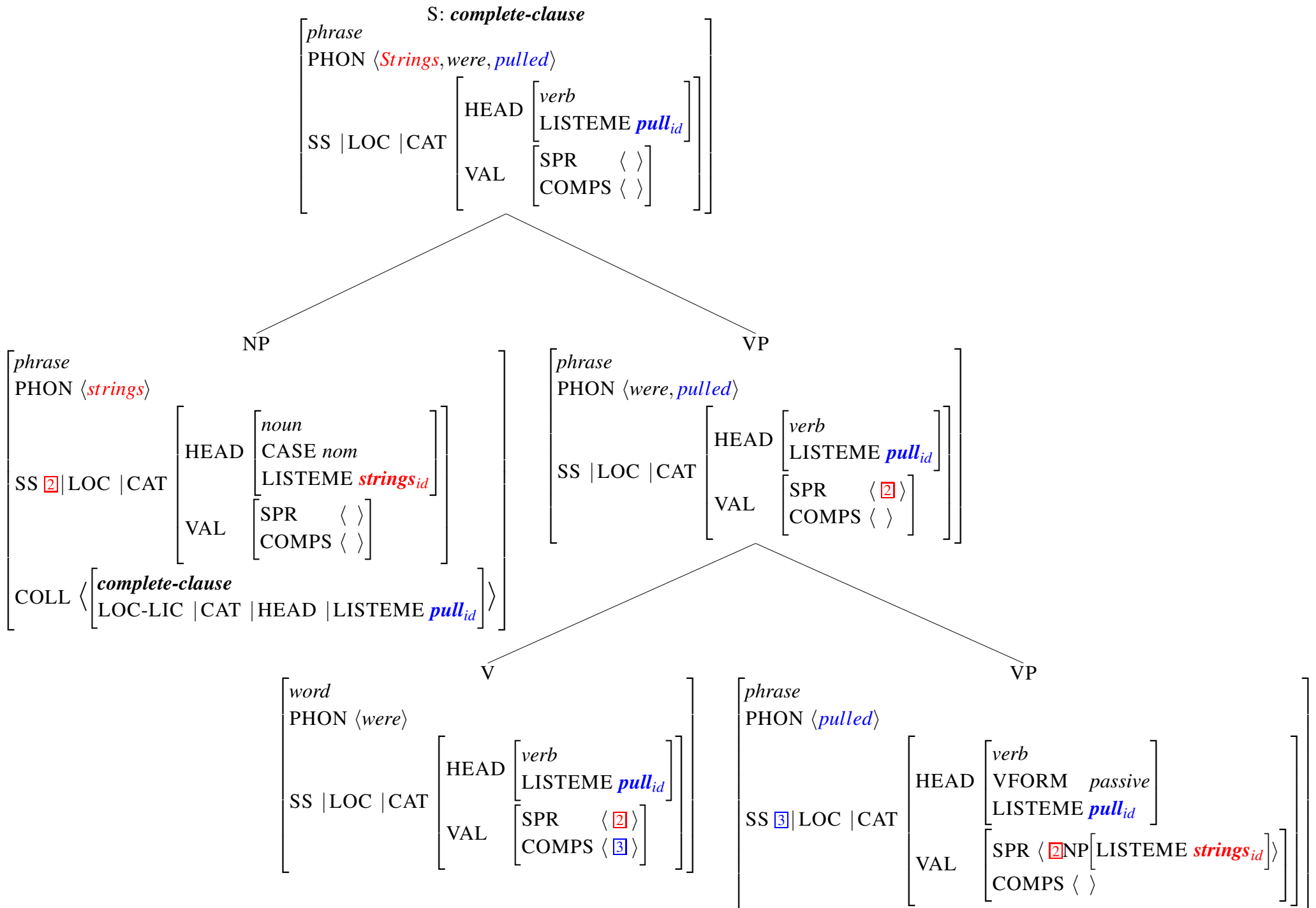
NB:

This problem also occurs in any sentence in which there is an auxiliary on top of *pull*!

Solution:

Söhn would have to assume that a raising verb not only adopts the SPR value of its complement (as is commonly assumed in HPSG), but also the LISTEME value.





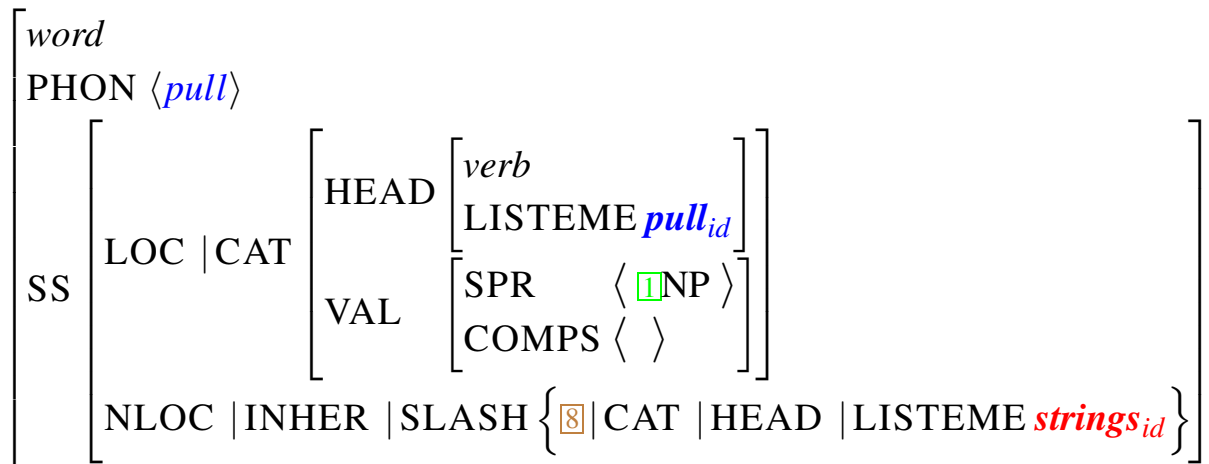
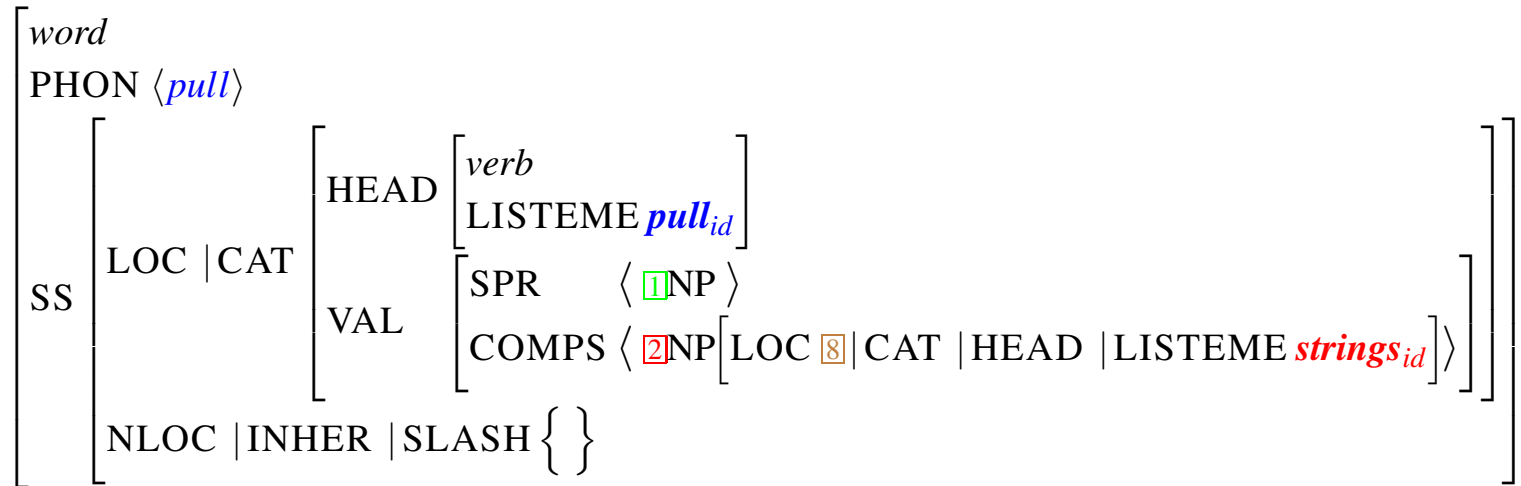
Topicalization

In HPSG, 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. **the 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. **the COMPLEMENT EXTRACTION LEXICAL RULE** on the next slide.

