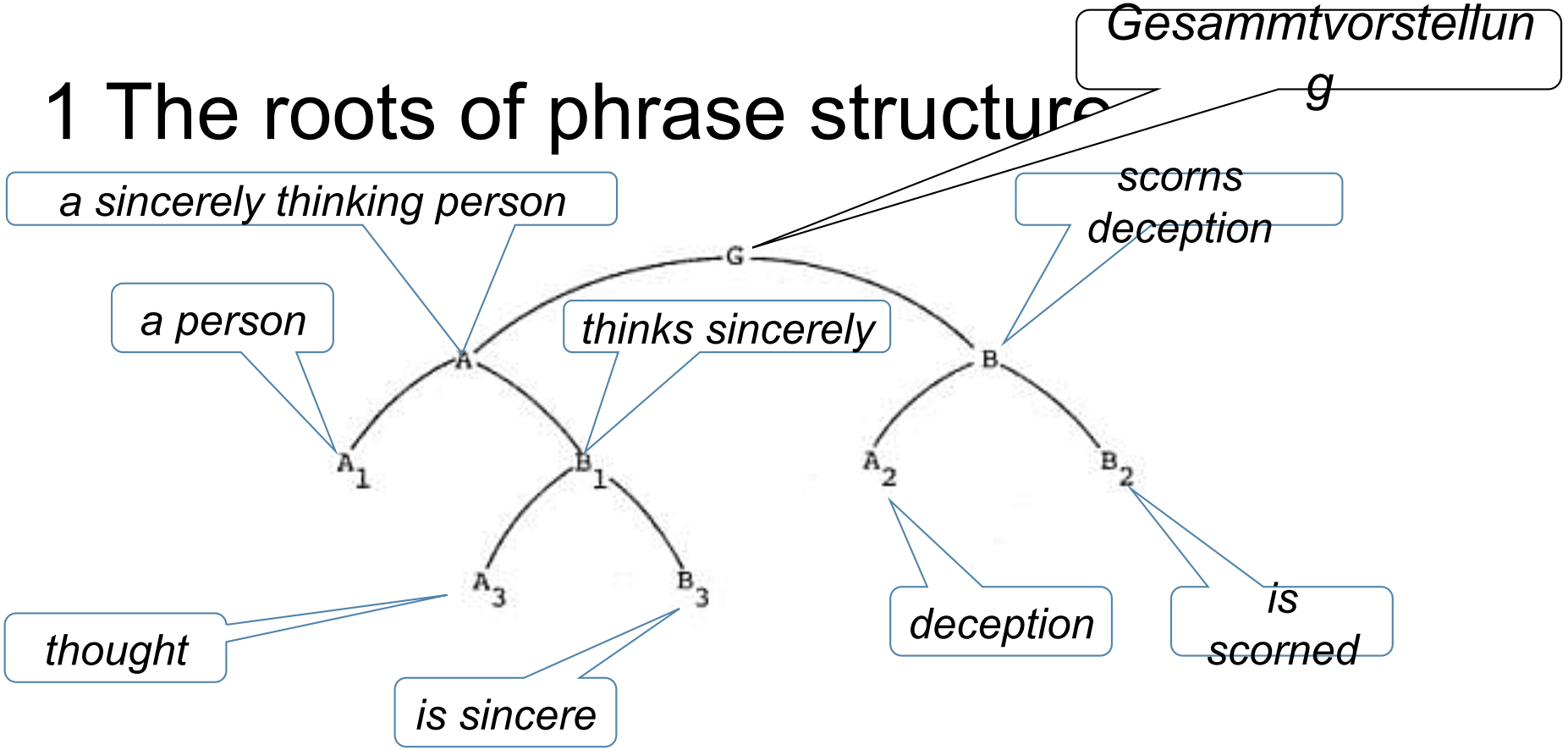
Syntax and cognition
or: Unifying Everything 2

Dick Hudson
Freie Universität Berlin, October 2015

Plan

1. How syntactic theory has been influenced by psychology
2. Why cognition?
3. Phrase structure and dependency structure
4. How to choose between PS and DS?
5. A challenge for DS, and a cognitive solution
6. Towards a new kind of DS
7. New-DS and PS: are they notational variants?
8. Conclusions

1 The roots of phrase structure



Ein redlich denkender Mensch verschmäht die Täuschung

A sincerely thinking person scorns deception

Who offered this analysis?

Wundt, Leipzig, 1900

So what?

- Bloomfield took this analysis from Wundt
  - who thought top-down analysis was psychologically real
  - but was looking at the meaning, not the words
- and turned it into his immediate-constituent analysis
- which Chomsky turned into his phrase-structure grammar
  - using ideas from mathematics which used brackets
  - which he turned into trees without crossing branches.
- So phrase structure is already based on assumptions about cognition.

2 Can we get away from cognition?

- We can certainly try.
  - E.g. Integrational linguistics
- But why would we want to try?
- After all, language is surely a kind of knowledge.
- So sooner or later our theories must meet theories about knowledge.
  - Jackendoff's "graceful integration" of language with the rest of cognition
- So we should at least try to build elementary ideas about other areas of cognition into our theories of language.

Cognitive linguistics

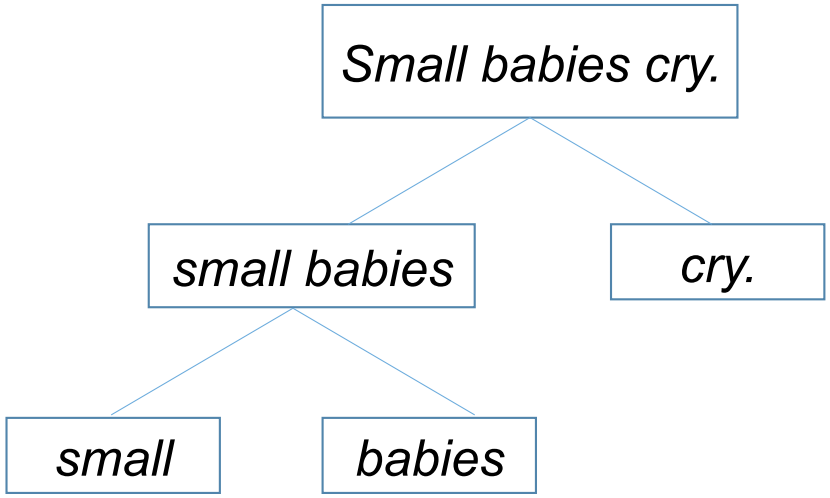
- This is one of the goals of cognitive linguistics
- Including
  - Cognitive Grammar (Langacker)
  - Construction Grammar (Goldberg, Croft)
  - Word Grammar
- My talk is about Word Grammar
  - developing since 1984
  - still changing
  - so some of this talk is new

