

Intermediate Grammaticality

Sandra Villata

Dept. of Linguistics and Psychological Sciences, University of Connecticut

Logic Colloquium, UConn, Nov 13, 2020

Collaborators



JON SPROUSE



WHIT TABOR

The boy love the book



The boy loves the book

What did you think the boy read?



What did you think that the boy read?

What did you wonder whether the boy read?



... ?

What did you smile because the boy read?



... ?

What cry if why east and the angry the?



... ?

Parsing Ungrammaticality

- All sentences above are ungrammatical, but there is no doubt that we understand at least some of them
- Not all grammatical violations are perceived in the same way: some of them are more severe than others
- This observation suggests that there is a sense in which we parse ungrammatical sentences (i.e. we compute and we extract their meaning)
- But how does this happen? Which mechanism or (set of mechanisms) underlie ungrammatical sentence processing?

Traditional assumptions

1. The **grammar** of a given language (e.g. English) is a system of *rules* that generates *all and only* the well-formed structures of that language
2. To explain processing phenomena (i.e. how we understand and/or produce in *real time* the sentences of a given language) we have to add a theory of **parsing** to the theory of grammar
3. The meaning of a sentence is computed on the output of the grammar (i.e. there are two distinct modules, one for syntax and one for **semantics**, which are in a feedforward relation, such that the output of syntax is the input of semantics)
4. The traditional **grammar+parsing theory** can generate output *incrementally* in real time (i.e. word-by-word as the sentence is processed) and for semantics to take the output of the parser and compute meaning for initial strings of well-formed sentences

Two implications ...

The view according to which the **grammar** of a given language (e.g. English) is a **system of rules** that generate *all and only* the well-formed structures of that given language has two main implications:

- (1) ungrammatical sentences are simply not generated by the grammar (the derivation stops as soon as the linguistic input cannot be accommodated by any rule of the grammar)
- (2) the output of the grammar is strictly binary (grammatical vs. ungrammatical)

... and corresponding challenges

1. Ungrammaticality

If the derivation of the sentence crashes as soon as ungrammaticality is detected, and if semantics is computed over syntactic representations, we shouldn't be able to derive the meaning of ungrammatical sentences at all (not even of seemingly easy ones, like “**John love the book*”)

2. Gradient effects

If the grammar produces strictly categorical outcomes (grammatical vs. ungrammatical) then there is no room for intermediacy

The traditional answer

The traditional view replies to these challenges by claiming that whichever system/mechanism is responsible for ungrammatical sentence processing and its gradient patterns, this must be outside of the grammar module

repair strategies

memory
limitations

semantics

reanalysis
mechanisms

interference

...

Outline

1. I will first try to substantiate the claim that **ungrammatical** sentences cannot be generated under the traditional view of grammar
2. I will then turn to **gradient effects**. I will show that gradience is a core phenomenon in language: it is *pervasive* and *measurable*

As a test case, I will focus on one of the most prototypical, and yet arguably most theoretically challenging, syntactic phenomena: **islands** (Ross, 1967). Islands are interesting because, although most linguistic theories claim that sentences violating island constraints are fully ungrammatical and uninterpretable, I will present experimental evidence revealing *gradient patterns of acceptability* and *interpretability for islands*, which calls for a theory of gradience

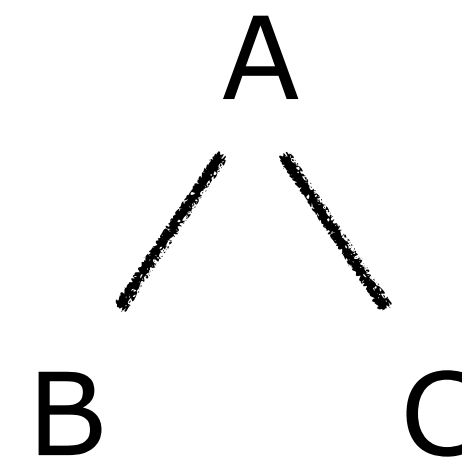
3. To account for these facts, I will present a view that consists in adopting a **more flexible rule-based system** in which sentential elements can be *coerced*, under specific circumstances, to play a role that does not fully fit them. In this system, unlike traditional ones, structure formation is forced even under sub-optimal circumstances, generating semi-grammatical structures

1. Ungrammaticality

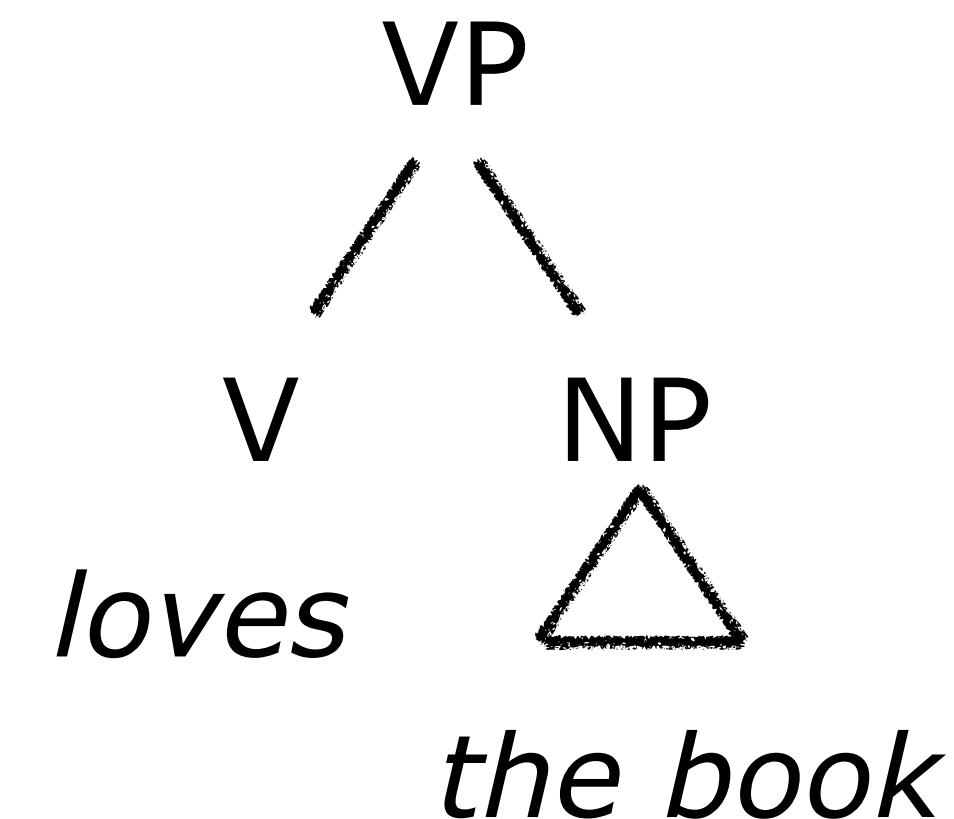
Phrase-structure rules

Grammatical rules are phrase-structure rules taking this general form:

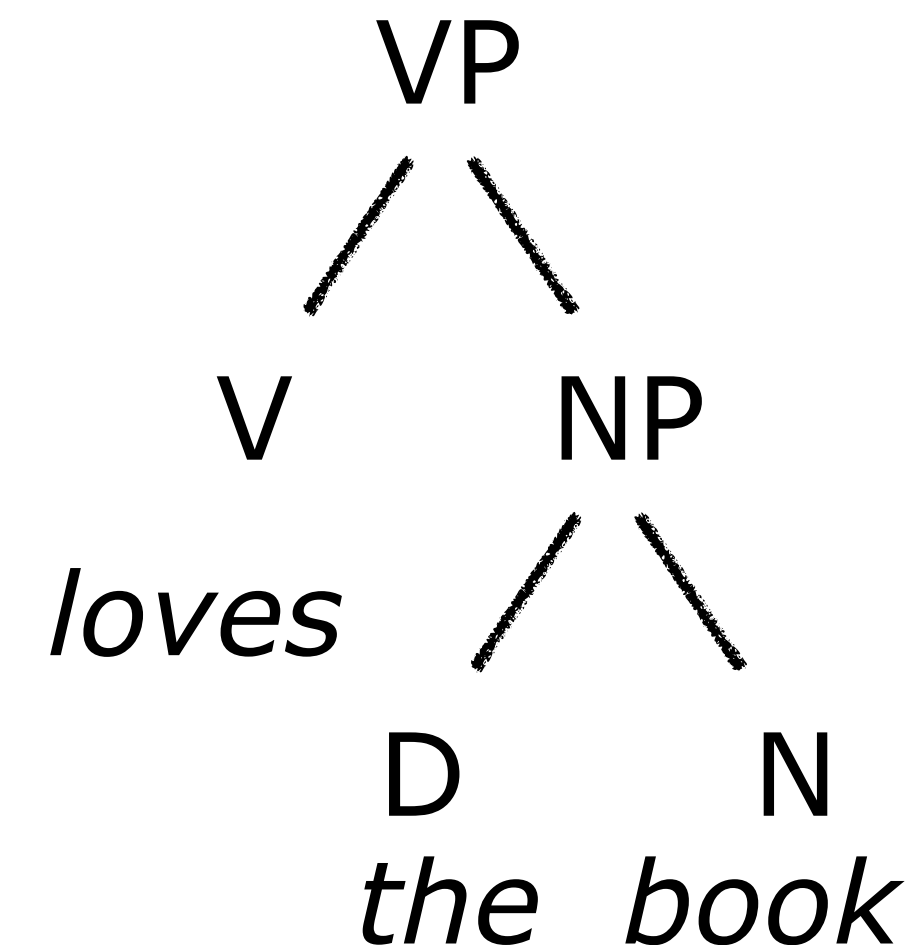
$A \rightarrow B \quad C$
("A rewrites as B and C")



$VP \rightarrow V \quad NP$
(**V**erbal **P**hrase rewrites as **V**erb, **N**oun **P**hrase)



$NP \rightarrow D \quad N$
(**N**oun **P**hrase rewrites as **D**eterminer, **N**oun)



Generating sentences

"John loves the book"