Topicalization

COMPLEMENT EXTRACTION LEXICAL RULE:



Pronominalization

Söhn's PRONOUN LISTEME CONSTRAINT:

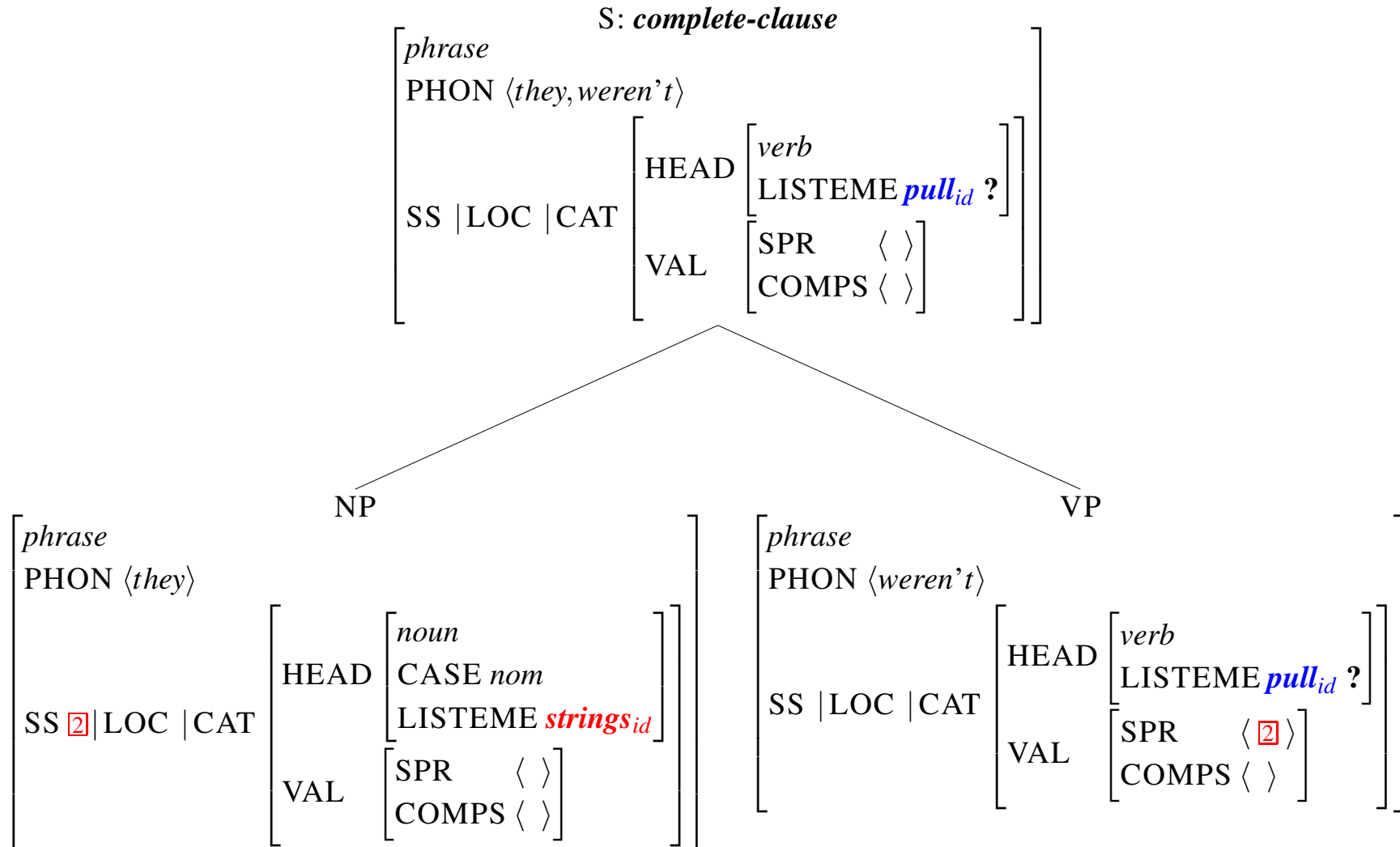
”If a pronoun is co-indexed with a specific antecedent, it adopts its LISTEME value. If this is not the case, the LISTEME value of the pronoun is identical with the PRO-LISTEME value.”

(Söhn 2006:100, my own translation)

For our idiom data, only the first clause/scenario plays a role, which simply states:

Pronouns adopt the LISTEME value of their antecedents.

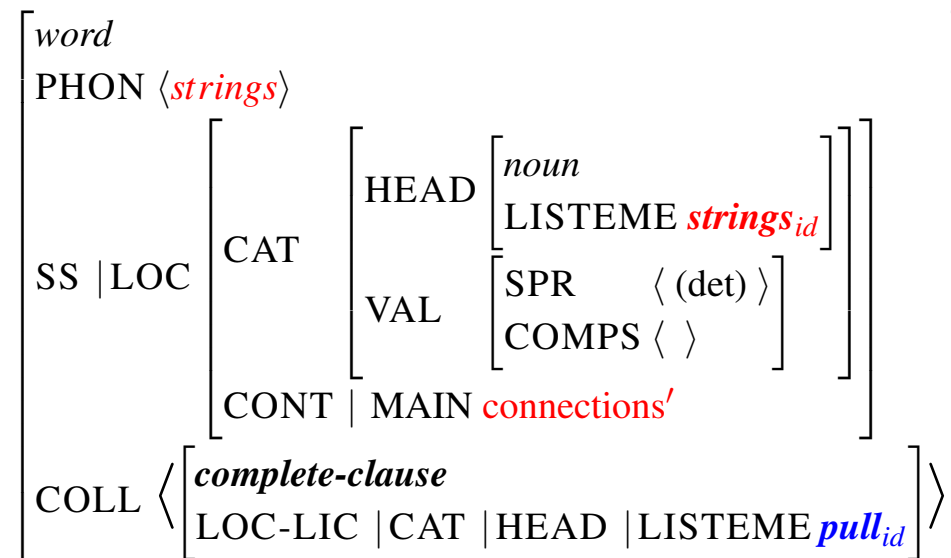
I was worried that *strings* might be *pulled*,
 but fortunately



Kim and Chris graduated from law school together with roughly equal records. Kim's uncle is a state senator, and he pulled strings to get Kim a clerkship with a state supreme court justice. Chris didn't have access to any strings, so he ended up hanging out a shingle.

(Wasow et al. 1983)

Söhn's approach cannot account for the last sentence of the above mini discourse, as the idiomatic noun *strings* with the unique LISTEME value *strings_{id}* wants to occur within a *complete-clause* that has the unique LISTEME value *pull_{id}*:



beans and *strings* also occur in the N-*after*-N construction (cf. Jackendoff 2008)

An example from the Web:

- (7) The whole idea of the really talented/successful person in their 20's isn't a real thing. Or at the very least, it isn't an actual attainable thing. All those people have people behind them **pulling string** after **string** for them. Rich parents, well-connected parents, well-connected god parents ... Whatever it is, I can guarantee you it's there somewhere.³

Dilemma for Söhn:

If he assigns LISTEME values to lexemes (not words), both the singular and the plural form of idiomatic *strings* have the LISTEME value *strings_{id}*. In that case, the above data can potentially be accounted for, but at the same time this causes a problem with the fact that a sentence like *Kim pulled a string* cannot be interpreted idiomatically.

If, on the other hand, he assigns LISTEME values to words (not lexemes), he can prevent *Kim pulled a string* from being grammatical in the idiomatic sense by assigning the LISTEME value *strings_{id}* to the idiomatic word *strings*, but then he cannot account for the N-*after*-N data.

³source: <http://doiwakeorsleep.tumblr.com/post/46393612094/25-things-i-learned-in-the-first-half-of-my-20s>

Evaluation of Söhn's COLL Account

What Söhn's approach can account for:

- Determiner Variation in the NP
- Adjectival Modification in the NP
- Topicalization of the NP (as long as it is not long-distance)
- Pronominalization of the NP

What Söhn's approach, as it stands, cannot account for:

- Passivization and any sentence in which the idiomatic verb is not the head
- Isolated Occurrence of the NP
- N-*after*-N Construction

End of the first talk ...

Thank you!