3 Phrase structure or dependency structure?

- Two traditions in syntax:
- Phrase structure
  - Born in the USA (but inspired by Germany)
  - 1933 Bloomfield
  - 1957 Chomsky etc
- Dependency structure
  - Much older
  - Born in the Middle East and Europe
  - But taught in the USA in the 19th century (Reed and Kellogg diagrams)
  - 1959 Tesnière

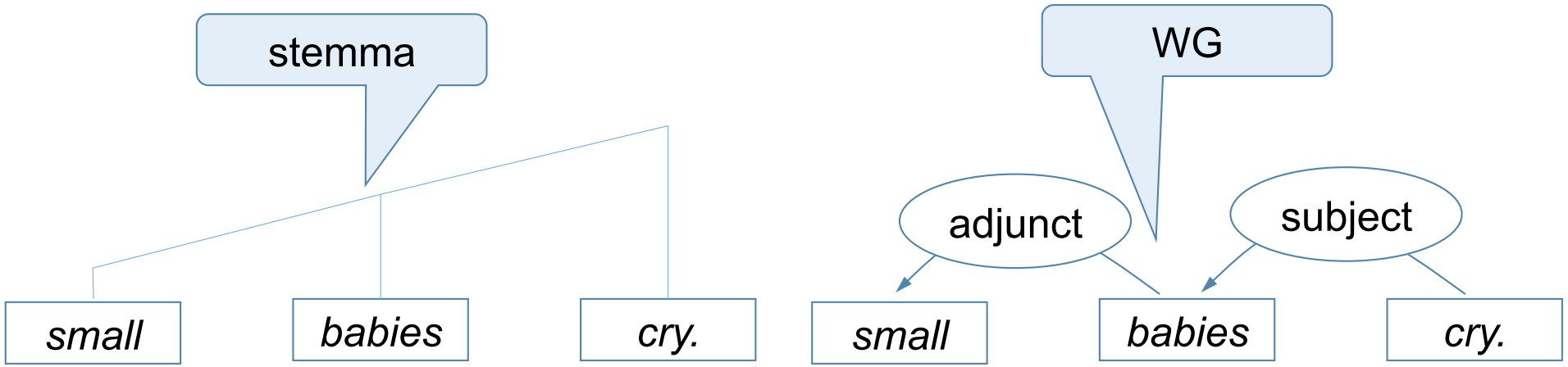
Phrase structure

- The only relation recognised is the part-whole relation



Dependency structure

- The only relation recognised is the dependency between two words.



How a DS grammar works

- Every word has a valency
  - the dependents that it needs
  - (WG only) also its need for a ‘parent’ (a word on which it depends)
- These needs must be satisfied by other words
- Totally ‘bottom-up’.
- Every word also has a meaning
  - lexical meaning
  - modified by dependents
    - *babies* means ‘small babies’ when modified by *small*
    - *cry* means ‘small babies cry’ when modified by *babies* modified by *small*

4 How to choose between PS and DS?

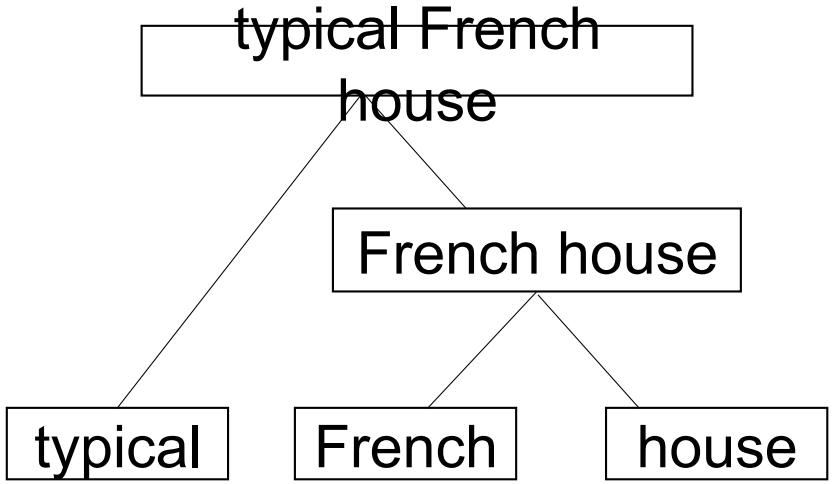
- Consider the facts
  - e.g. maybe c-command is important and requires PS?
- Look for elegance
  - count the nodes
- Look at general cognition
  - what kinds of relations can we recognise in general?
  - answer: many different kinds!!!
    - part-whole relations
    - social relations among individuals
    - spatial relations among objects
    - relations between events and their participants etc etc etc

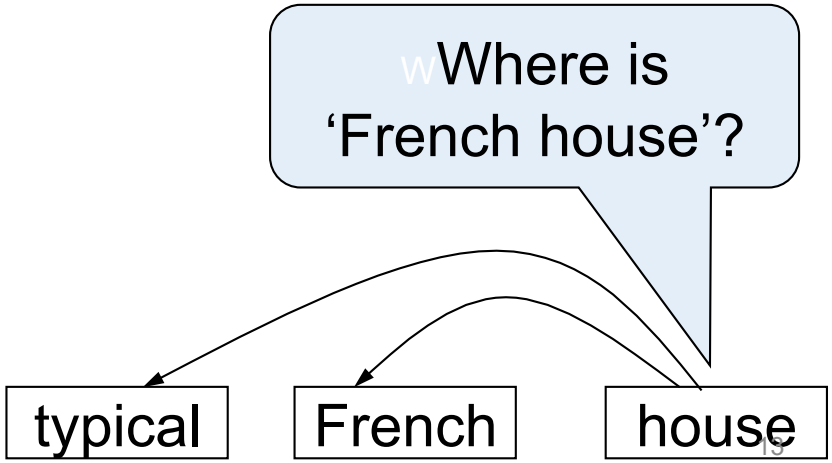
So the winner is ...

- Neither PS nor DS
  - because they both recognise only one kind of relation
  - and we know that our minds can recognise many different kinds.
- But DS is better than PS
  - because the evidence for word-word relations is stronger than for phrases:
  - lexical selection, e.g. DEPEND + ON
  - idioms, e.g. TAKE + CARE
  - government, e.g. MIT + Dative
- But there is a little evidence for phrase-like units ...