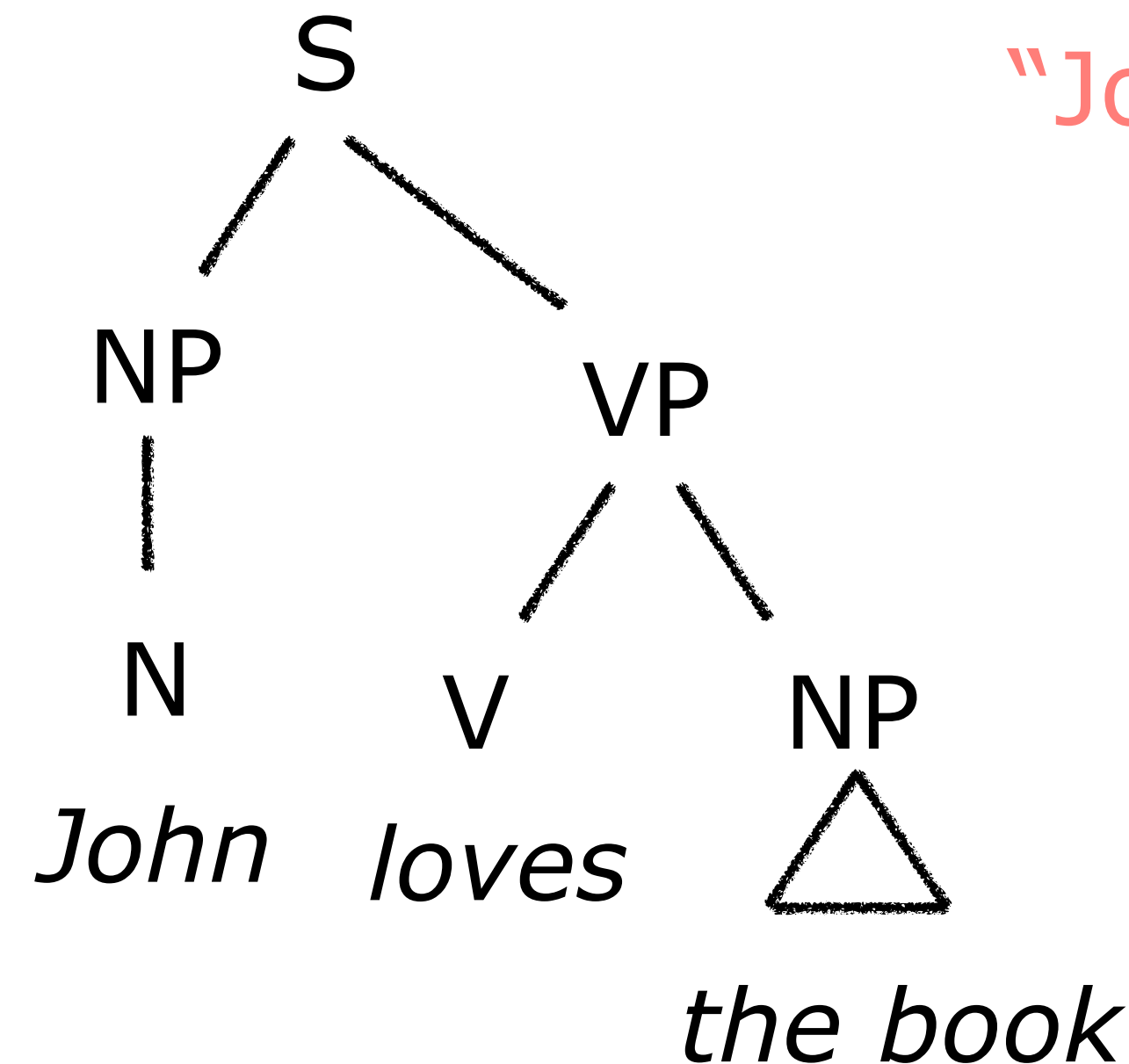
NP

VP

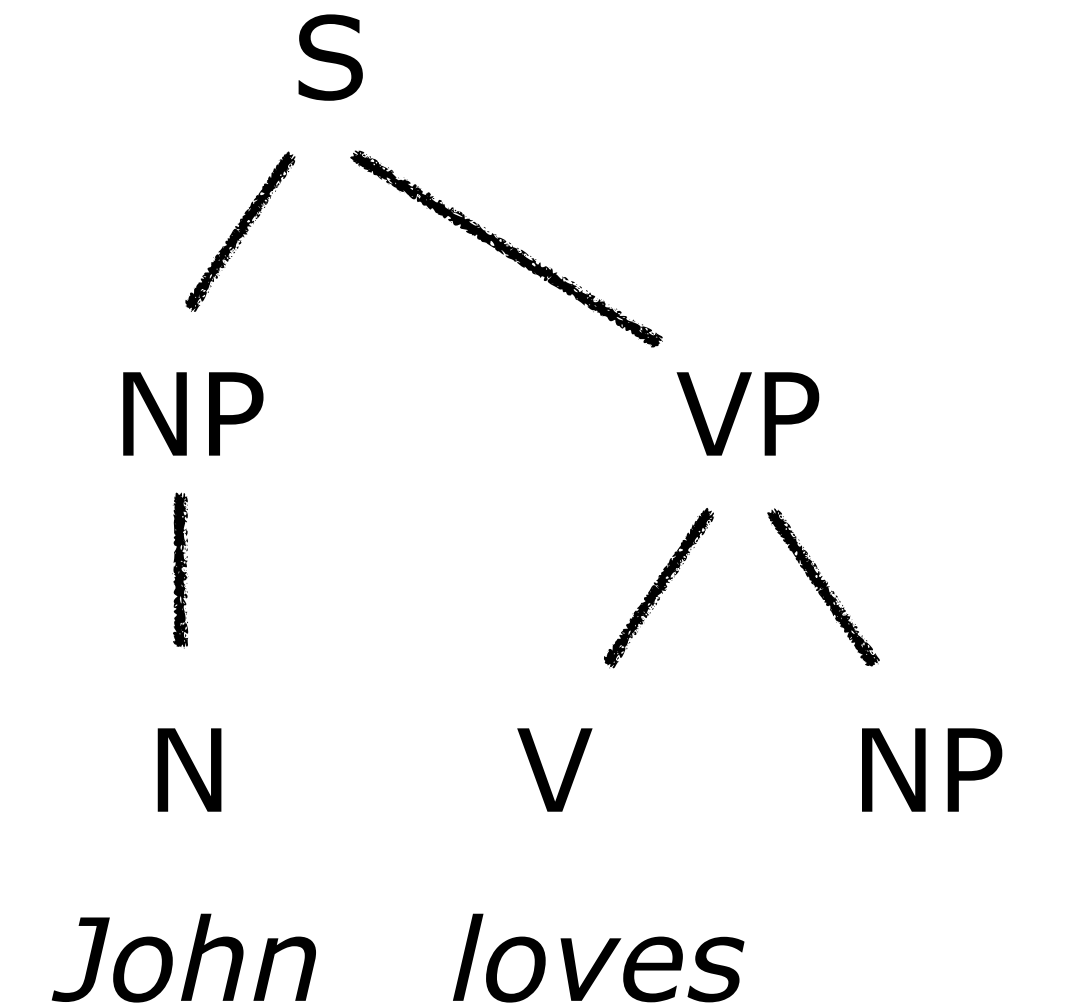
$S \rightarrow NP \quad VP$

$NP \rightarrow N$

$VP \rightarrow V \quad NP$



"John loves fell"



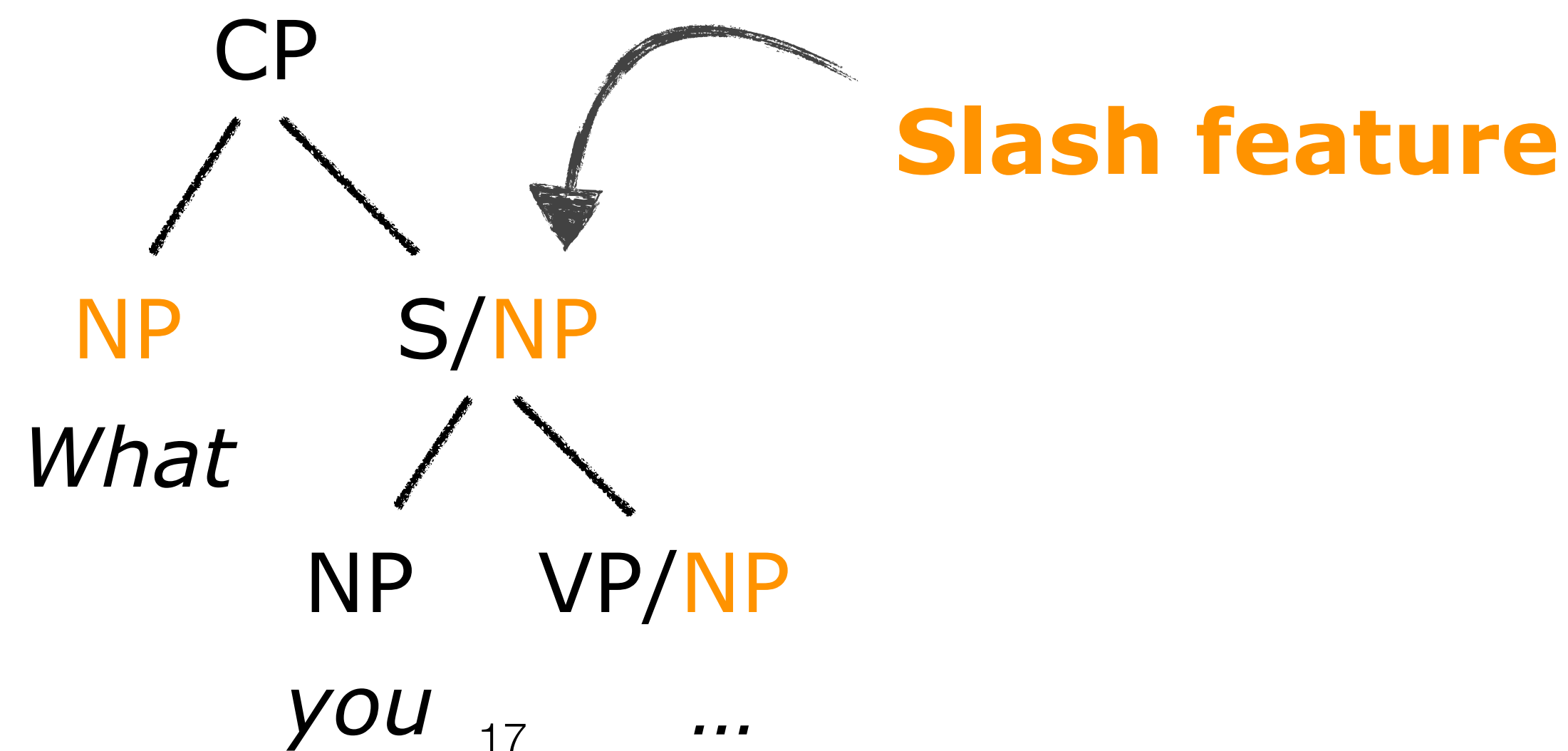
Ungrammatical sentences cannot be generated under traditional grammatical and parsing accounts

Long-distance dependencies

Long-distance dependencies: Dependencies between two elements that are in a syntactic and semantic relation but not adjacent in the sentence (the original position of the wh-element is indicated by an underscore called a *gap*)

What_i did you think that John read ___i?

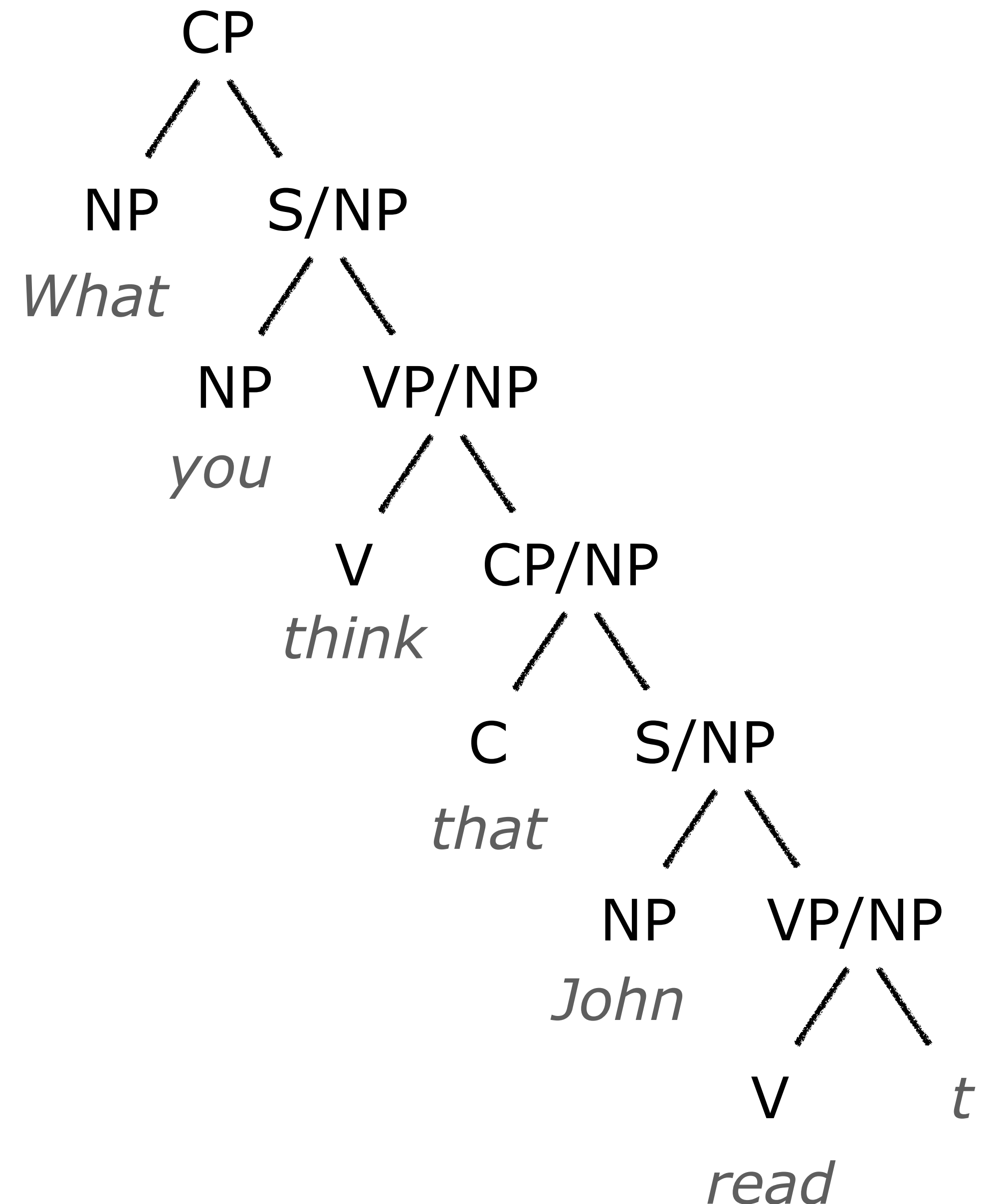
One way to encode long-distance dependencies in phrase structure rules is through the slash feature notation (Gadzar, 1981)