Approaching Idioms Less Syntactically: The Semantic Representation Theory

Sascha Bargmann
University of Frankfurt

TAU-GU

29 November 2014

Based on:
“Idioms as Evidence for the Proper Analysis of Relative Clauses”
(Webelhuth, Bargmann, and Götze to appear)

Outline

1. The SEMANTIC REPRESENTATION (SR) Account
 - (a) Theoretical Assumptions
 - (b) Application to the English Idiom Data
2. Comparative Evaluation of Söhn's COLL and the SR Account
3. Overall Summary

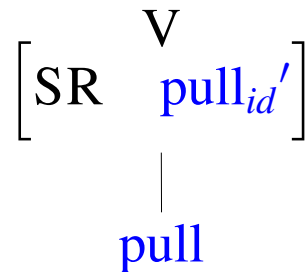
The SEMANTIC REPRESENTATION (SR) Account

Theoretical Assumptions on Syntax, Semantics, and Idioms:

- **Syntax:** Head-Driven Phrase Structure Grammar (HPSG)
- **Semantics:**
 - Representation language: predicate logic + generalized quantifiers + lambda calculus
 - Combinatorics: a version of Flexible Montague Grammar with lexical type shifting and functional application at phrasal nodes
 - Semantic contribution of a sign: expression of the semantic representation language
- **Idioms:**
 - Each part of the idiom makes a unique contribution to the SEMANTIC REPRESENTATION of the larger linguistic context.
 - A part of an idiom can 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.

pull strings

Lexical Entry



Constraint

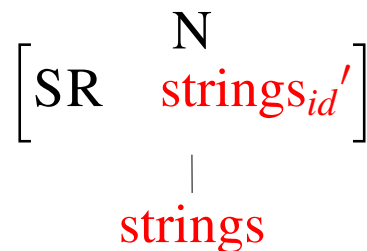
In the SEMANTIC REPRESENTATION, the verb's second argument is specified by a quantifier Qx that is restricted by $\text{strings}_{id'}(x)$ (after anaphor resolution).

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

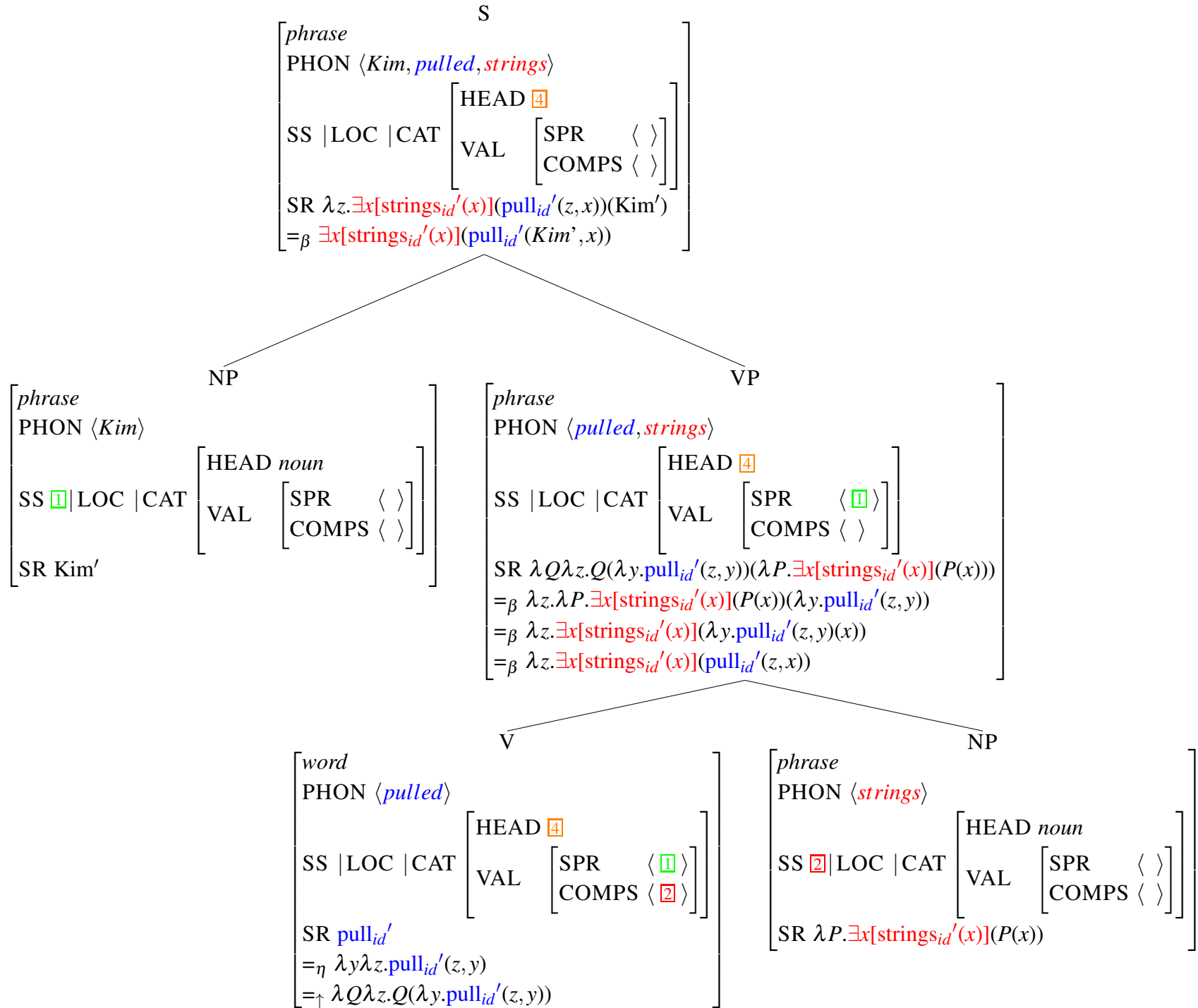
a. Qx binds the second argument of $\text{pull}_{id'}$ (after anaphor resolution)

or

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

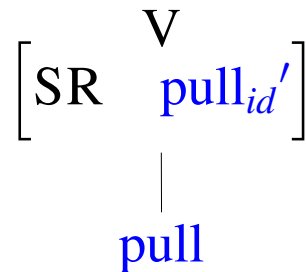


Crucial Assumption: The above-mentioned SEMANTIC REPRESENTATIONS only occur in the lexical entries themselves and in pronouns and ellipsis sites licensed by them.



pull strings

Lexical Entry



Constraint

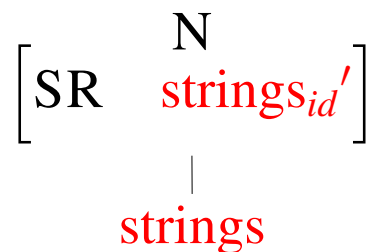
In the SEMANTIC REPRESENTATION, the verb's second argument is specified by a quantifier Qx that is restricted by $\text{strings}_{id'}(x)$ (after anaphor resolution).