5 A challenge for DS

- *a big French house* = a house which is big and French
- But: *a typical French house* = a house which is typical of French houses.
  - Noticed by Oesten Dahl 1980

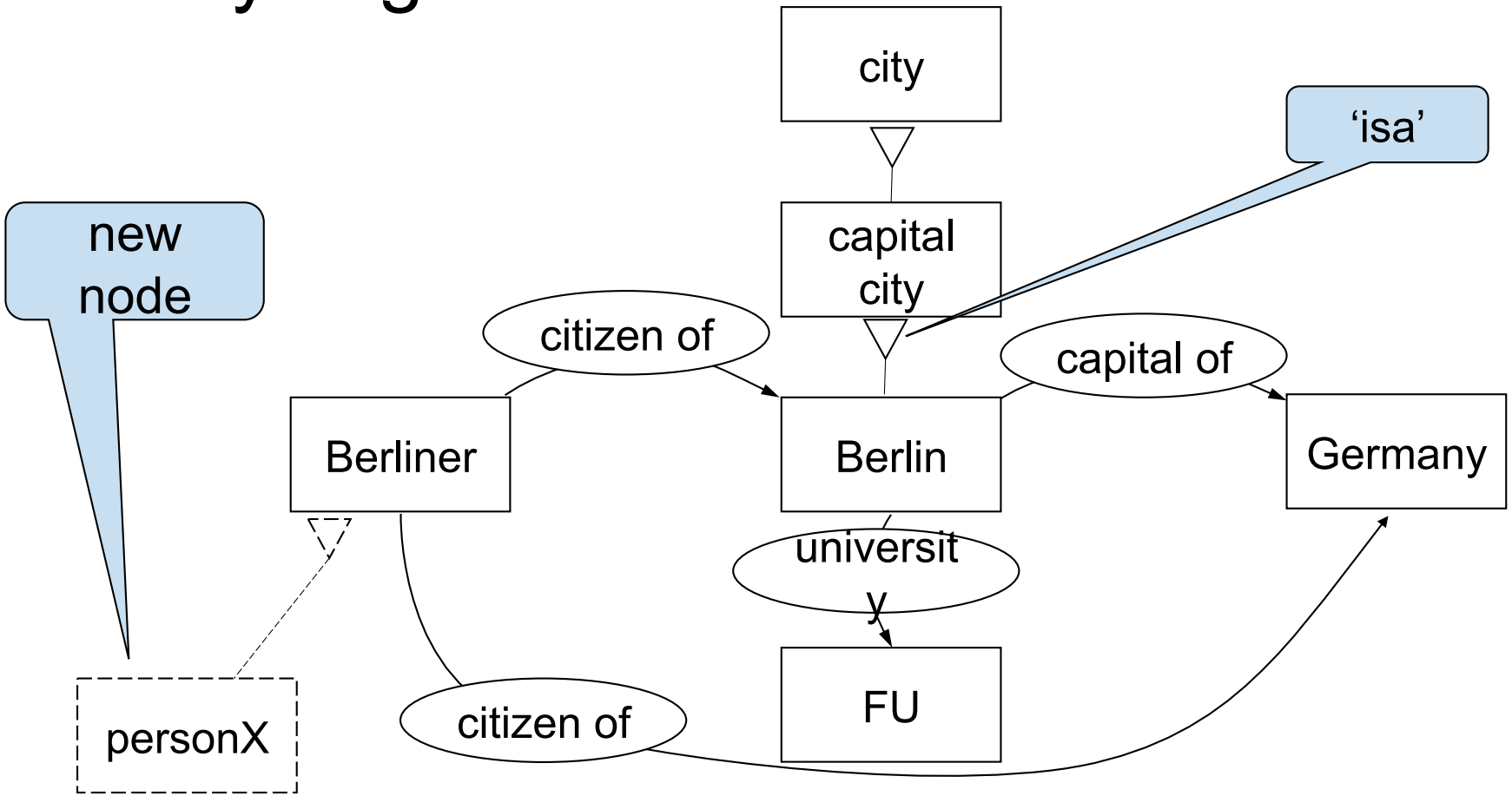




Solution: use what general cognition offers

- Knowledge is a network of atomic nodes.
- The network distinguishes different kinds of relation
  - including a vast and open collection of ‘relational concepts’, created as needed
- The ‘isa’ relation allows default inheritance.
- We create temporary nodes for experiences.
  - e.g. someone I saw on the street, ‘personX’