The traditional view (in a nutshell)

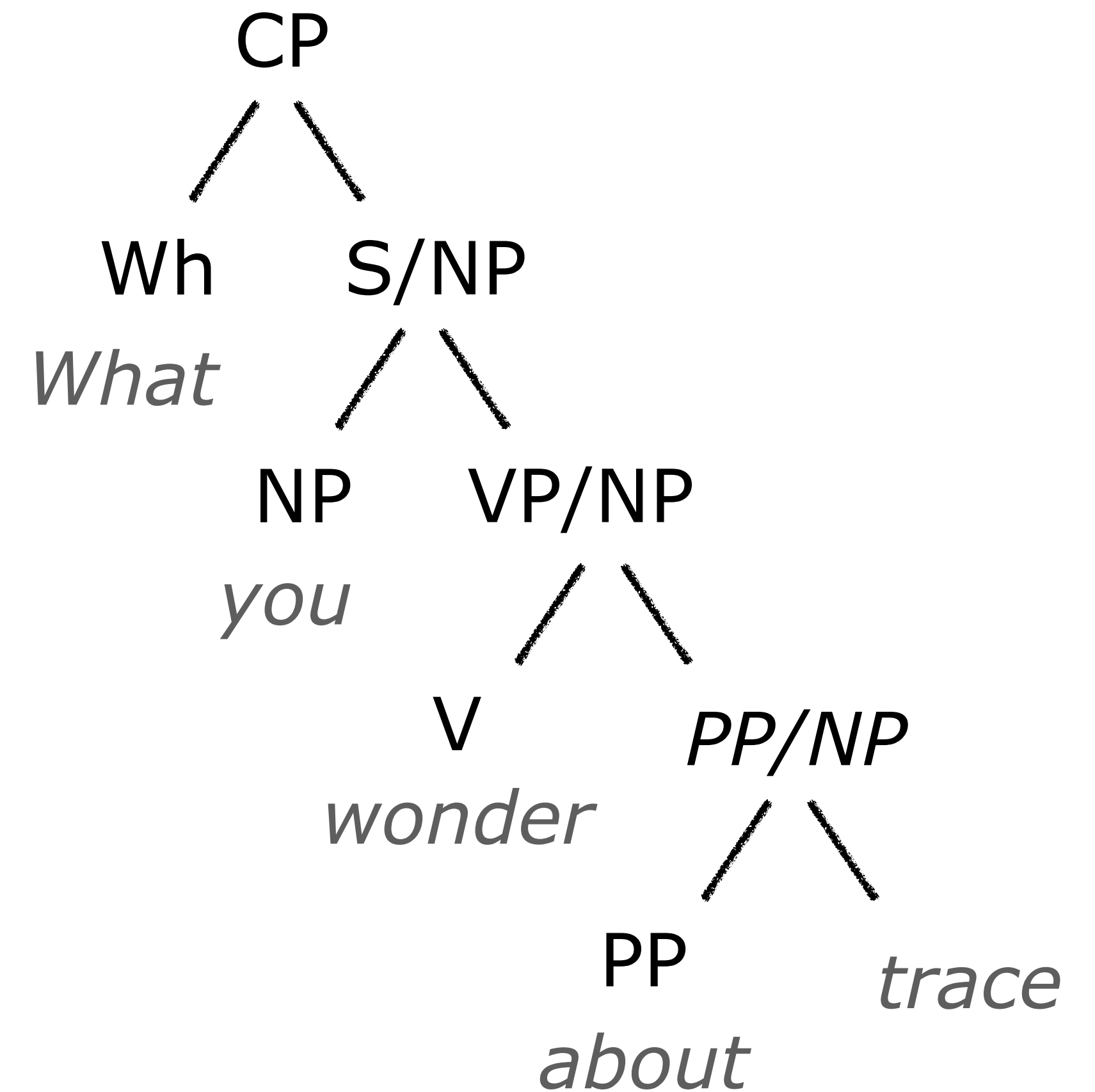
What did you think that John read?

- | | | | |
|----|----------------------|-----------------------|----------------------|
| 1. | CP | → NP[<i>what</i>] | S/NP[<i>what</i>] |
| 2. | S/NP[<i>what</i>] | → NP | VP/NP[<i>what</i>] |
| 3. | VP/NP[<i>what</i>] | → V[<i>think</i>] | CP/NP[<i>what</i>] |
| 4. | CP/NP[<i>what</i>] | → C[<i>that</i>] | S/NP[<i>what</i>] |
| 5. | VP/NP[<i>what</i>] | → V | trace |
| 6. | VP/NP[<i>what</i>] | → V[<i>wonder</i>] | trace |
| 7. | VP/NP[<i>what</i>] | → V[<i>wonder</i>] | PP/NP |
| 8. | PP/NP[<i>what</i>] | → PP | trace |
| 9. | CP | → C[<i>whether</i>] | S |



The traditional view (in a nutshell)

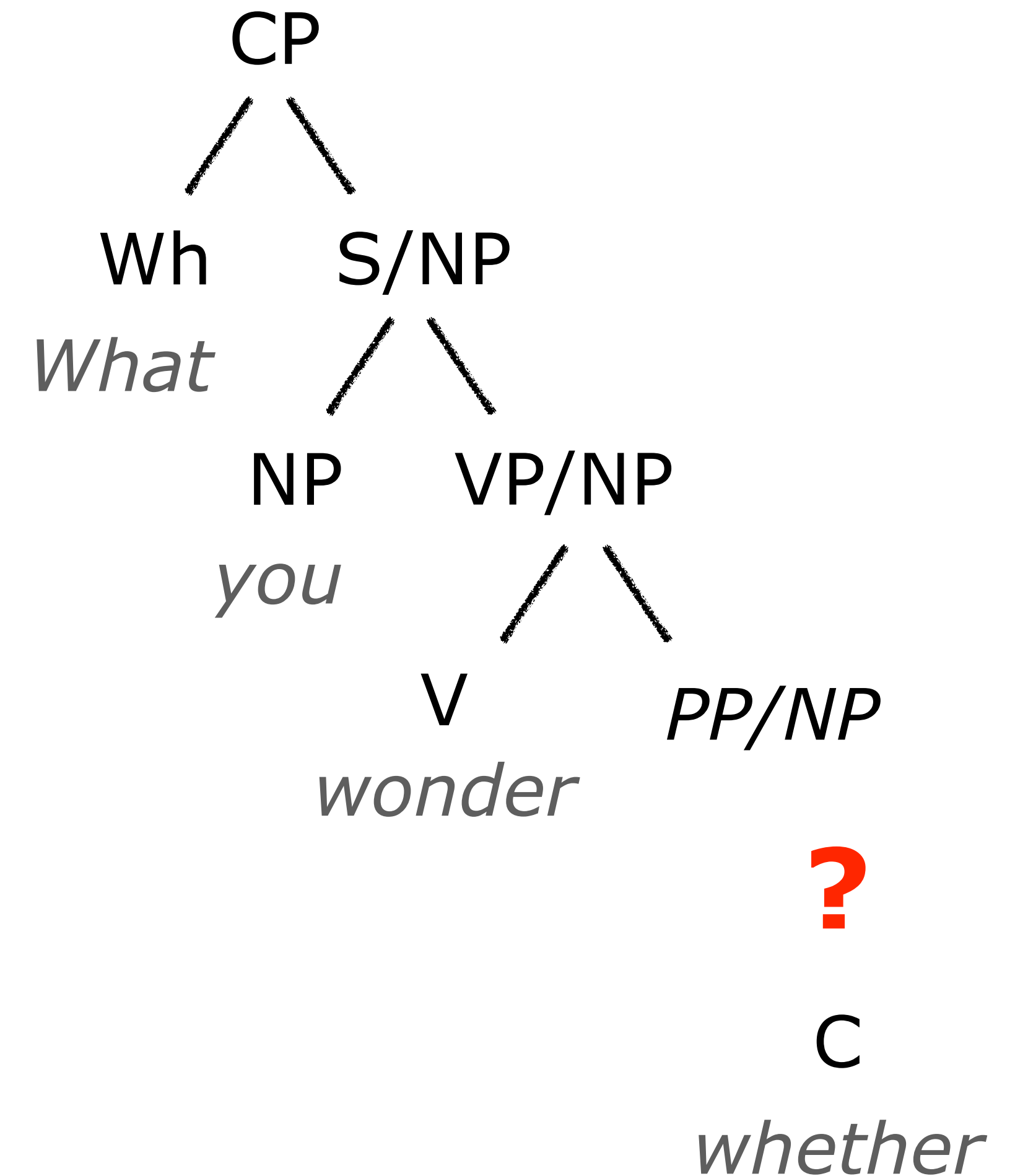
What did you wonder about?



- | | | | | |
|----|----------------------|---|---------------------|----------------------|
| 1. | CP | → | NP[<i>what</i>] | S/NP[<i>what</i>] |
| 2. | S/NP[<i>what</i>] | → | NP | VP/NP[<i>what</i>] |
| 3. | VP/NP[<i>what</i>] | → | V[<i>think</i>] | CP/NP[<i>what</i>] |
| 4. | CP/NP[<i>what</i>] | → | C[<i>that</i>] | S/NP[<i>what</i>] |
| 5. | VP/NP[<i>what</i>] | → | V | trace |
| 6. | VP/NP[<i>what</i>] | → | V[<i>wonder</i>] | trace |
| 7. | VP/NP[<i>what</i>] | → | V[<i>wonder</i>] | PP/NP |
| 8. | PP/NP[<i>what</i>] | → | PP | trace |
| 9. | CP | → | C[<i>whether</i>] | S |

The traditional view (in a nutshell)

**What did you wonder whether John read?*



- | | | | |
|----|----------------------|-----------------------|----------------------|
| 1. | CP | → NP[<i>what</i>] | S/NP[<i>what</i>] |
| 2. | S/NP[<i>what</i>] | → NP | VP/NP[<i>what</i>] |
| 3. | VP/NP[<i>what</i>] | → V[<i>think</i>] | CP/NP[<i>what</i>] |
| 4. | CP/NP[<i>what</i>] | → C[<i>that</i>] | S/NP[<i>what</i>] |
| 5. | VP/NP[<i>what</i>] | → V | trace |
| 6. | VP/NP[<i>what</i>] | → V[<i>wonder</i>] | trace |
| 7. | VP/NP[<i>what</i>] | → V[<i>wonder</i>] | PP/NP |
| 8. | PP/NP[<i>what</i>] | → PP | trace |
| 9. | CP | → C[<i>whether</i>] | S |

The derivation crashes as soon as the sentence turns out to be ungrammatical (i.e. when no rule can be applied to integrate the linguistic input in the tree)

2. Gradient in islands

Syntactic Islands

Islands = Encapsulated syntactic domains that prohibit the establishment of a long-distance dependency inside of them (the island domain is in red; the asterisk indicates ungrammaticality)

(Ross, 1967)

WHETHER ISLAND

*What_i did you wonder whether John read ___i?