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

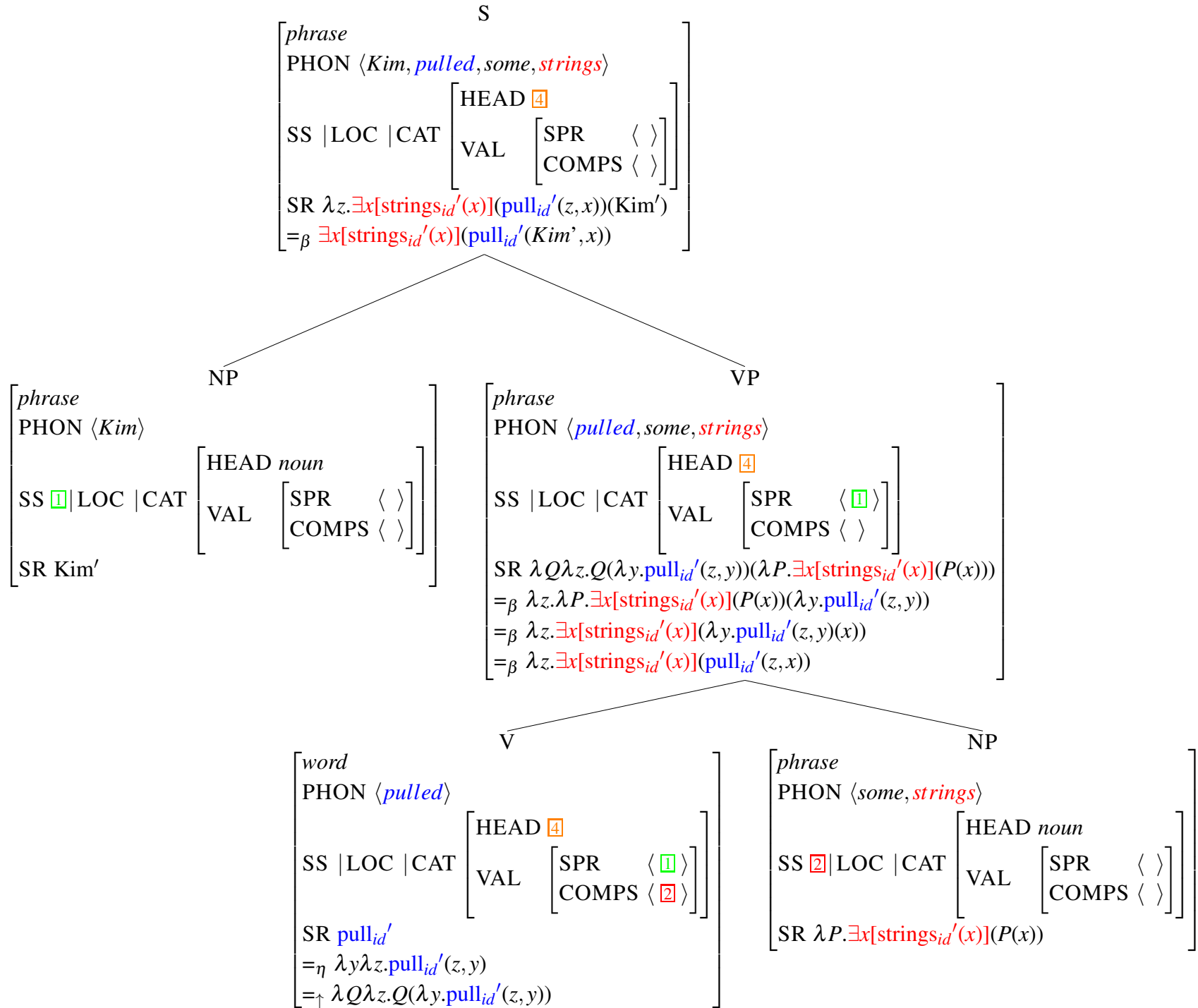
a. Qx binds the second argument of $\text{pull}_{id'}$ (after anaphor resolution)

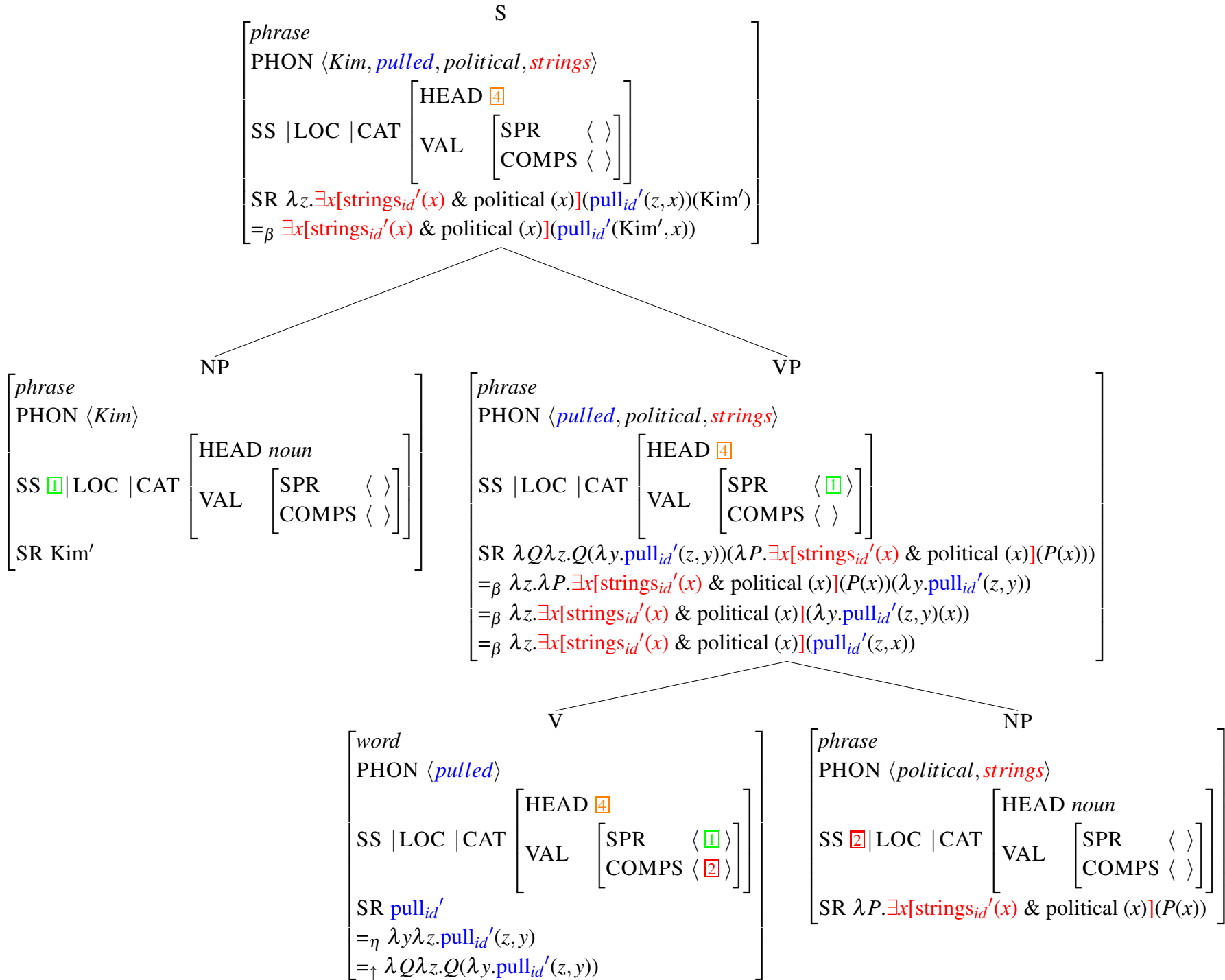
or

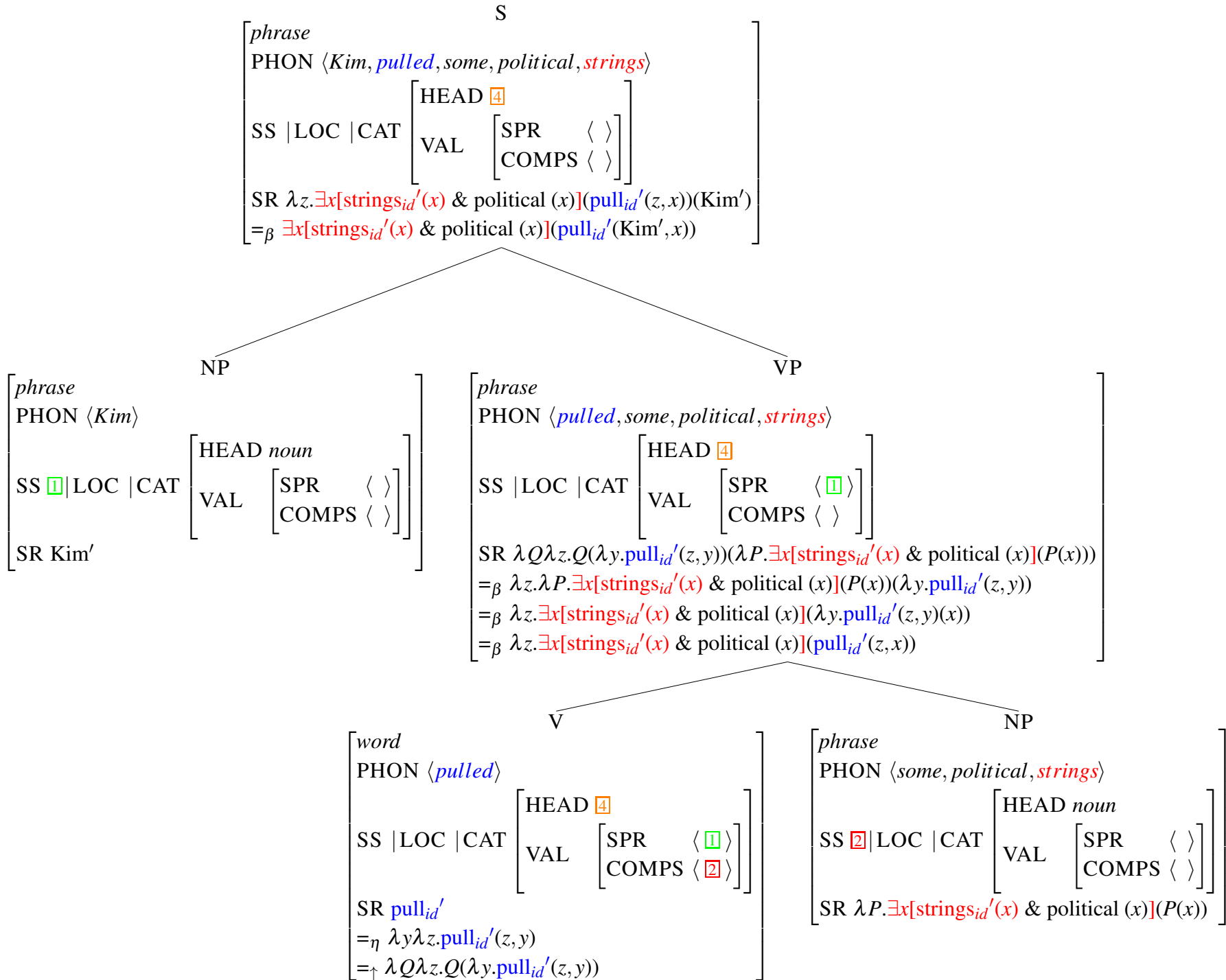
b. $\text{strings}_{id'}$ is salient in the present discourse.