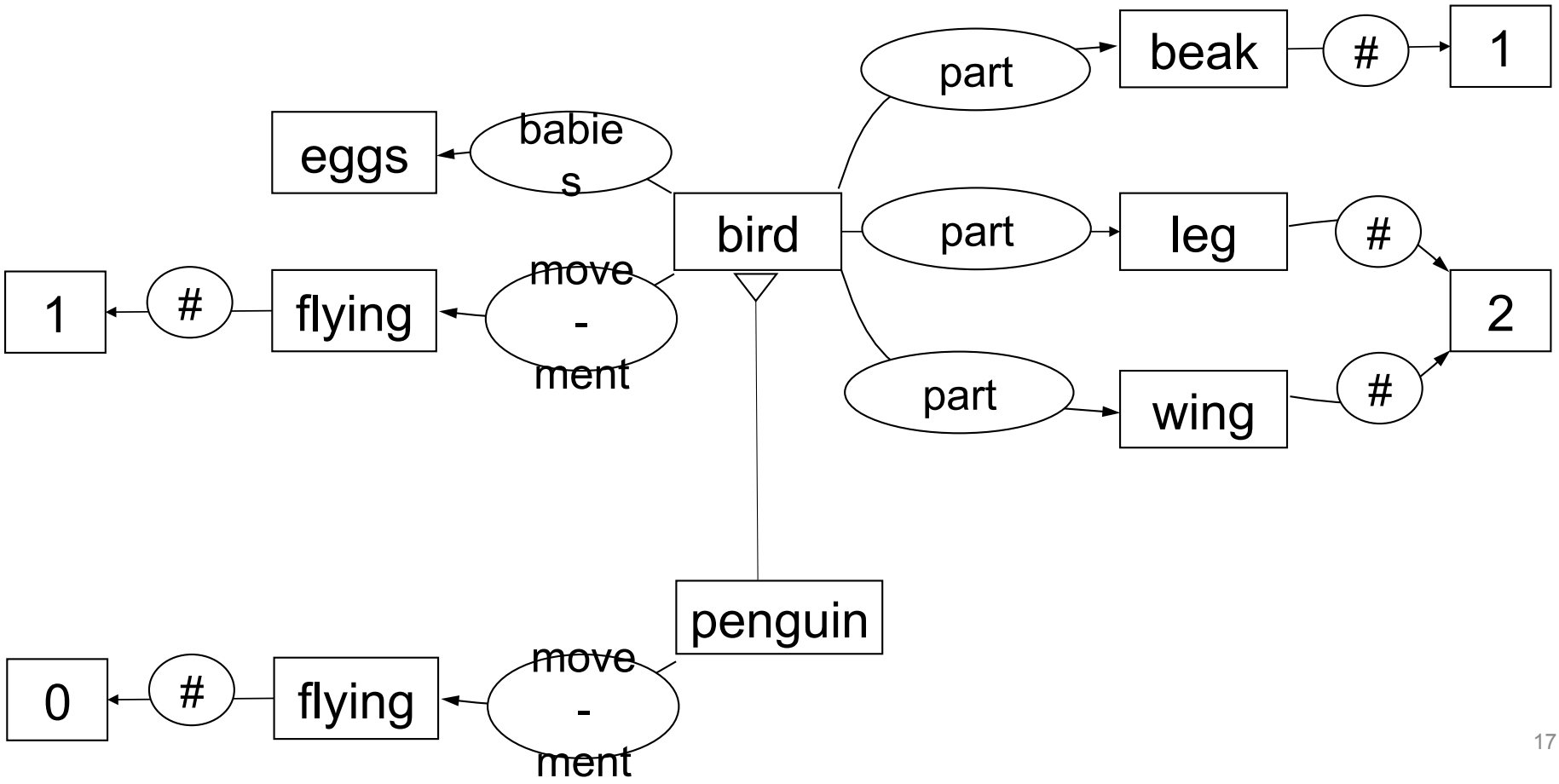
A tiny cognitive network



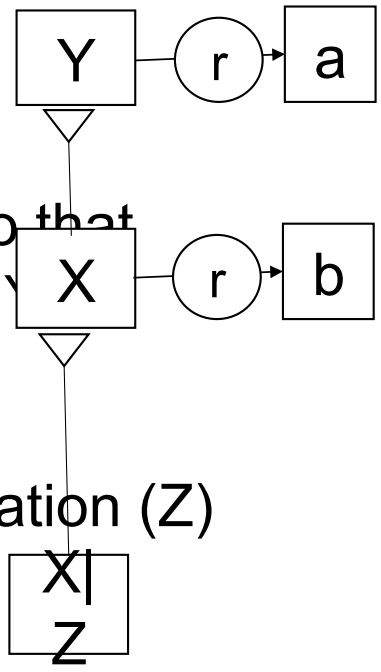
Default inheritance

- This is part of node-creation.
- If A isa B, then the properties of A always override those of B.
- So we can make generalisations even when there are exceptions.

Default inheritance in birds



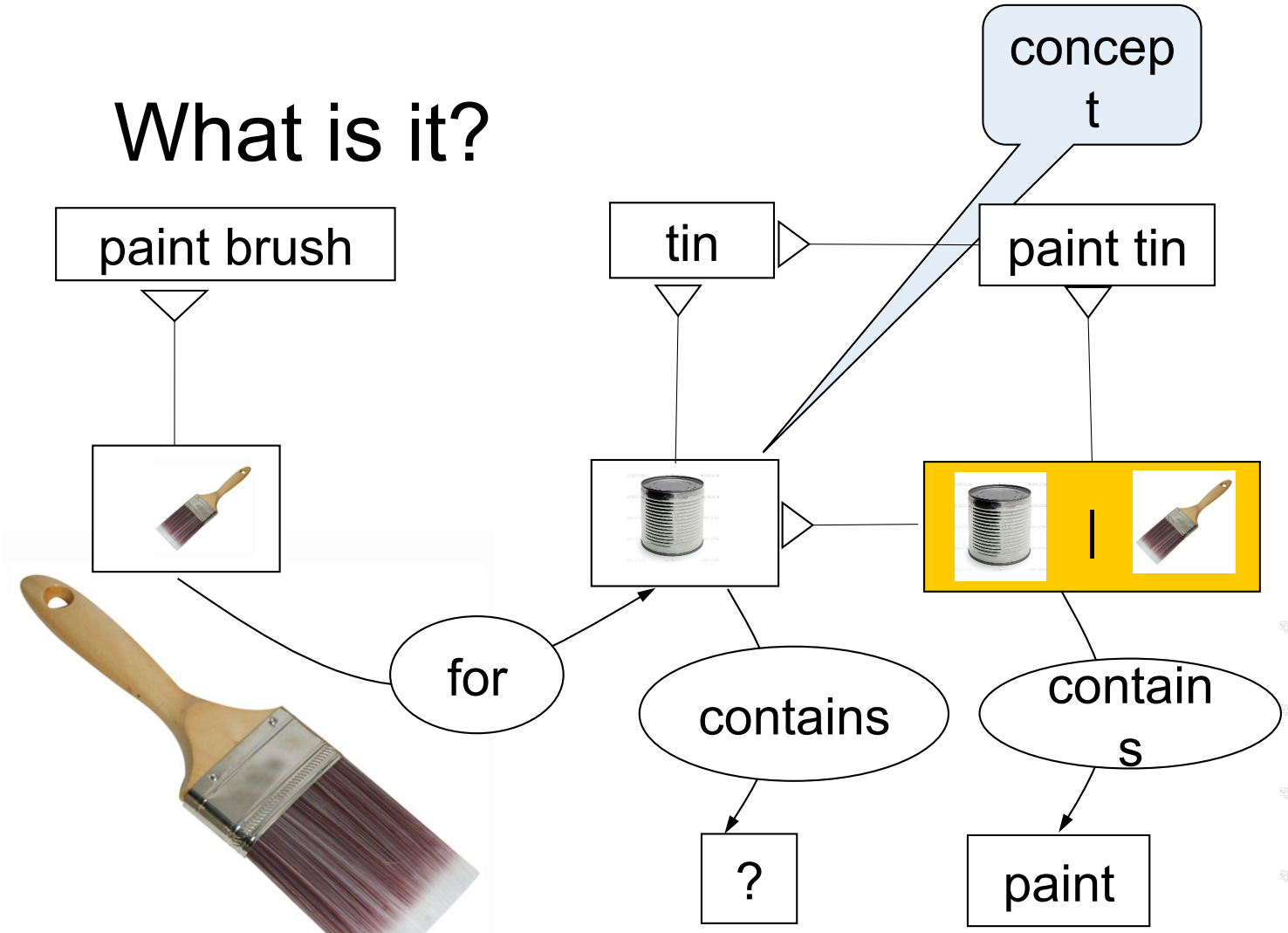
Node-creation



- We create a new node (‘X’) for ongoing experience so that
  - we can enrich it by classifying it as a Y, and inheriting from Y
  - we can distinguish the experience from Y
  - we can accommodate irregularities.
- And we continue to enrich it in the light of new information (Z)
  - by creating a further node (‘X|Z’, ‘X with Z’)
  - with isa links to X.
  - These nodes allow us to remember earlier states
  - and they’re the material of detailed ‘constructions’.
- For example ...