COMPLEX NP ISLAND

*What_i did you hear the news that John read ___i?

ADJUNCT ISLAND

*What_i did you smile because John read ___i?

Gradient effects in islands' acceptability

Intriguingly, the acceptability of some of these islands has been shown to *improve* when the wh-element is made them lexically specific (e.g. *which book*)

(e.g. Sprouse et al. 2012, 2016; Sprouse & Messick 2015; Villata et al. 2016; Atkinson et al. 2016)

WHETHER ISLAND

?Which book_i did you wonder **whether John read ____i?**

COMPLEX NP ISLAND

?Which book_i did you hear the news **that John read ____i?**

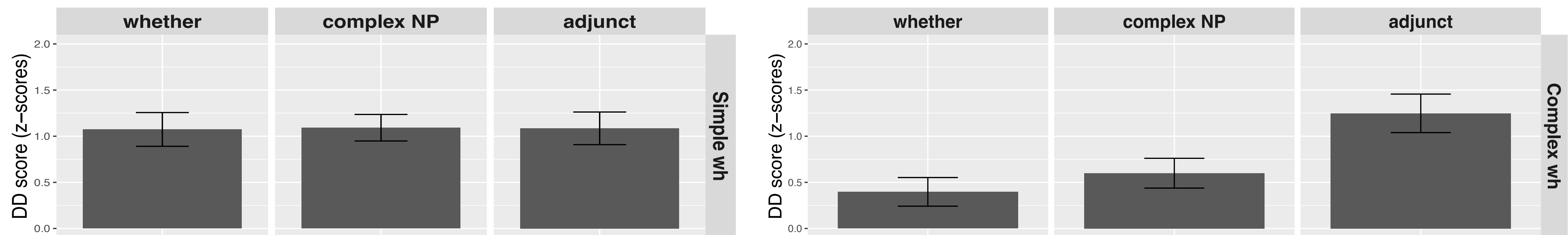
ADJUNCT ISLAND

*Which book_i did you smile **because John read ____i?**

Gradient islands' effects in acceptability

This is illustrated in the plots below (data from Sprouse & Messick 2015): the island effect is significantly reduced (but not eliminated) for Whether and Complex NP islands with complex wh, but not for Adjunct islands

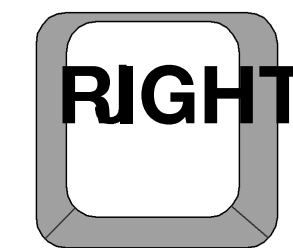
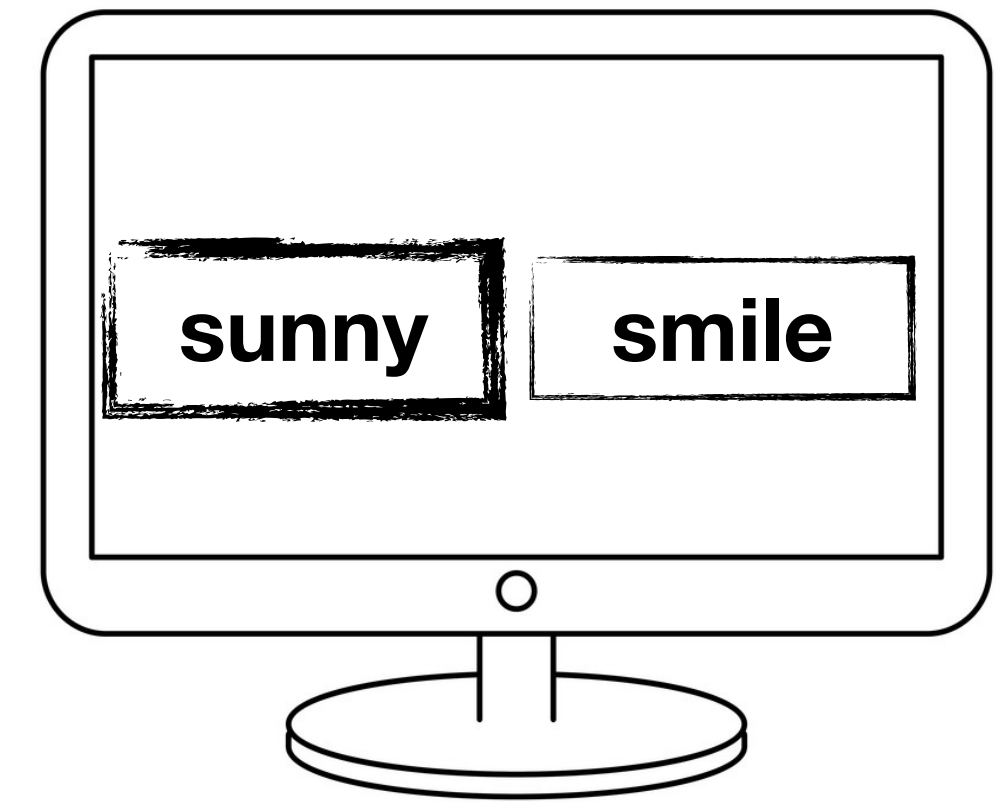
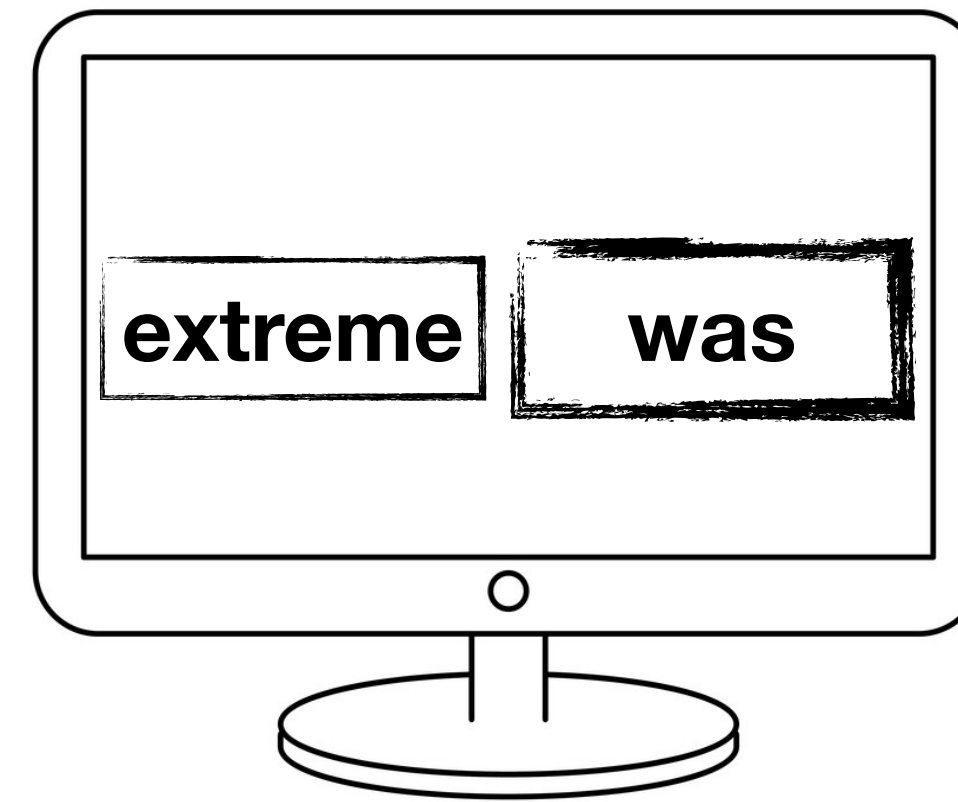
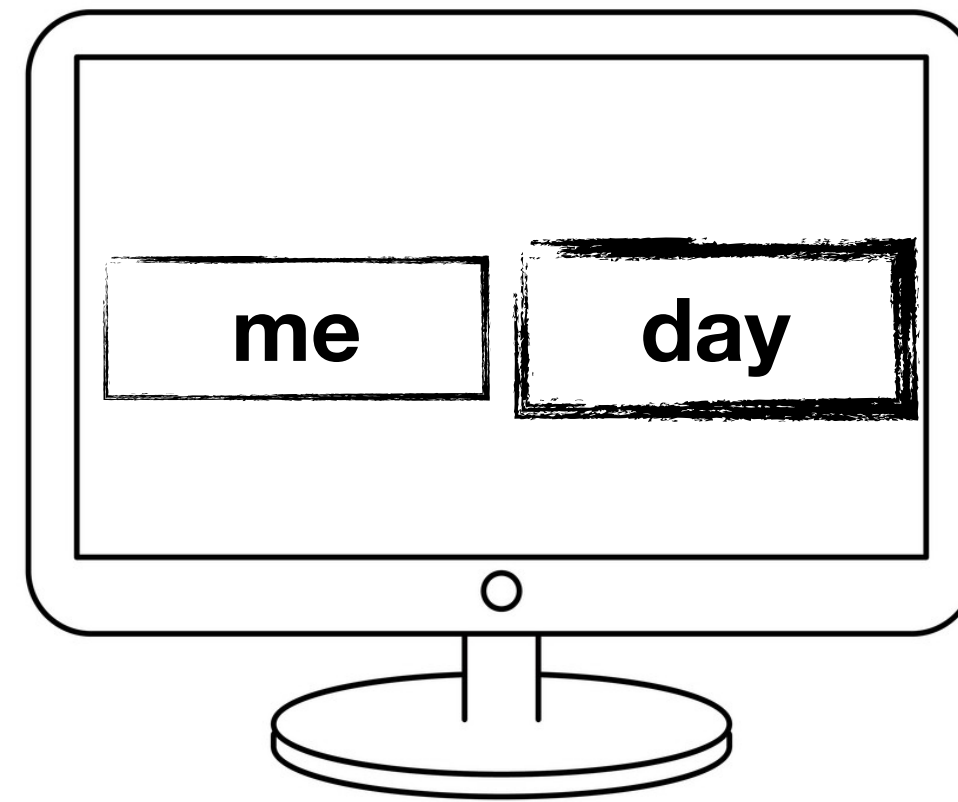
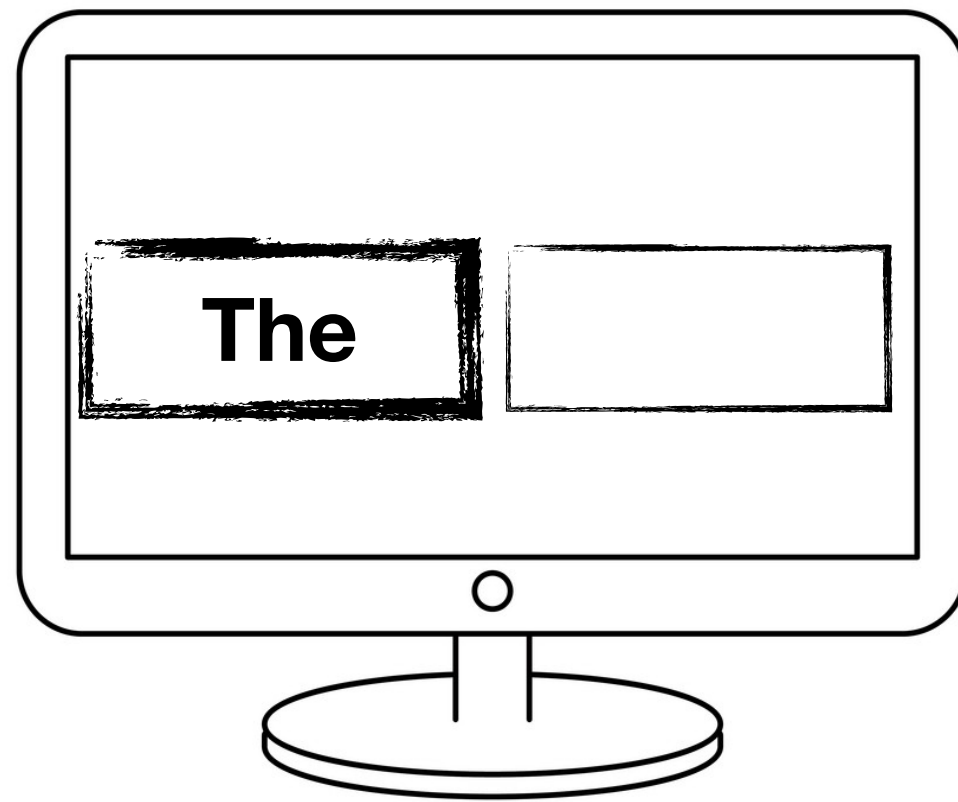
(DD scores are a measure of the *strength of the island effect*, so the higher the bar, the stronger the island effect)