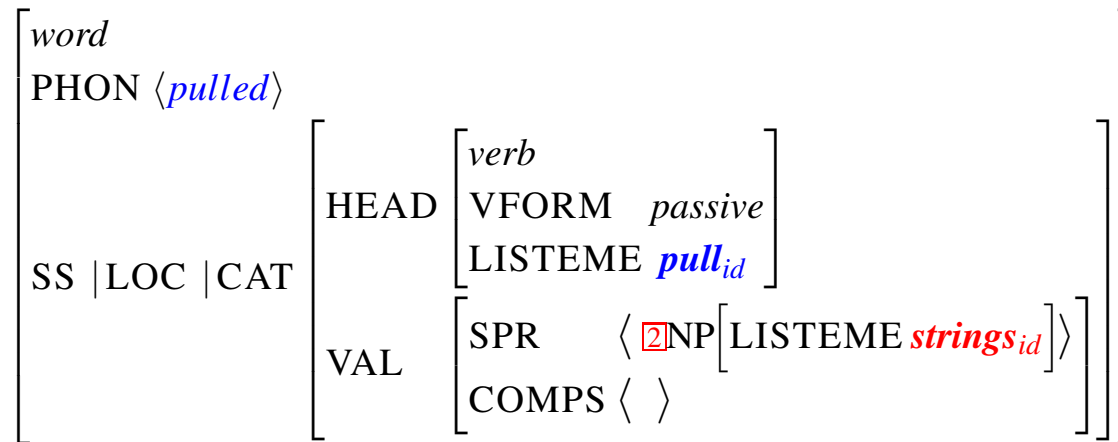
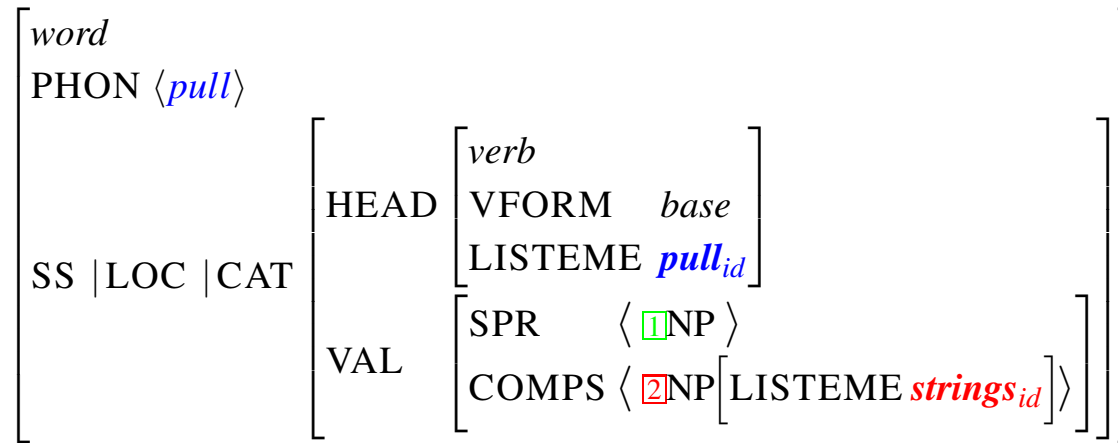
Crucial Assumption: The above-mentioned SEMANTIC REPRESENTATIONS only occur in the lexical entries themselves and in pronouns and ellipsis sites licensed by them.



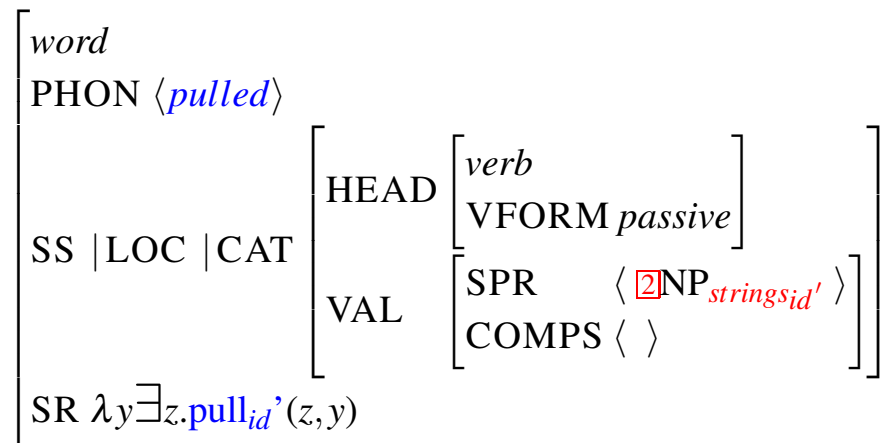
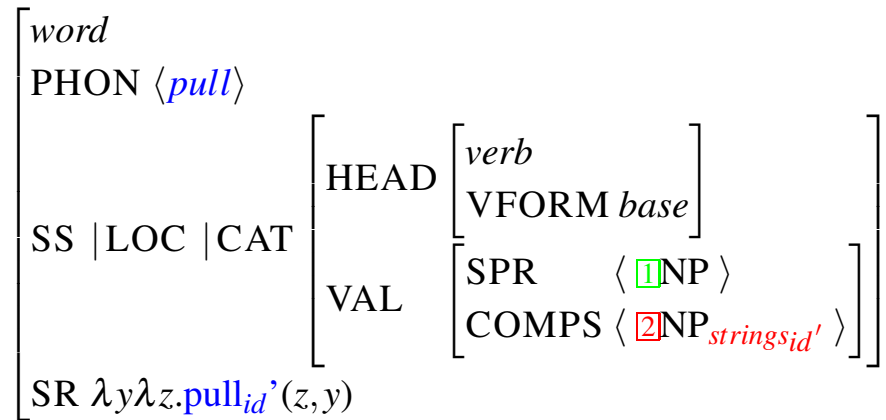