NB | ‘with’, not + ‘and’. Statistics: p(x|y) means ‘the probability of x in the context of y’

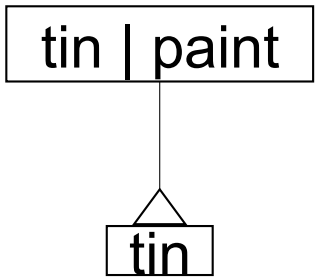
What is it?


concep
t
paint brush
tin
paint tin
for
contains
contain
s
?
paint

6 Towards a new kind of DS

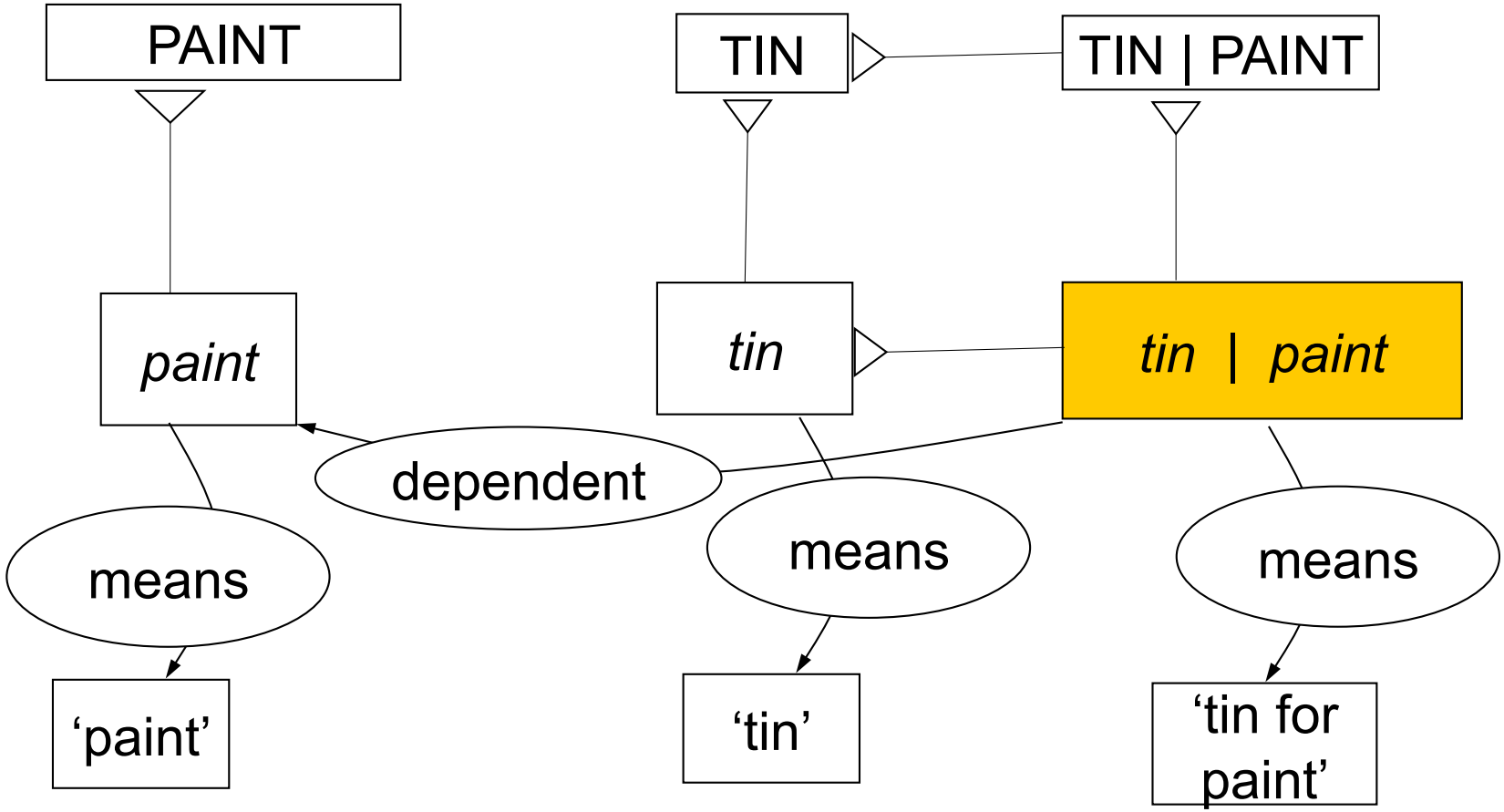
- Assume one initial node per word.
  - This inherits directly from some lexeme in the grammar.
  - e.g. in *paint tin*, we create one concept for each word token:

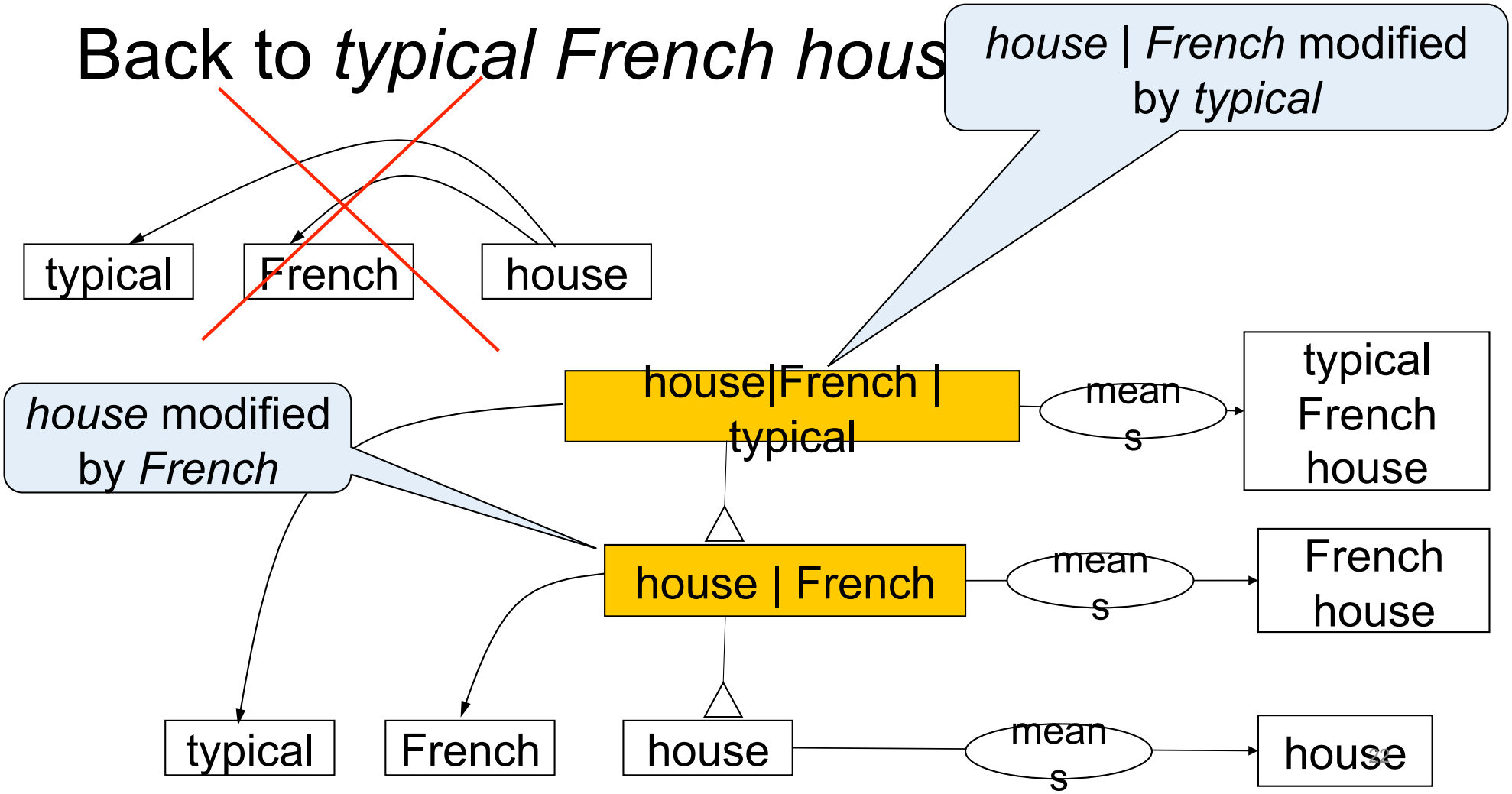
    paint tin
- But add an extra node for each dependent.
  - This shows how the word's meaning is affected by each dependent.
  - e.g. we create an extra concept for '*tin* as modified by *paint'*:



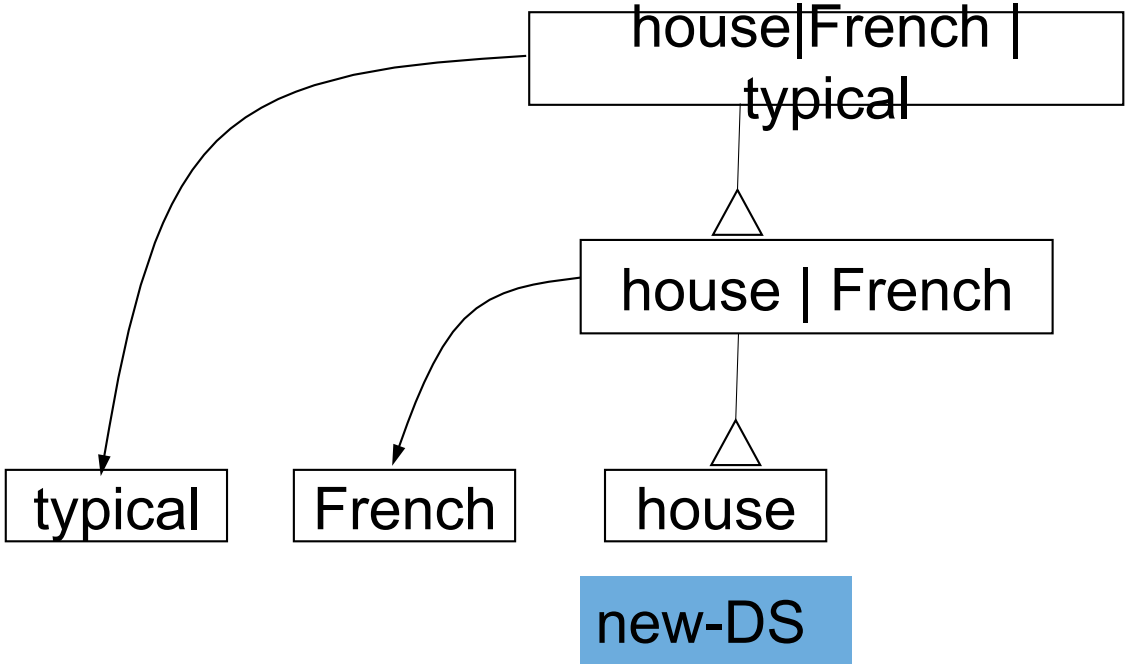
- And we link the two nodes by 'isa'.