Beyond gradient acceptability

- The data reported above provide evidence for gradient acceptability patterns in island-violating sentences
- Interestingly, we also have evidence showing that the increased islands' acceptability is the result of comprehenders' ability to interpret the island by forming a dependency inside of it

The Maze Task



Task rationale

- Participants read the island preamble word-by-word up to a critical point (the verb inside the island, e.g. *solved*):

What did you wonder whether the candidate solved ...

preamble

... **before** the interview in Paris?

... **the** problem before the interview?

- At this point they were asked to decide how to continue the sentence:
 - select a preposition (e.g. "before"), compatible with establishing a dependency inside of the island
 - select a determiner ("the"), compatible with not establishing a dependency inside the island

Task rationale

- Notice that both options are ungrammatical:
 - *violate the island constraint*, but interpreting the wh-word as the object of the verb thus assigning it a thematic role (theme), which renders the sentence interpretable:

What did you wonder whether the candidate **solved** __ before the interview?

- *vacuous quantification* (i.e. the dependency between the wh-word and the verb is not established, and thus the wh-word lacks a thematic role)

What did you wonder whether the candidate **solved** *the problem*?

Materials

- We tested Whether, Complex NP, and Adjunct islands with both simple and complex wh

WHETHER ISLAND

What did you wonder **whether the candidate solved...**

COMPLEX NP ISLAND

What did you hear the news **that the candidate solved...** ... before the interview?
... the problem?

ADJUNCT ISLAND

What did you smile **because the candidate solved...**

Results

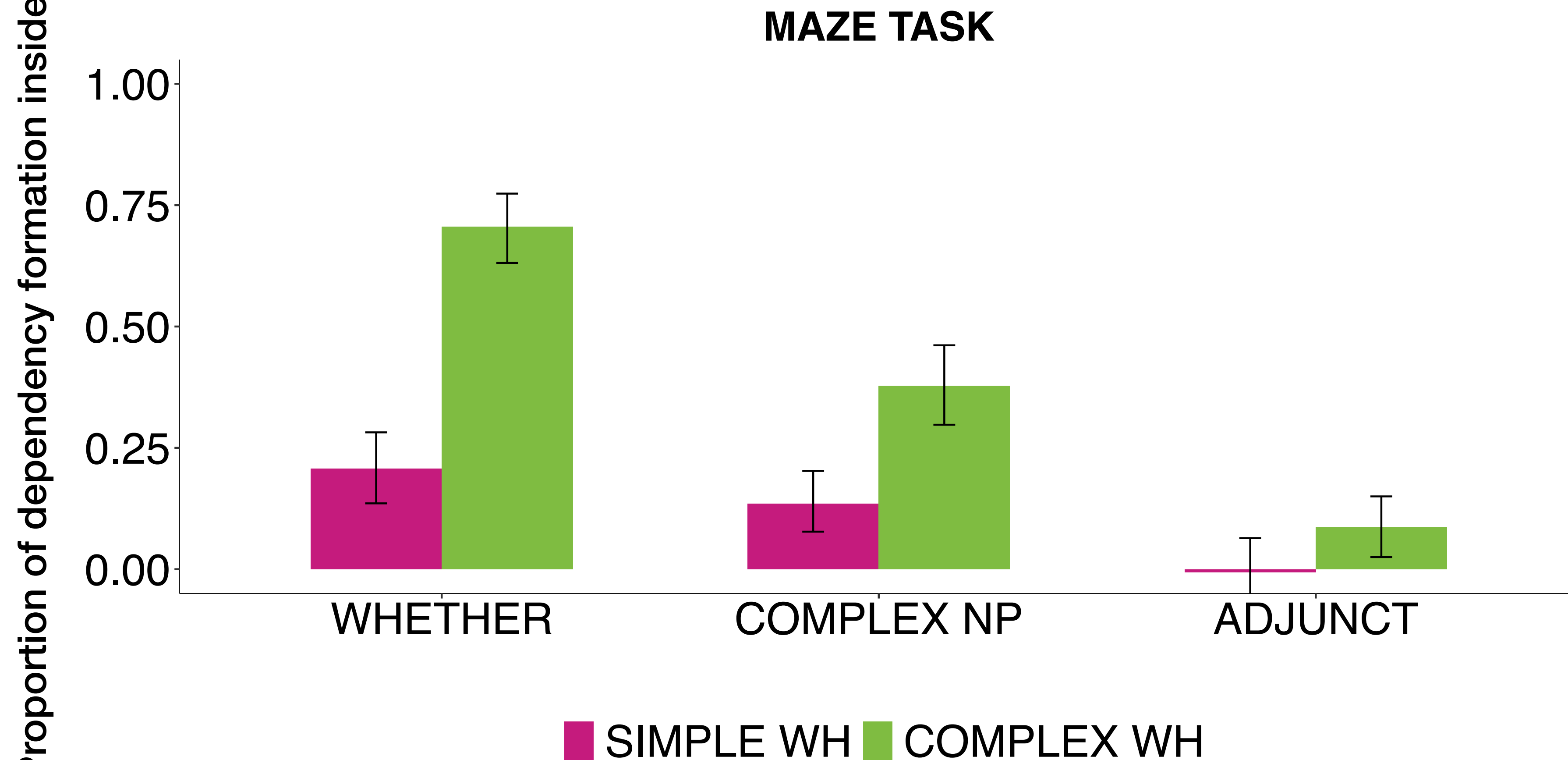


Fig. 1 Proportions of preposition selection (compatible with establishing a dependency inside an island) corrected for the control condition establishing the floor of gap selection in sentences in which gap selection is not expected (no long-distance dependency)

Interim Discussion

- These results suggest that there is a **strong correlation between gradient acceptability and the formation of island violating dependencies**: islands that receive higher acceptability rates are those for which participants are more willing to form a dependency between the wh-word and the verb inside the island
- What are the **mechanisms** that generate gradience: are they **grammar-internal** or **grammar-external**?
- The findings reported here from the Maze Task do not allow us to provide an answer to this question
 1. The syntactic derivation might fail as soon as the violation is detected (in line with standard assumptions) and some extra-grammatical mechanism might then kick in to cobble the sentence together
 2. The syntactic tree is generated in a more flexible rule-based system that allows the generation of semi-grammatical sentences

Coercion

- The mechanism we will invoke to account for ungrammatical sentence processing is **coercion**
- Coercion intervenes when there is no way to generate a full-formed syntactic tree by following the rules of the grammar
- In these cases, what the system does is to force one or more sentential elements to play a role that does not fully fit them
- There are two kinds of coercion:
 1. **Interpretable coercion** = the system forms a thematically coherent tree (i.e. all thematic roles are assigned), despite the feature mismatch on some nodes, and the sentence results interpretable
 2. **Uninterpretable coercion** = the system forms a tree but not all element get their thematic roles, and the sentence thus results uninterpretable
- Interpretable coercion is triggered by the existence of a fully grammatical sentence that is analogue (i.e. syntactically and semantically similar) to the deviant one
- In the absence of such analogy, uninterpretable coercion applies

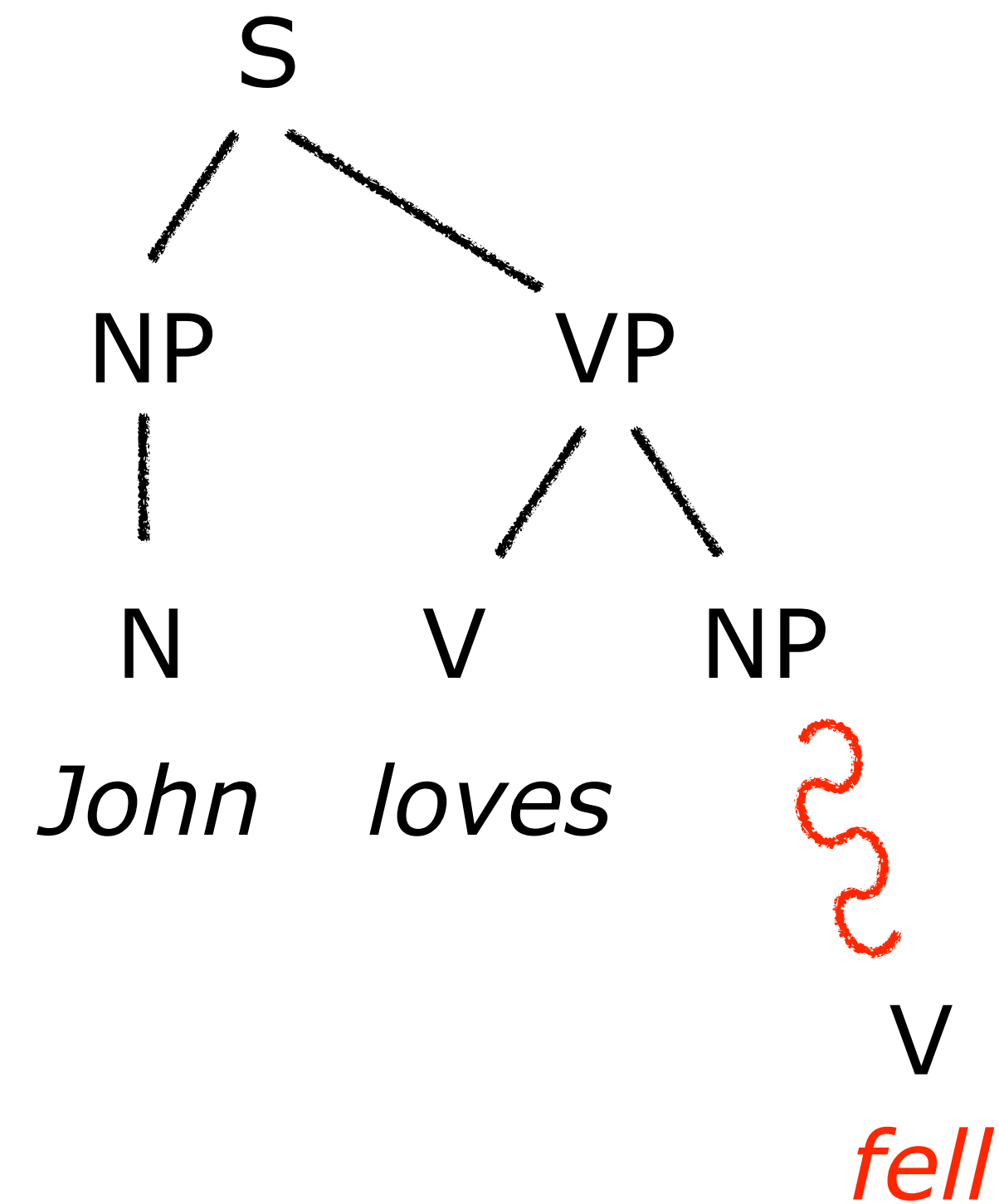
Uninterpretable coercion

"John loves fell"