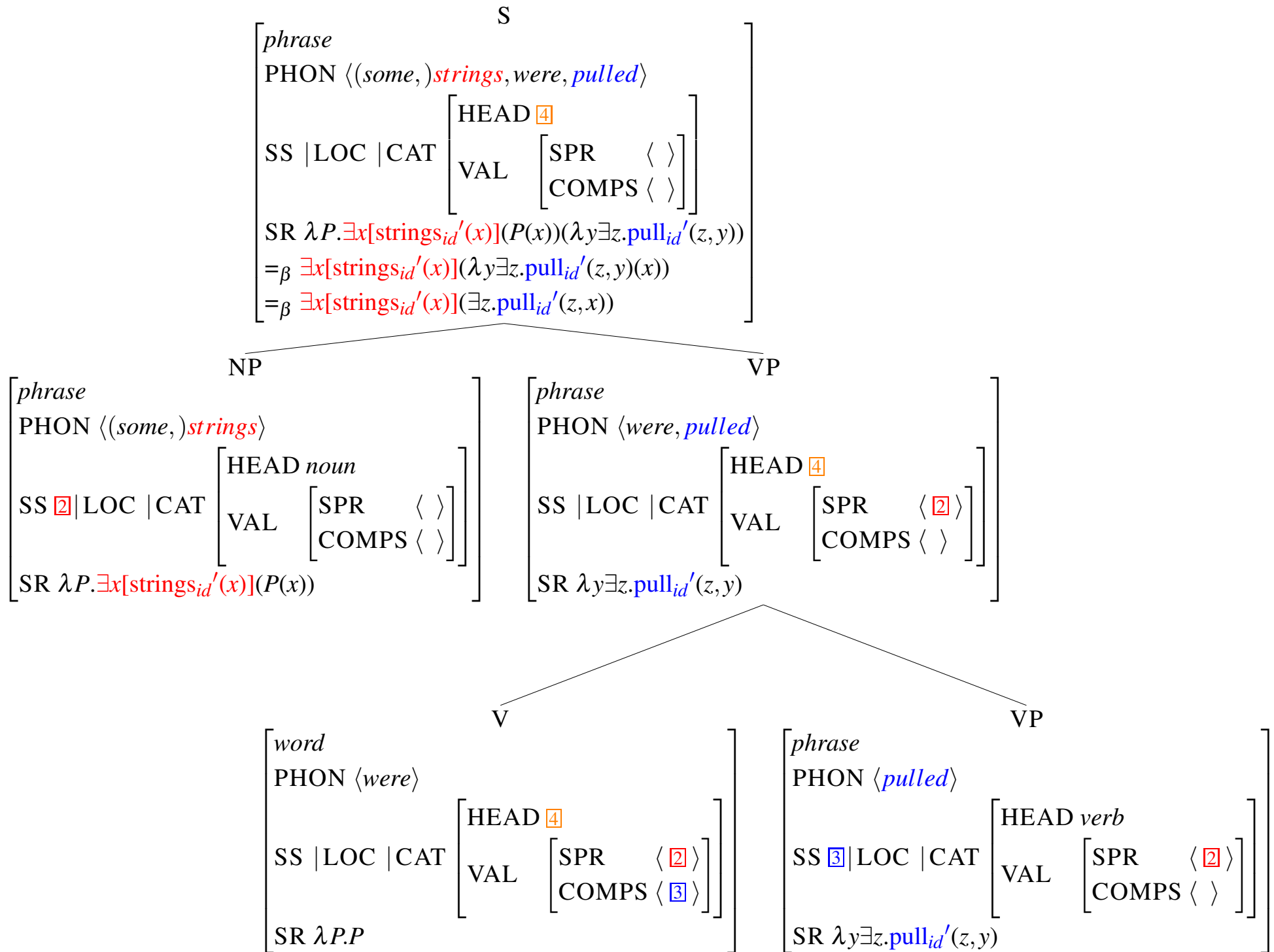


pull undergoing the PASSIVE LEXICAL RULE – LISTEME Version:

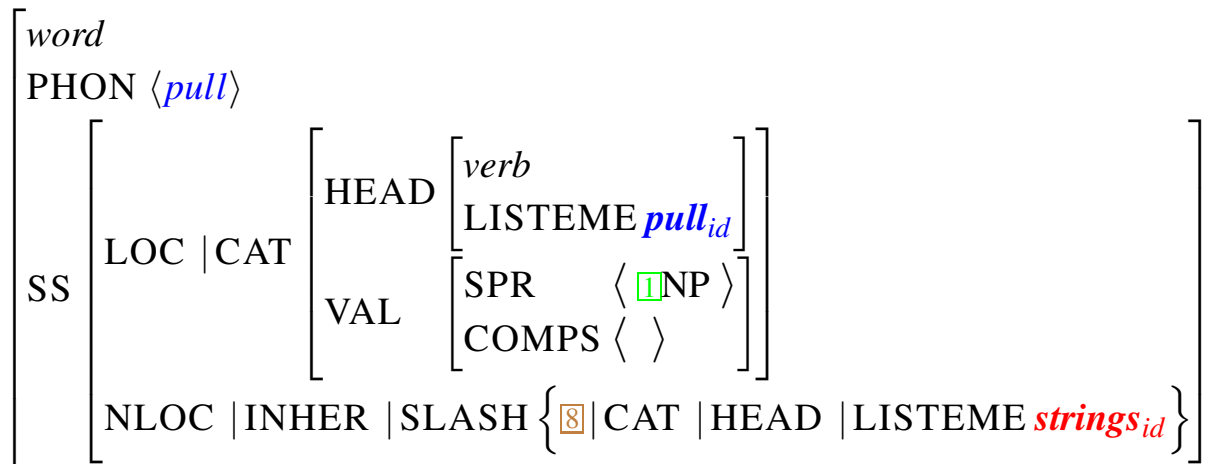
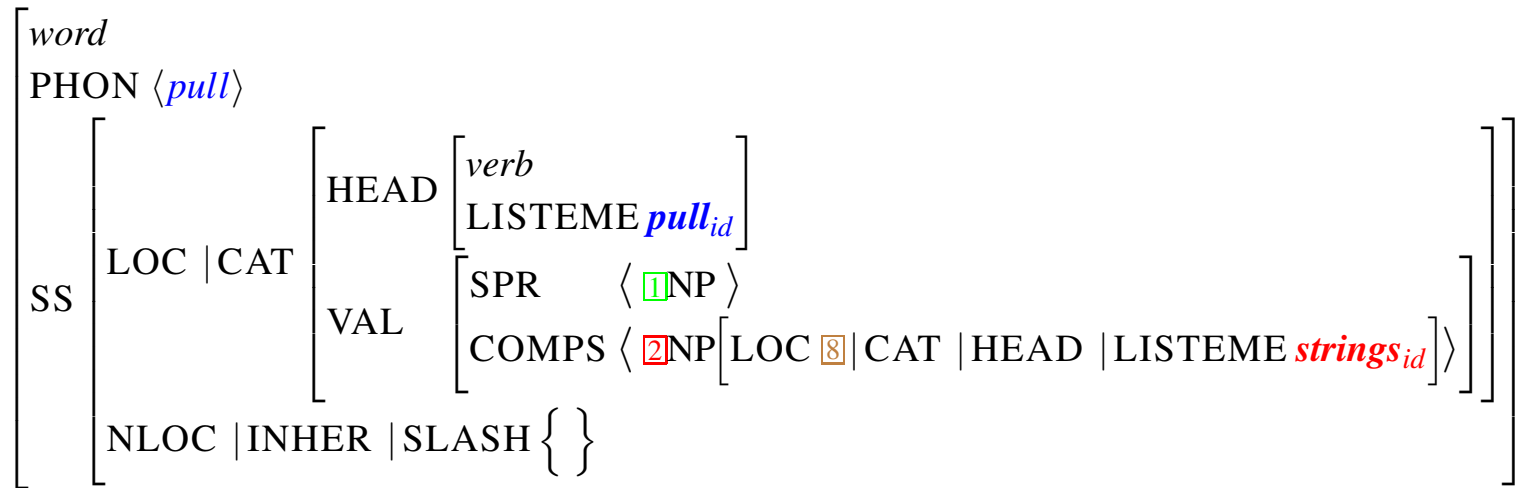


pull undergoing the PASSIVE LEXICAL RULE – SR Version:

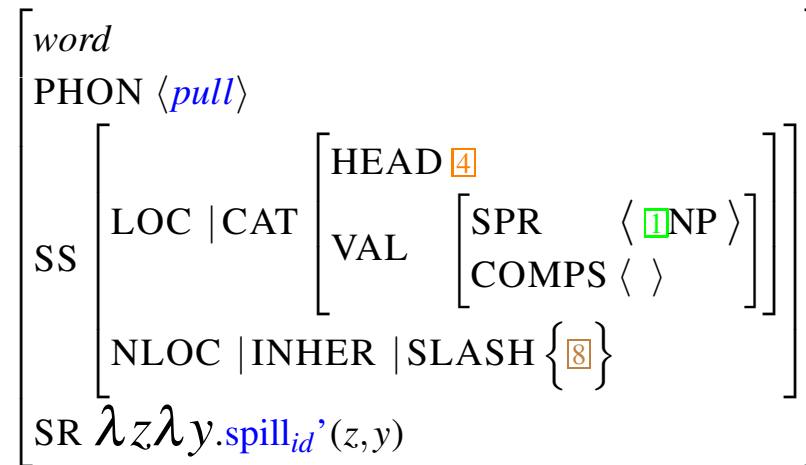
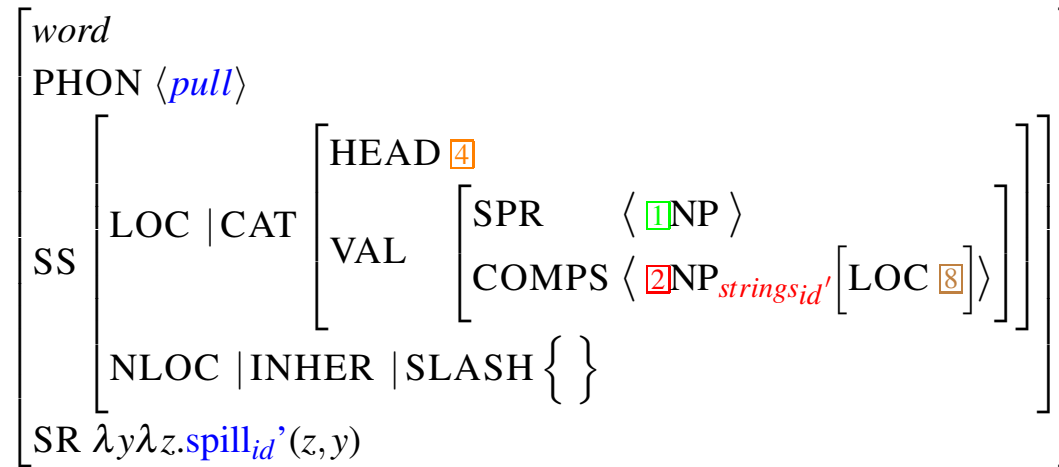


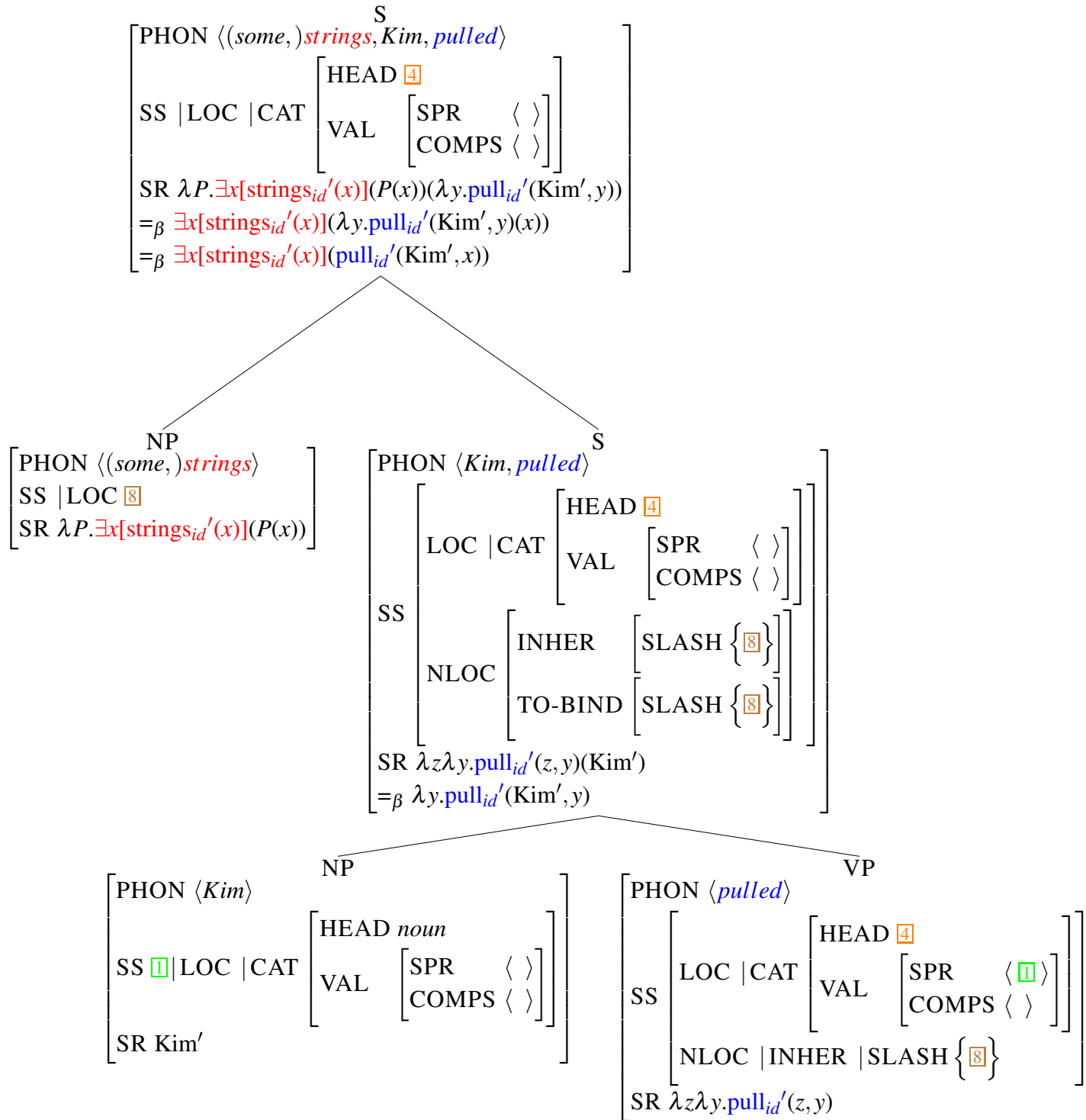


pull undergoing the COMPLEMENT EXTRACTION LEXICAL RULE – LISTEME Version:



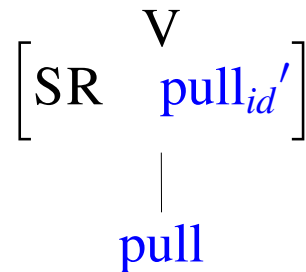
pull undergoing the COMPLEMENT EXTRACTION LEXICAL RULE – SR Version:





pull strings

Lexical Entry



Constraint

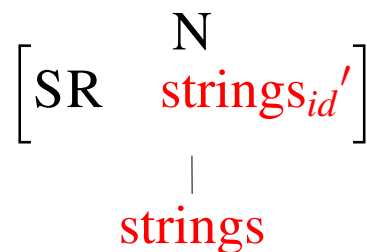
In the SEMANTIC REPRESENTATION, the verb's second argument is specified by a quantifier Qx that is restricted by $\text{strings}_{id'}(x)$ (after anaphor resolution).

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

a. Qx binds the second argument of $\text{pull}_{id'}$ (after anaphor resolution)

or

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



Crucial Assumption: The above-mentioned SEMANTIC REPRESENTATIONS only occur in the lexical entries themselves and in pronouns and ellipsis sites licensed by them.

The SEMANTIC REPRESENTATION (SR) Account

Theoretical Assumptions on Syntax, Semantics, and Idioms:

- **Discourse:**

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.

