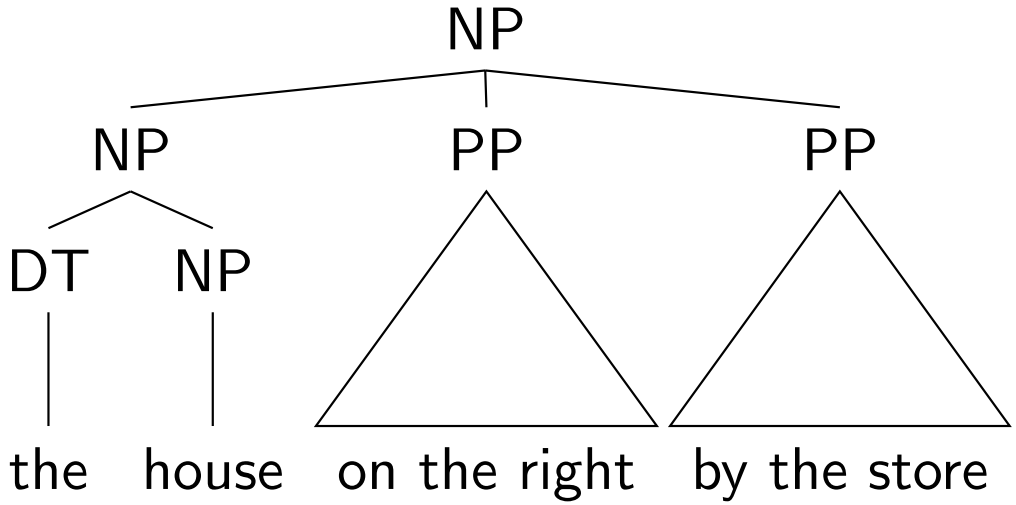
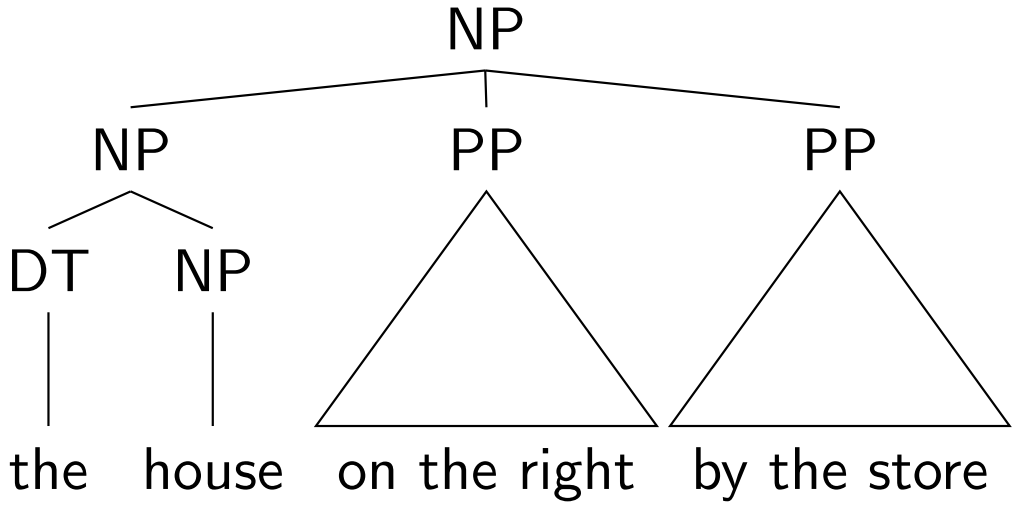
Subsumption

The PTB standard for multiple PPs is "two-level attachment"



Subsumption

The PTB standard for multiple PPs is "two-level attachment"