New-DS: *paint tin*

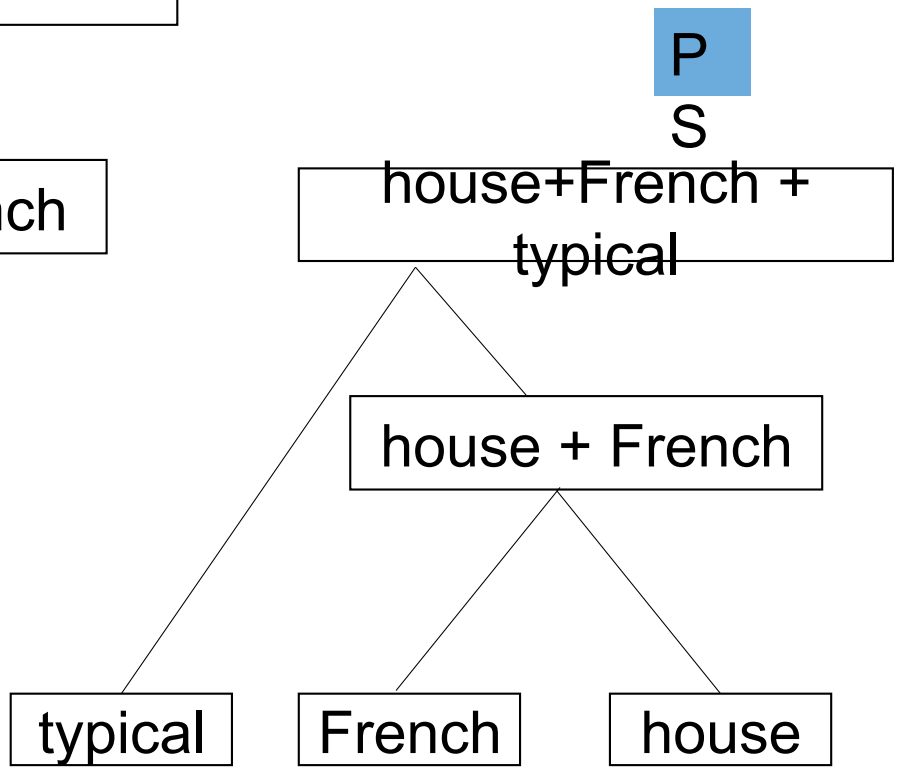



Back to typical French hous
house | French modified by typical
typical
French
house
house modified by French
house|French | typical
means
typical French house
house | French
means
French house
typical
French
house
means
house

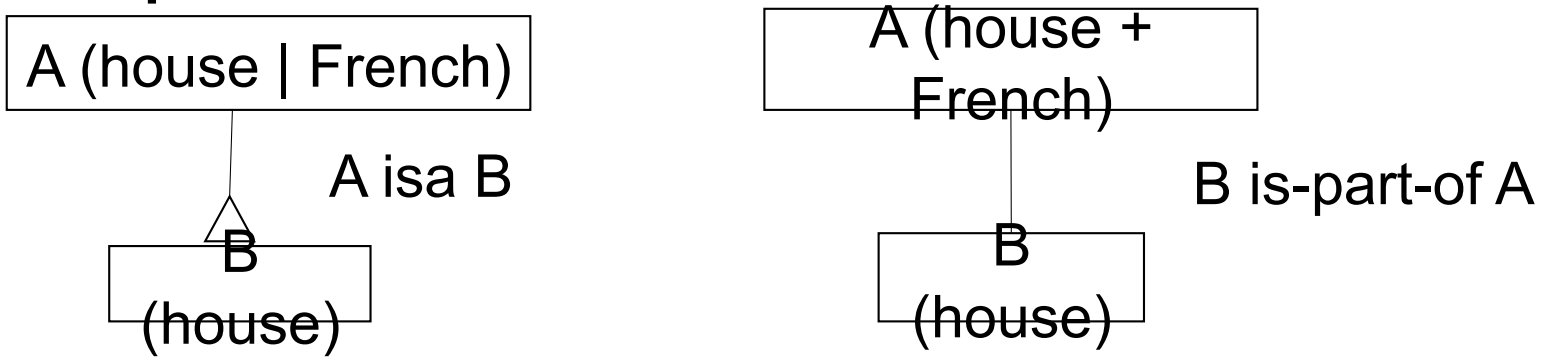
New-DS and PS



Notational variants??



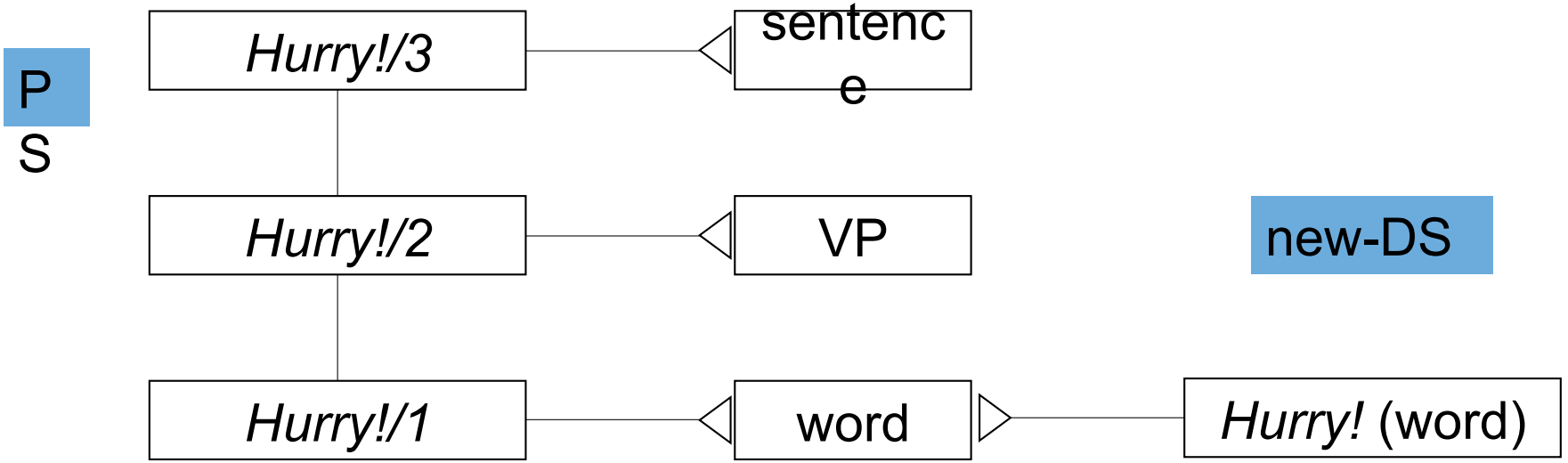
Isa, not part-of



- A isa B: like ‘Mary isa linguist’
  - shared properties
  - same size
- B is-part-of A: like ‘Mary’s foot is-part-of Mary’
  - different properties
  - different size

No unary branching in new-DS.

- PS needs both A and B because they have different properties, even when they have the same size.
- New-DS doesn't and can't.
  - 'higher' nodes are only needed where there's a dependent.