- **Pronouns:**

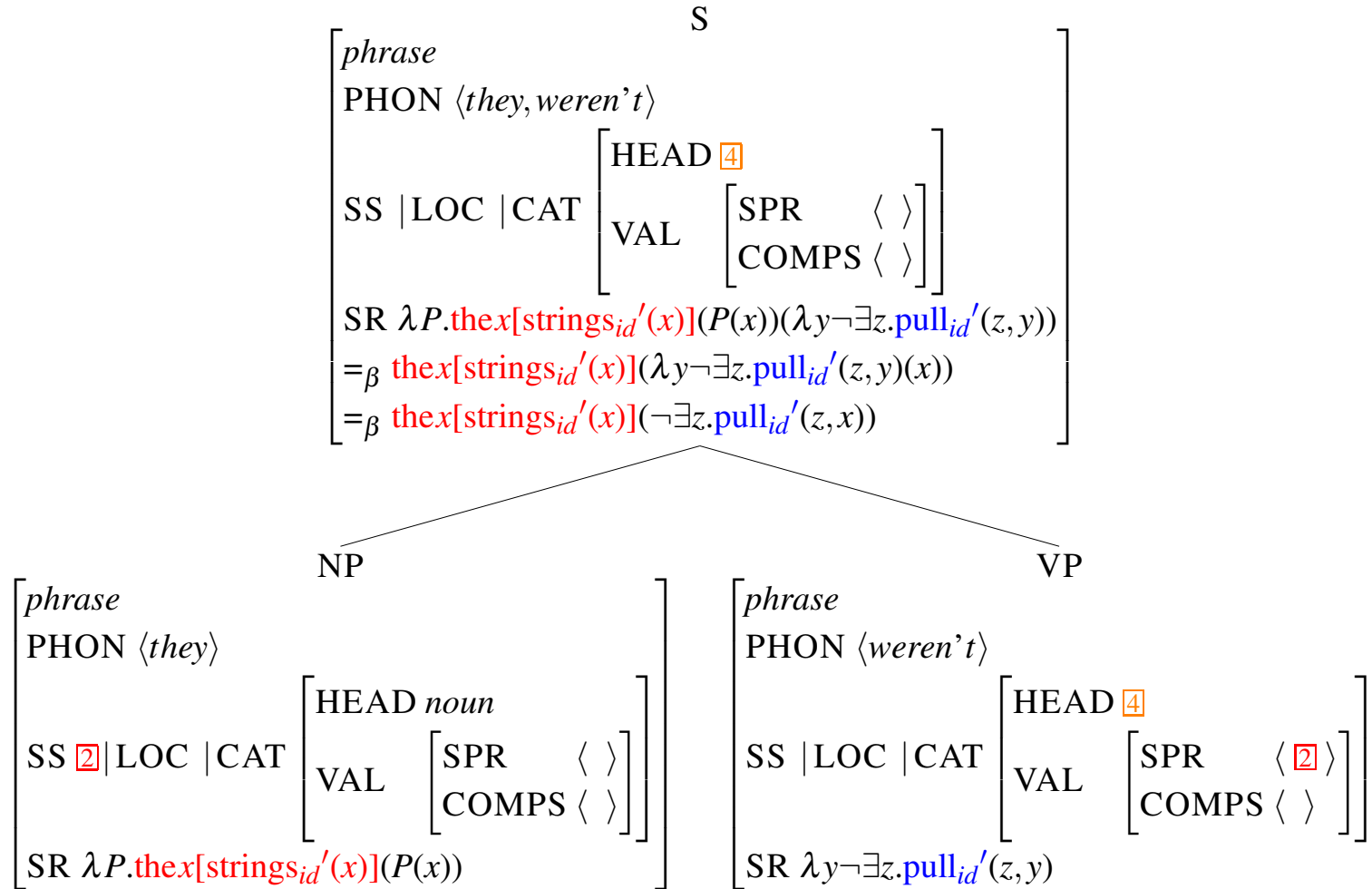
A personal pronoun is interpreted as a definite NP whose restrictor is identical to the restrictor of its antecedent:

- (8) 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** $x[\text{woman}'(x)](\text{whistle}'(x))$

I was worried that *strings* might be *pulled*,

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

but fortunately



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

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

$$\begin{array}{c}
 \text{S} \\
 \left[\begin{array}{l}
 \textit{phrase} \\
 \text{PHON } \langle \textit{Chris, didn't, have, access, to, any, strings} \rangle \\
 \text{SS} \mid \text{LOC} \mid \text{CAT} \left[\begin{array}{l}
 \text{HEAD } \textit{verb} \\
 \text{VAL} \left[\begin{array}{l}
 \text{SPR} \quad \langle \rangle \\
 \text{COMPS} \langle \rangle
 \end{array} \right]
 \end{array} \right] \\
 \text{SR } \neg \exists x [\textit{strings}_{id'}(x)] (\textit{access}'(\textit{chris}', x))
 \end{array} \right]
 \end{array}$$

The N-*after*-N construction

Our example from the Web:

- (9) The whole idea of the really talented/successful person in their 20's isn't a real thing. Or at the very least, it isn't an actual attainable thing. All those people have people behind them **pulling string** after **string** for them. Rich parents, well-connected parents, well-connected god parents ... Whatever it is, I can guarantee you it's there somewhere.

As the NP *string after string* is only morphosyntactically singular:

- (10) [_{NP} String after string] HAS been pulled.

... but semantically plural:

Whenever you have pulled string after string, you have pulled strings.

... a more semantically oriented approach like the SR theory is much better suited to account for these data.

Evaluation of the SR Theory

What the SR theory can account for:

- Determiner Variation in the NP
- Adjectival Modification in the NP
- Passivization and sentences in which the idiomatic verb is not the head
- Topicalization of the NP
- Pronominalization of the NP
- Ellipsis of the idiomatic verb
- Isolated Occurrence of the idiomatic NP

What the SR theory, as it stands, cannot account for:

- N-*after*-N Construction

Comparative Evaluation of Söhn's COLL and the SR Account

| | SR Account |
|-----------------------------------|------------|
| Determiner Variation in the NP | ✓ |
| Adjectival Modification in the NP | ✓ |
| NP Movement | ✓ |
| <i>Wh</i> Movement | ✓ |
| Pronominalization of the NP | ✓ |
| Ellipsis of the Verb | ✓ |
| Isolated Occurrence of the NP | ✓ |
| N- <i>after</i> -N Construction | (*) |

Comparative Evaluation of Söhn's COLL and the SR Account

| | SR Account | Söhn's COLL |
|-----------------------------------|------------|-------------|
| Determiner Variation in the NP | ✓ | ✓ |
| Adjectival Modification in the NP | ✓ | ✓ |
| NP Movement | ✓ | * |
| <i>Wh</i> Movement | ✓ | (✓) |
| Pronominalization of the NP | ✓ | ✓ |
| Ellipsis of the Verb | ✓ | ✓ |
| Isolated Occurrence of the NP | ✓ | * |
| N- <i>after</i> -N Construction | (*) | * |

Comparative Evaluation of Söhn's COLL and the SR Account

| | SR Account | Söhn's COLL |
|-----------------------------------|------------|-------------|
| Determiner Variation in the NP | ✓ | ✓ |
| Adjectival Modification in the NP | ✓ | ✓ |
| NP Movement | ✓ | * |
| <i>Wh</i> Movement | ✓ | (✓) |
| Pronominalization of the NP | ✓ | ✓ |
| Ellipsis of the Verb | ✓ | ✓ |
| Isolated Occurrence of the NP | ✓ | * |
| <i>N-after-N</i> Construction | (*) | * |

Conclusion:

The SR theory can account for a larger subset of the English idiom data.

Overall Summary

1. Idioms differ from each other in their degree of syntactic flexibility, and there are at least two types of English VP idioms:
 - (a) Syntactically almost entirely frozen idioms like *kick the bucket*
 - (b) Syntactically very flexible idioms like *spill beans* or *pull strings*
2. Both Söhn's COLL account and the SR theory assign each of the parts of a syntactically flexible idiom its own lexical entry. This entry indicates what other element(s) the idiom part has to co-occur with. In Söhn's COLL account this is done via the value of the morphosyntactic feature LISTEME, in the SR theory via the SEMANTIC REPRESENTATION contribution.
3. The SR theory can account for a larger subset of the English idiom data presented, which is partly, but not entirely, due to the fact that it is compatible with discourse theory and can therefore also account for intersentential phenomena of idioms.
4. Söhn's COLL account involves the stipulation of the LISTEME feature and a special pronoun rule. The LF theory involves neither of these and, least to me, seems more intuitive, as the entire subject has a semantic feel to it.

End of the second talk ...

You've made it!

References

- Rodney Huddleston and Geoffrey K. Pullum. *The Cambridge Grammar of the English Language*. Cambridge University Press, Cambridge, 2002.
- Ray Jackendoff. Construction after construction and its theoretical challenges. *Language*, 84 (1):8–28, 2008.
- Jan-Philipp Söhn. *Über Bärendienste und erstaunte Bauklötze. Idiome ohne freie Lesart in der HPSG*. Peter Lang, Frankfurt am Main, 2006. Ph.D. thesis, Friedrich-Schiller-Universität Jena.
- 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.