$S \rightarrow NP \quad VP$

$NP \rightarrow N$

$VP \rightarrow V \quad NP$



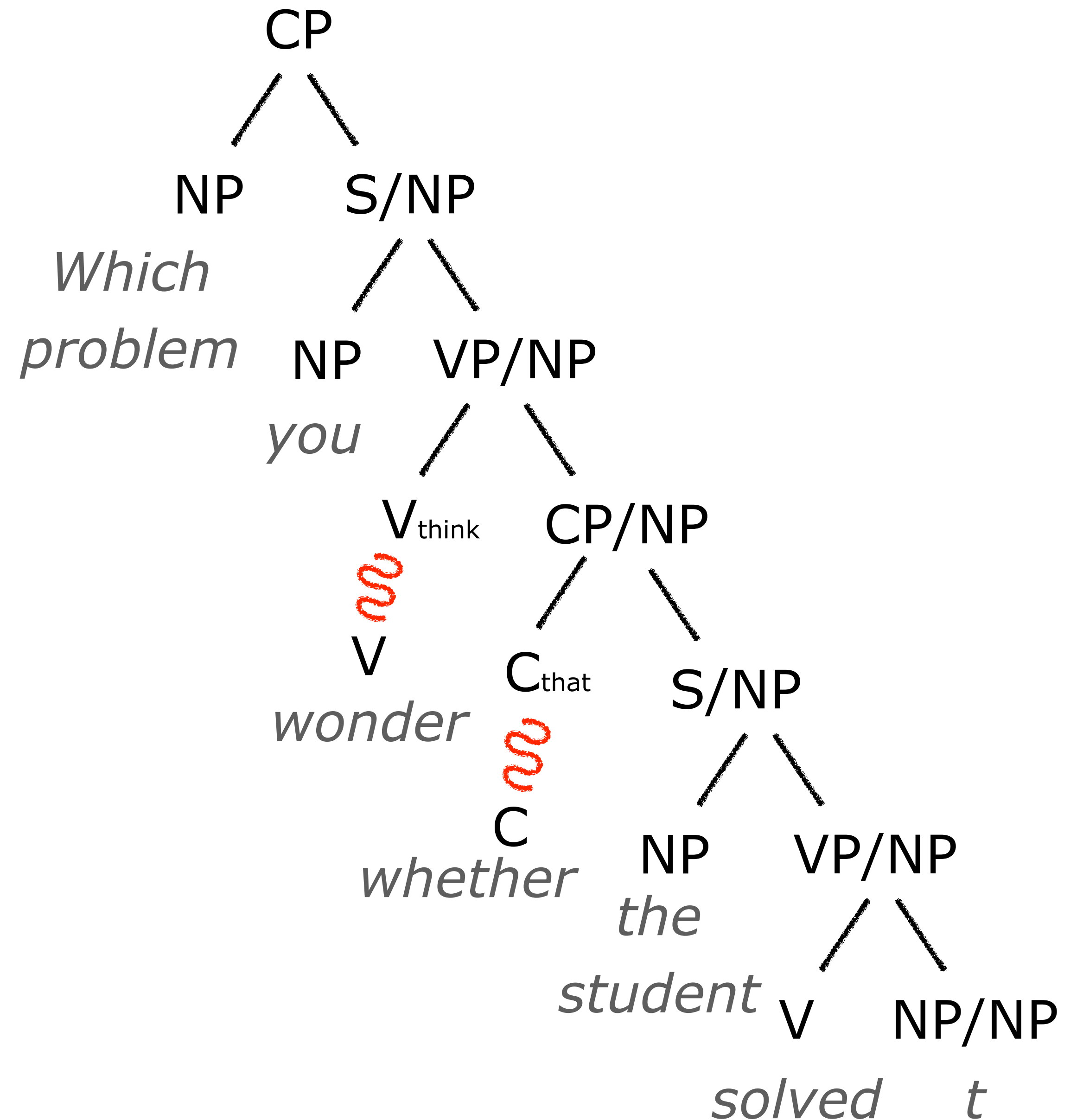
Coercion in Whether islands

- Whether-islands resemble a fully grammatical long-distance dependency, namely extraction from a declarative:

 Which problem did you wonder whether the student solved?
Which problem did you think that the student solved?


1. “Wonder” and “Think” are mental-process verbs that subcategorise for a propositional complement
 2. Both “wonder” and “think” are followed by a complementiser
 3. Both “wonder” and “think” can refer to the subject’s degree of certainty about the embedded proposition: “wonder” indicates a high degree of uncertainty about the truth/falsehood of the complement, while “think” is more biased toward the truth of the complement (even though this bias can be reduced if “think” is focused, as in “I THINK that the student solved the problem, but I can’t swear to it”)
- Because of the analogue syntactic structures in which “wonder”-“think” and “whether”-“that” appear, and their close semantics, interpretable coercion occurs (most of the time)

- When (interpretable) coercion occurs, “wonder” is coerced into “think” (a slash-propagator verb) and “whether” into “that” (a slash-propagator complementizer), thus allowing the propagation of the slash feature down the tree



Complex vs. Simple wh

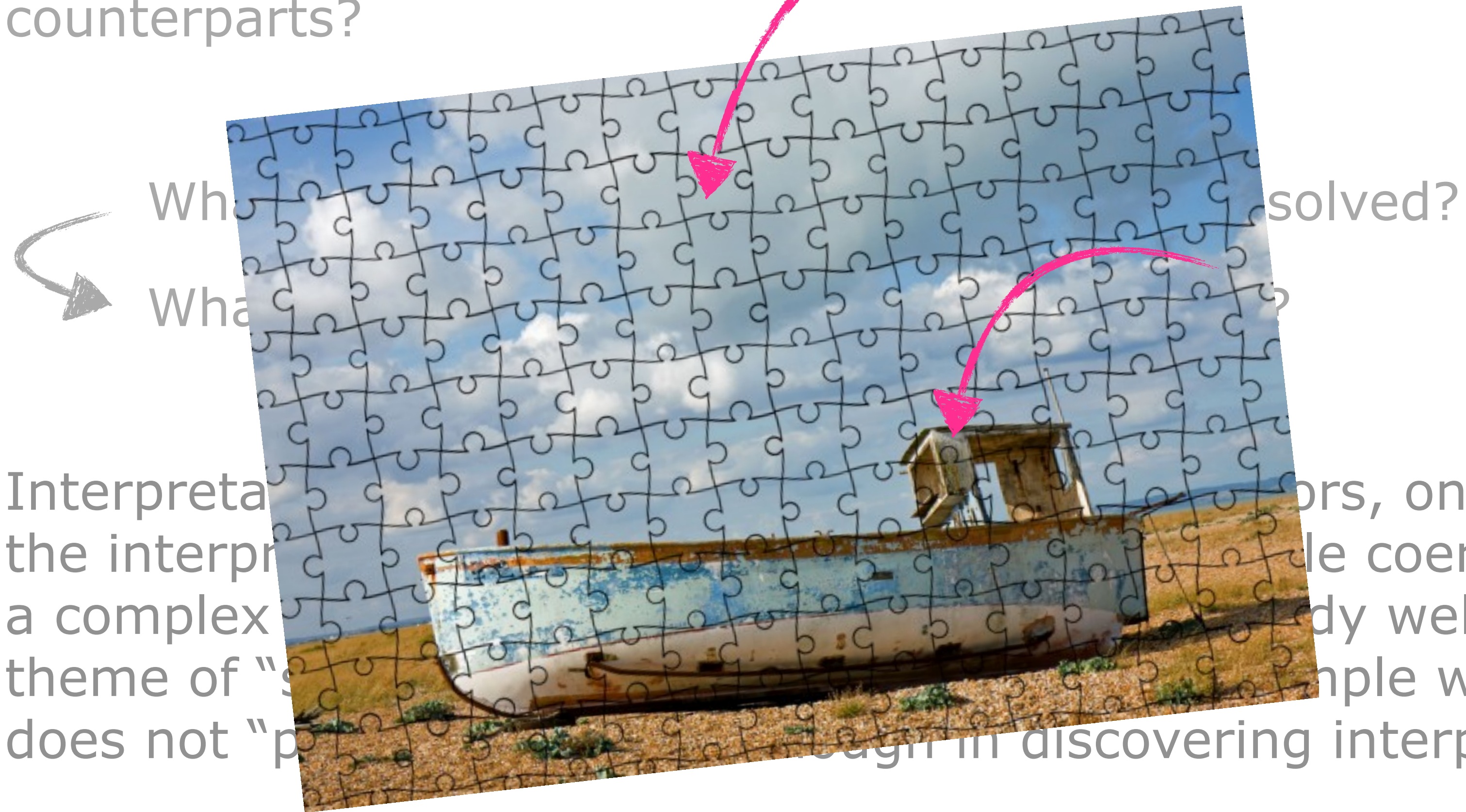
- Why Whether-islands with complex wh are more coercible than their simple counterparts?

 What did you **wonder whether** the student solved?
What did you **think that** the student solved?

- Interpretable coercion is favoured by several factors, one of which is **semantics**: the interpretative pressure underlying interpretable coercion is stronger for a complex wh like “which problem”, which is already well suited to fill the role of the theme of “solved”, than for a semantically light simple wh like “what”. This lightness does not “push” the system enough in discovering interpretable coercion

Complex vs. Simple wh


- Why Whether-islands with complex wh are more coercible than their simple counterparts?



- Interpretation of the interpretation of a complex theme of "s" does not "p" ... ors, one of which is semantics: ... le coercion is stronger for ... dy well suited to fill the role of the ... mple wh like "what". This lightness ... begin in discovering interpretable coercion