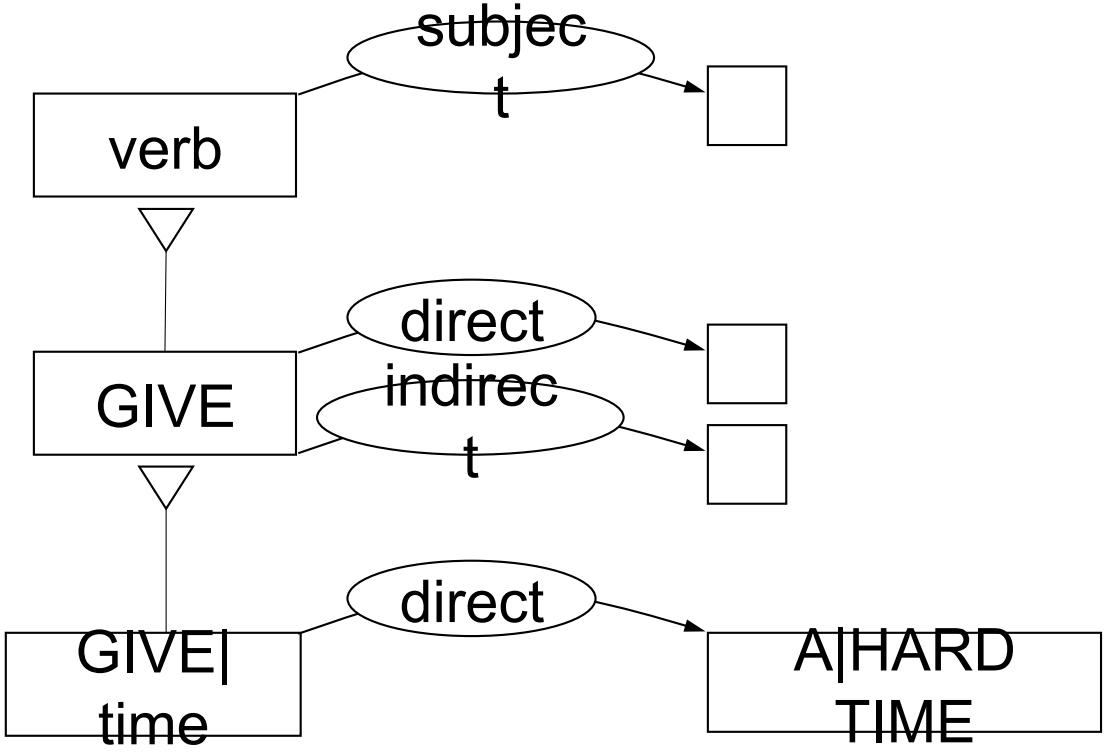


New-DS guarantees headedness

- A problem for DS?
  - *Student after student came in*. (Jackendoff)
- What is the head?
  - Answer: the first *student*, just like *tin of paint*.
- But why no determiner?
  - Stipulated, as in *to school, at home*
  - *wine from France* but: *the wine of France*
  - A construction definable, as usual, in terms of dependencies

In new-DS single dependencies are constructions

e.g. GIVE NP A HARD TIME

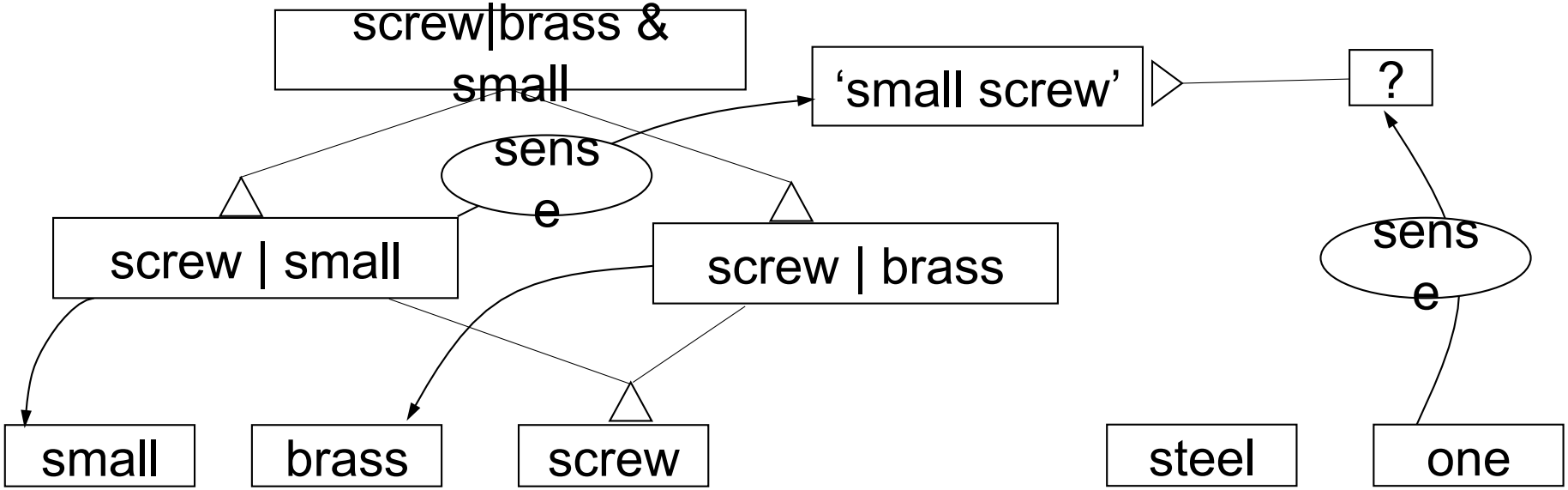


New-DS allows meaning-order mismatches

- *People are sometimes very tall*.
  - = Some people are very tall.
  - = sometimes (people are very tall)
- *John is typically late again.*
  - = John is late again, and John being late is typical.
- *I needed a small brass screw, but I could only find a steel one.*
  - *one* = ‘small screw’, not ‘brass screw’ or ‘small brass screw’!

How does New-DS allow this mismatch?

- *I needed a small brass screw, but I could only find a steel one.*
  - *one* = ‘small screw’.



8 Conclusions

- Syntactic theory should build on cognitive science.
- We should assume that our minds can apply any general-purpose cognitive machinery to language.
- This affects our assumptions about syntactic structure.
- It throws new light on the old dispute about PS versus DS.
- It allows us to develop a new version of DS which is more similar to PS.
- But even new-DS is different from PS
  - and better!

Danke für Ihre Aufmerksamkeit und Geduld!

- This slide show is available at

  dickhudson.com/talks

- Word Grammar offers much, much more ...
  - see dickhudson.com/word-grammar/