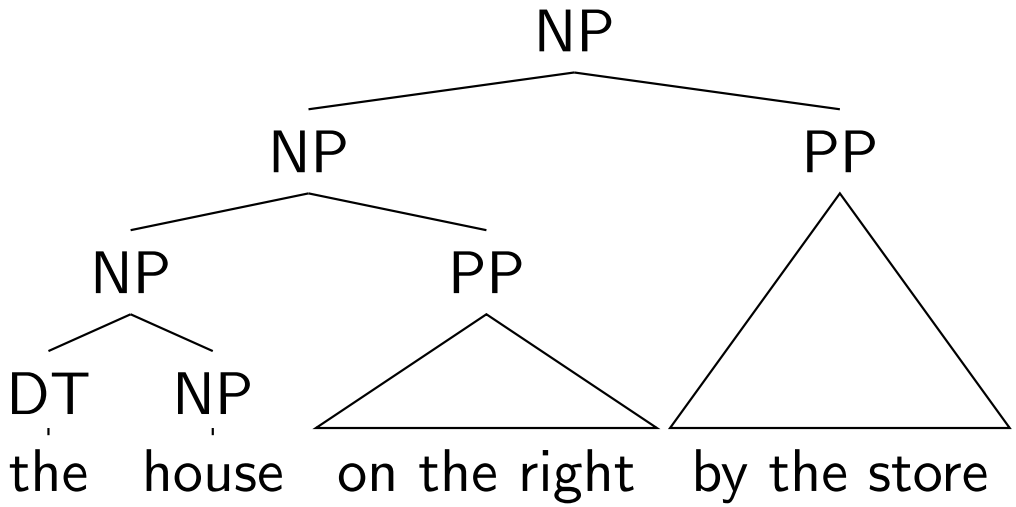


- However, multiple PPs are rare, and we have observed probabilities:

$$P(\mathrm{NP} \rightarrow \mathrm{NP\ PP}) = 0.112$$
$$P(\mathrm{NP} \rightarrow \mathrm{NP\ PP\ PP}) = 0.006$$

Subsumption

The PTB standard for multiple PPs is "two-level attachment"

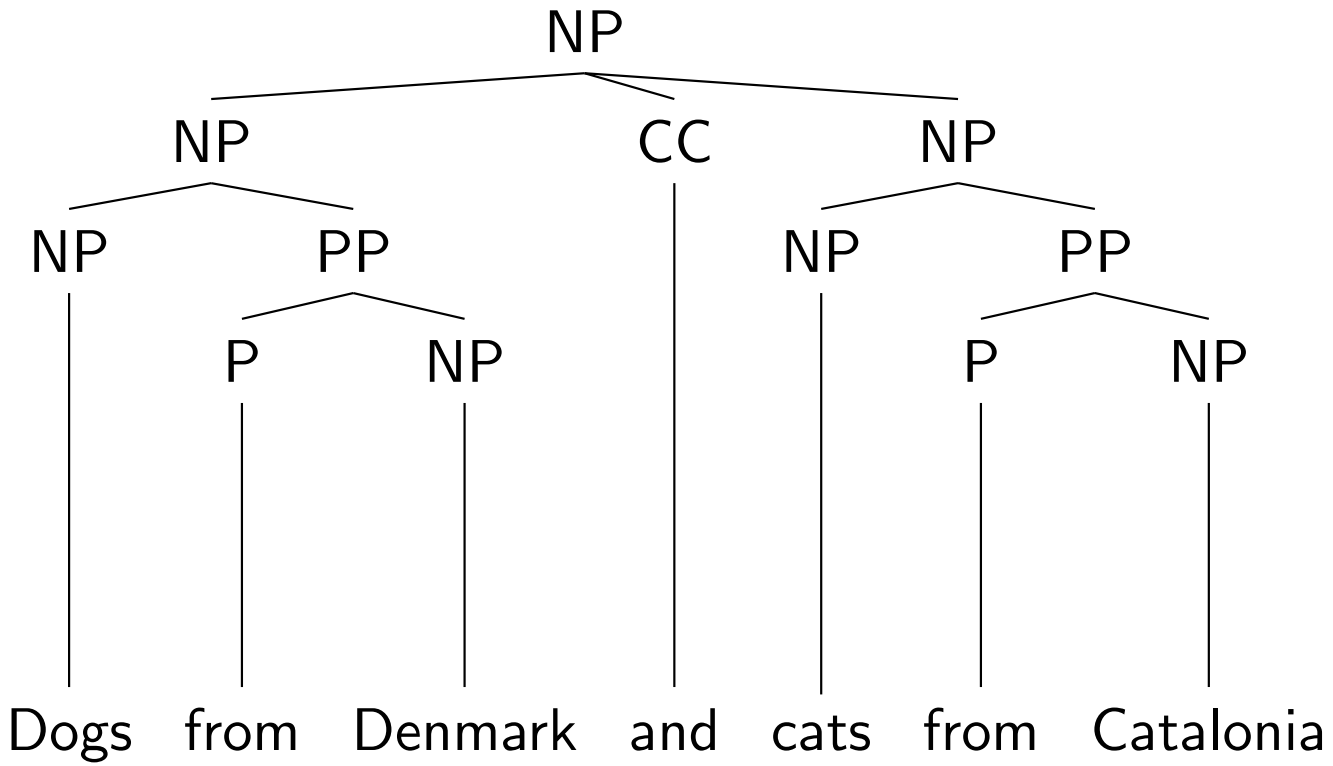


- However, multiple PPs are rare, and we have observed probabilities:

$$P(\text{NP} \rightarrow \text{NP PP}) = 0.112$$
$$P(\text{NP} \rightarrow \text{NP PP PP}) = 0.006$$
$$P(\text{NP} \rightarrow \text{NP PP})^2 = 0.013$$