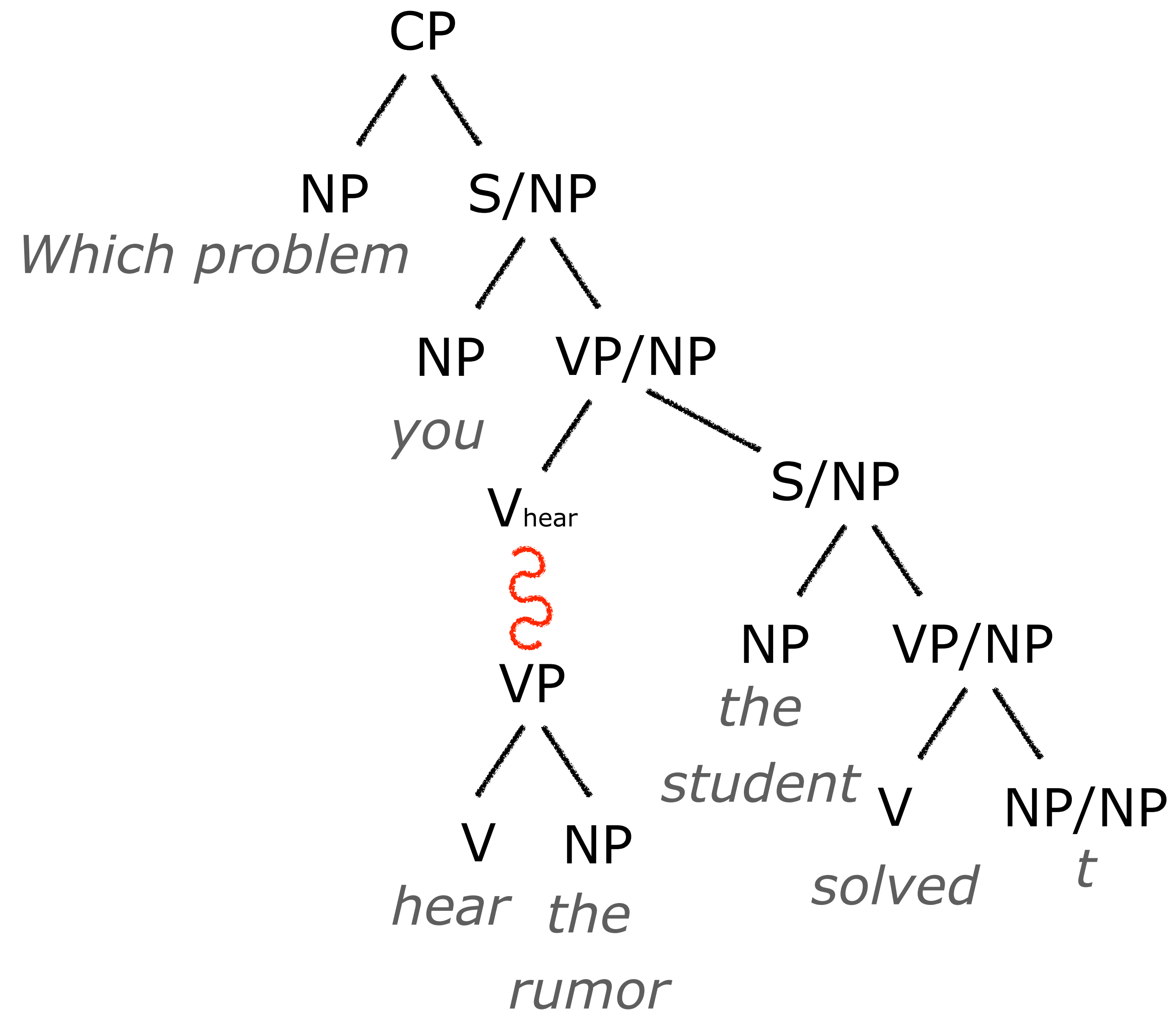
Coercion in Complex NP islands

- Complex NP islands also have an analogue fully grammatical long-distance dependency:


 Which problem did you **hear the news** that the student solved?
Which problem did you **hear** that the student solved?

- “Hear the news” and “Hear” denote the same *hearing event*
- “Hear the news” and “Hear” appear in the same syntactic context
- interpretable coercion occurs, and even more so with complex wh than with simple one for the same reason discussed above for Whether-islands

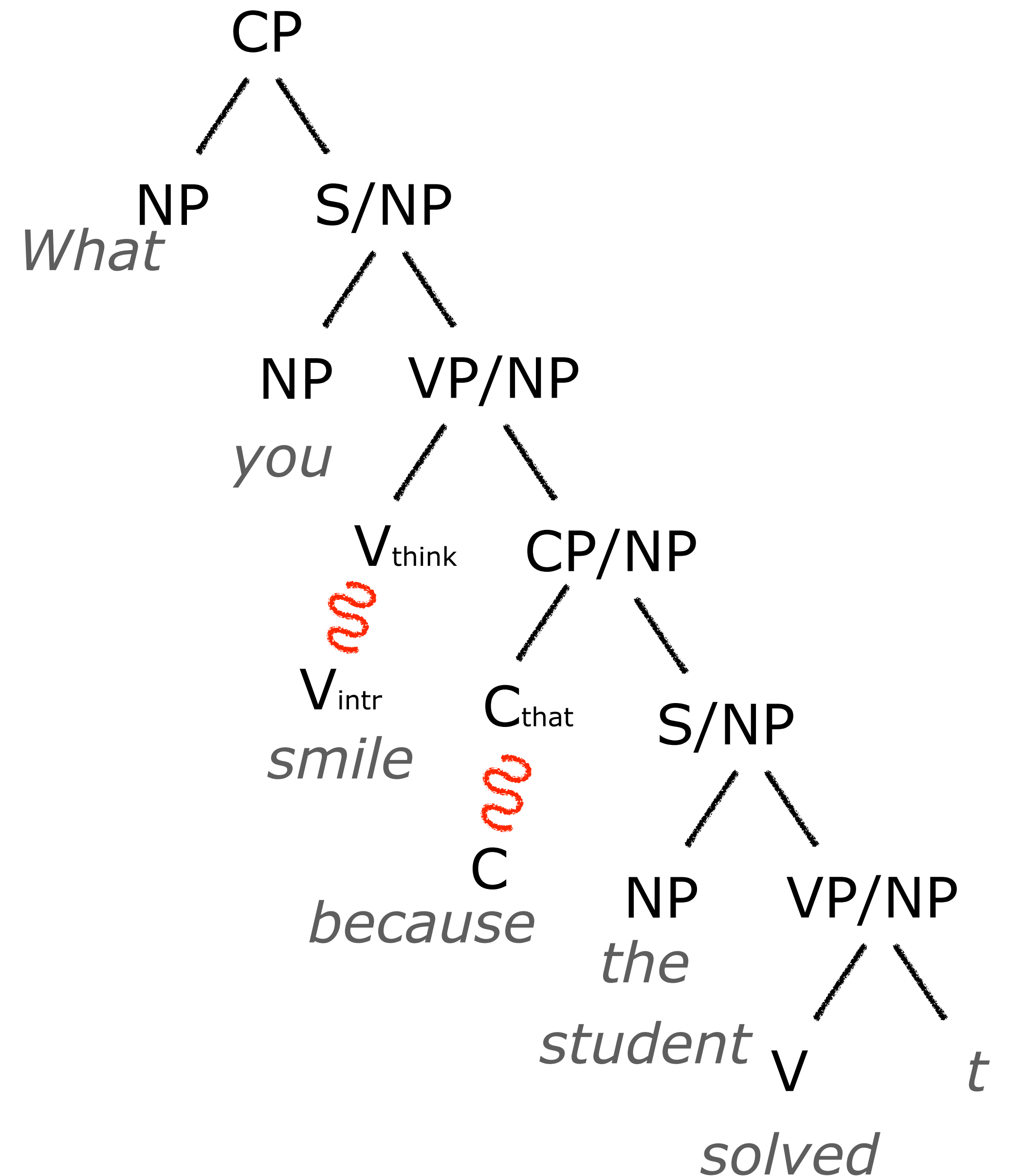
- When (interpretable) coercion occurs, “hear the rumor” is coerced into “hear”, which allows the propagation of the slash feature down the tree



Coercion in Adjunct islands

- Adjunct islands, unlike Whether and Complex NP islands, turn out not to have an analogue structure that is fully grammatical, and for this reason they are not (interpretably) coercible
- The extraction from a declarative is not a valid analogy in this case:
 -  Which problem did you **smile because** the student solved?
 - Which problem did you **think that** the student solved?
- “Smile” is unergative, while “think” is a sentence complement verb
- “That” introduces a complement, while “because” introduces an adjunct

- For Adjunct islands, (interpretable) coercion would imply coercing an intransitive verb (e.g. smile) into a verb taking a sentential complement (e.g. think). This coercion however puts a lot of strain into the system because of the different argument structure of the two verbs. As a result, this type of coercion is very unlikely to occur



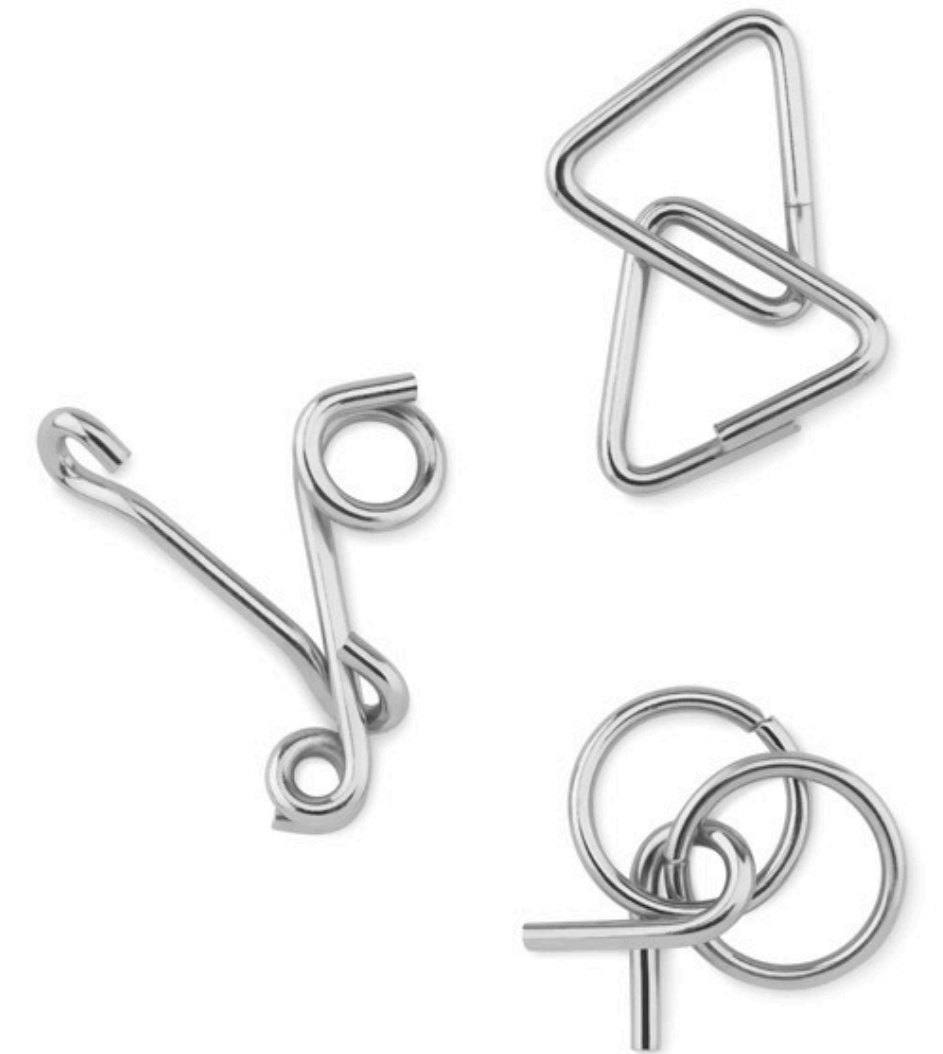
Flexible rule-following system

- Gradience can be generated through a rule-based system under the assumption that *syntax is coercible* - when no rule can accomodate the input, then the system can be made more *flexible*
- *Flexibility* is an option when the deviant structure resembles in some relevant respect to a fully grammatical one (the system *coerces* sentential elements to play a role that doesn't fully fit them but that, at the same time, it's not too far away from their actual role)



Similarity

- I claimed that *similarity/analogy* between sentences play a key role in processing: comprehenders' ability to parse an island fundamentally relies on the existence of a fully grammatical sentence that resembles the island structure
- One might wonder, however, why similarity should be relevant in the first place in accounting for language processing phenomena
- Language processing brings some commonalities with *brain teasers*: when it comes to parse particularly cumbersome sentences, the mind must find "its path" through them (this is similar to what we have to do when we have to solve a metal brain teaser: there is a precise sequence of moves that we have to do in order to get it right)
- Once the mind finds this sequence, this can be successfully applied to solve similar "brain teasers"



Linguistic brain teasers

Successfully parsing cumbersome sentences, even grammatical ones, might be challenging

- **Center embeddings:**

The rat that the cat that the dog loved chased died

The rat that the cat that the dog loved chased died

The rat (who was chased by the cat who, in its turn, was loved by the dog) died

- **Garden-path sentence:**

When the men hunt the birds that cheetahs eat typically scatter

When the men hunt, the birds that cheetahs eat typically scatter

The horse raced past the barn fell

The horse, (that was) raced past the barn, fell

Similarity in solving brain teasers

○ Islands

If you know how to “solve” this ...

What do you think that the student read ____ ?

... you might know how to “solve” this

*What do you wonder whether the student read ____ ?

○ Center embeddings

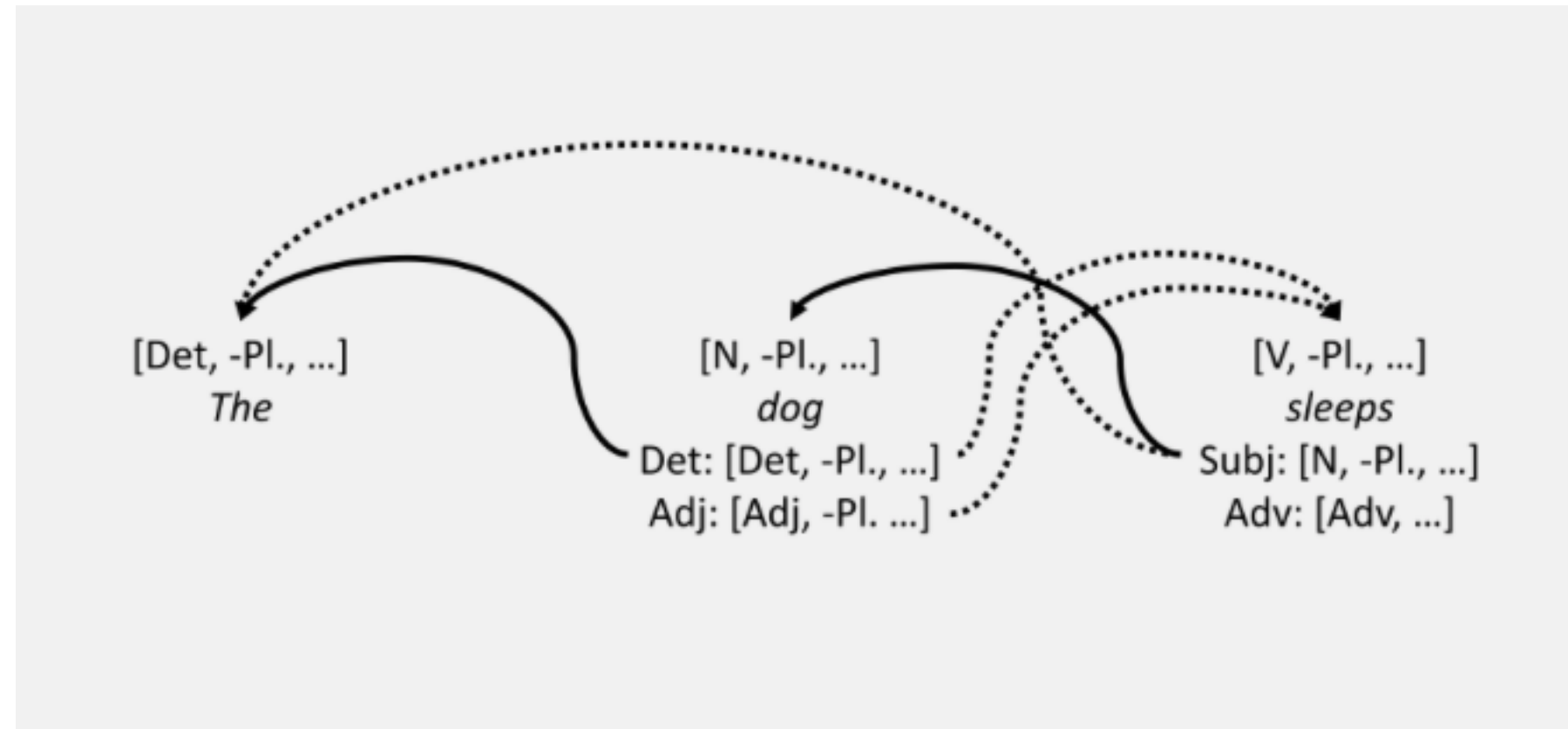
If you know how to “solve” this ...

The rat that the cat chased died

... you might know how to “solve” this

The rat that the cat that the dog loved chased died

SOSP



Modeling Ungrammaticality: A Self-Organizing Model of Islands

Sandra Villata (sandra.villata@uconn.edu)

Department of Linguistics, 365 Fairfield Way
Storrs, CT 06269-1145 USA

Jon Sprouse (jon.sprouse@uconn.edu)

Department of Linguistics, 365 Fairfield Way
Storrs, CT 06269-1145 USA

Whitney Tabor (whitney.tabor@uconn.edu)

Department of Psychology and Haskins Laboratories, 406 Babbidge Road
Storrs, CT 06269-1020 USA

Toward a Theory of Timing Effects in Self-Organized Sentence Processing

Garrett Smith (garrett.smith@uconn.edu)

Department of Psychological Sciences, 406 Babbidge Road, Unit 1020
Storrs, CT 06269 USA

Whitney Tabor (whitney.tabor@uconn.edu)

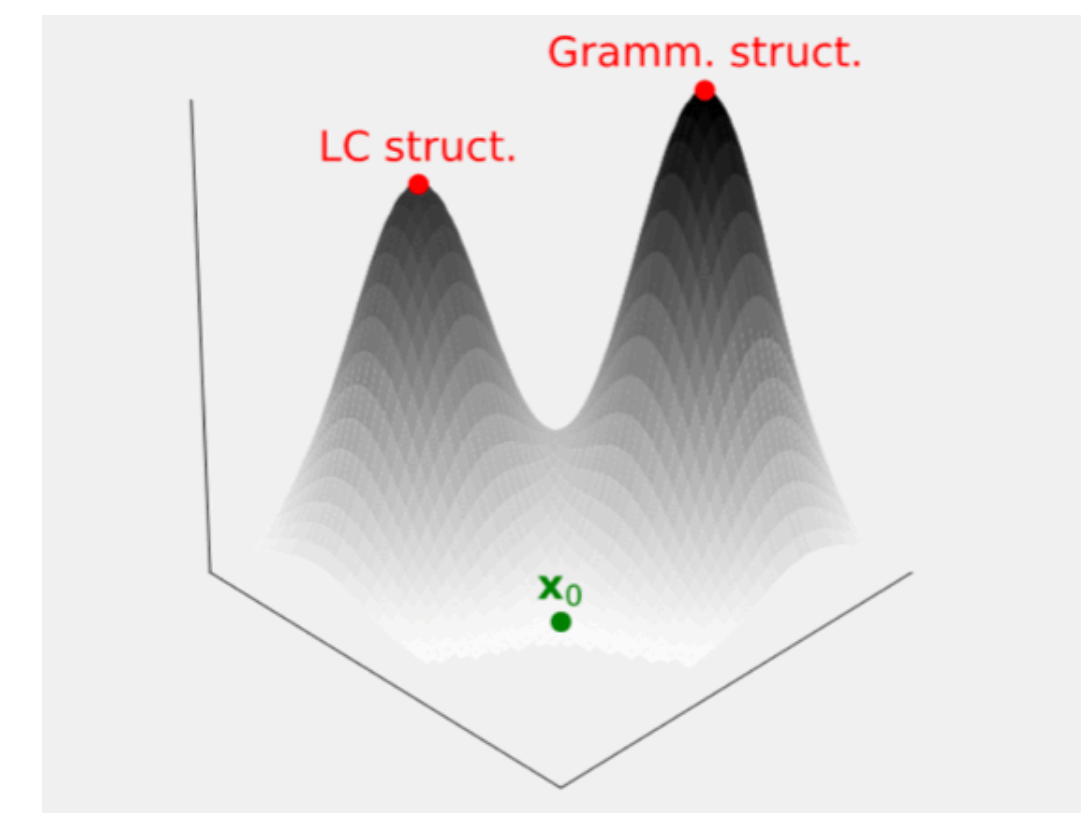
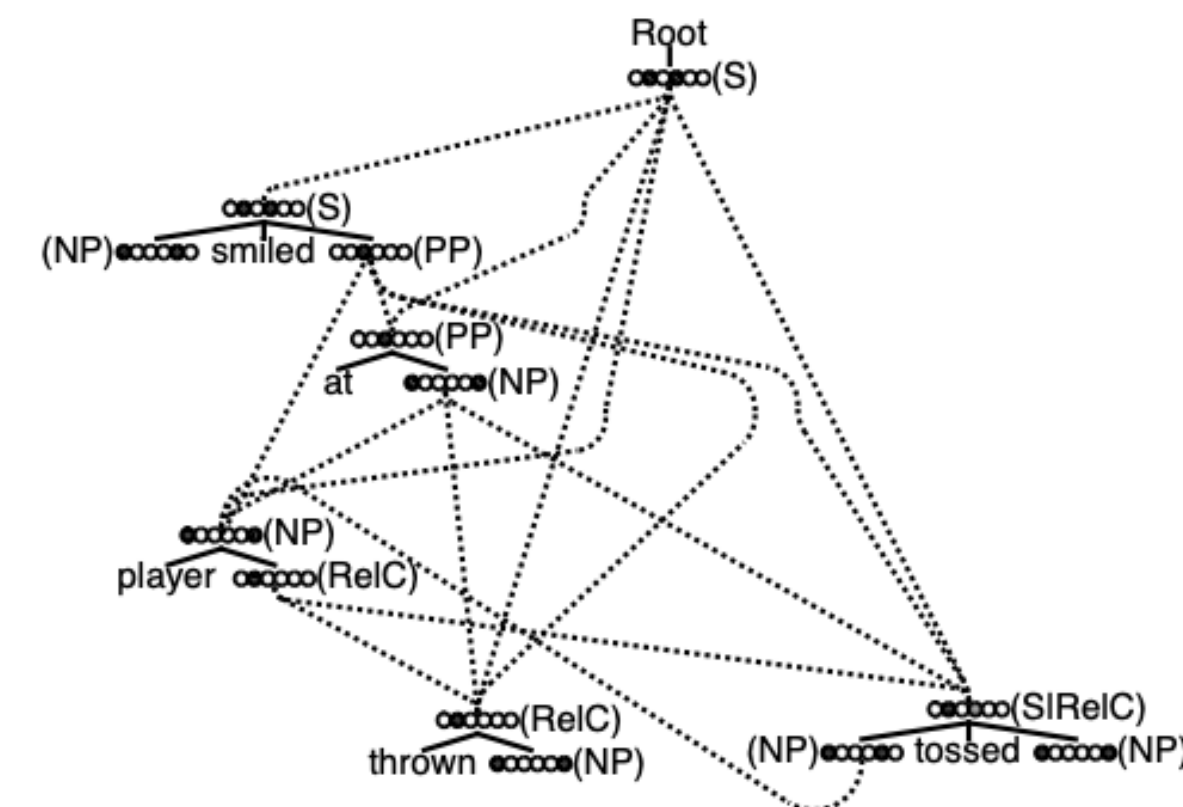
Department of Psychological Sciences, 406 Babbidge Road, Unit 1020
Storrs, CT 06269 USA

Journal of Experimental Psychology:
Learning, Memory, and Cognition
2004, Vol. 30, No. 2, 431–450

Copyright 2004 by the American Psychological Association, Inc.
0278-7393/04/\$12.00 DOI: 10.1037/0278-7393.30.2.431

Evidence for Self-Organized Sentence Processing: Digging-In Effects

Whitney Tabor and Sean Hutchins
University of Connecticut



Check also: Tabor, Villata, Sprouse, Talk virtually presented at at AMLaP 2020
<https://mediaup.uni-potsdam.de/Play/Chapter/259> (go to: 1h 39 min 07 sec)



UConn NBL

Thank you

This work has been supported by a postdoctoral grant to Sandra Villata from the Marica de Vincenzi Foundation as well as by research grants from the Institute for Brain and Cognitive Sciences (IBACS), University of Connecticut and the Neurobiology of Language Program (NBL), University of Connecticut.