- The PCFG parser will choose the "adjunction representation," even though it is never annotated this way.
- Johnson (1998) showed the 9% of all productions are **subsumed**.

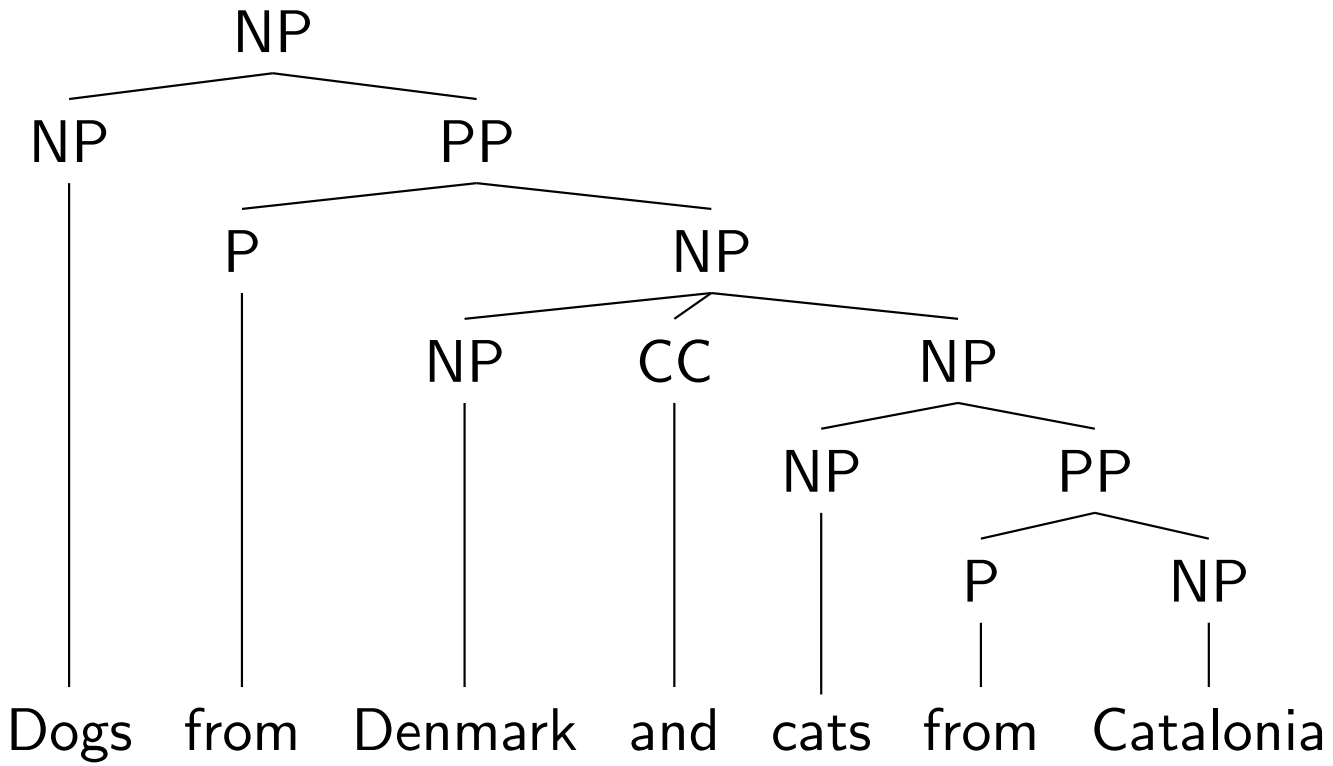
Equivalent parses under PCFGs



$\mathrm{NP} \rightarrow \mathrm{NP\ CC\ NP}$	1
$\mathrm{NP} \rightarrow \mathrm{NP\ PP}$	2
$\mathrm{PP} \rightarrow \mathrm{P\ NP}$	2

All parses get the same probability, regardless of the training data.

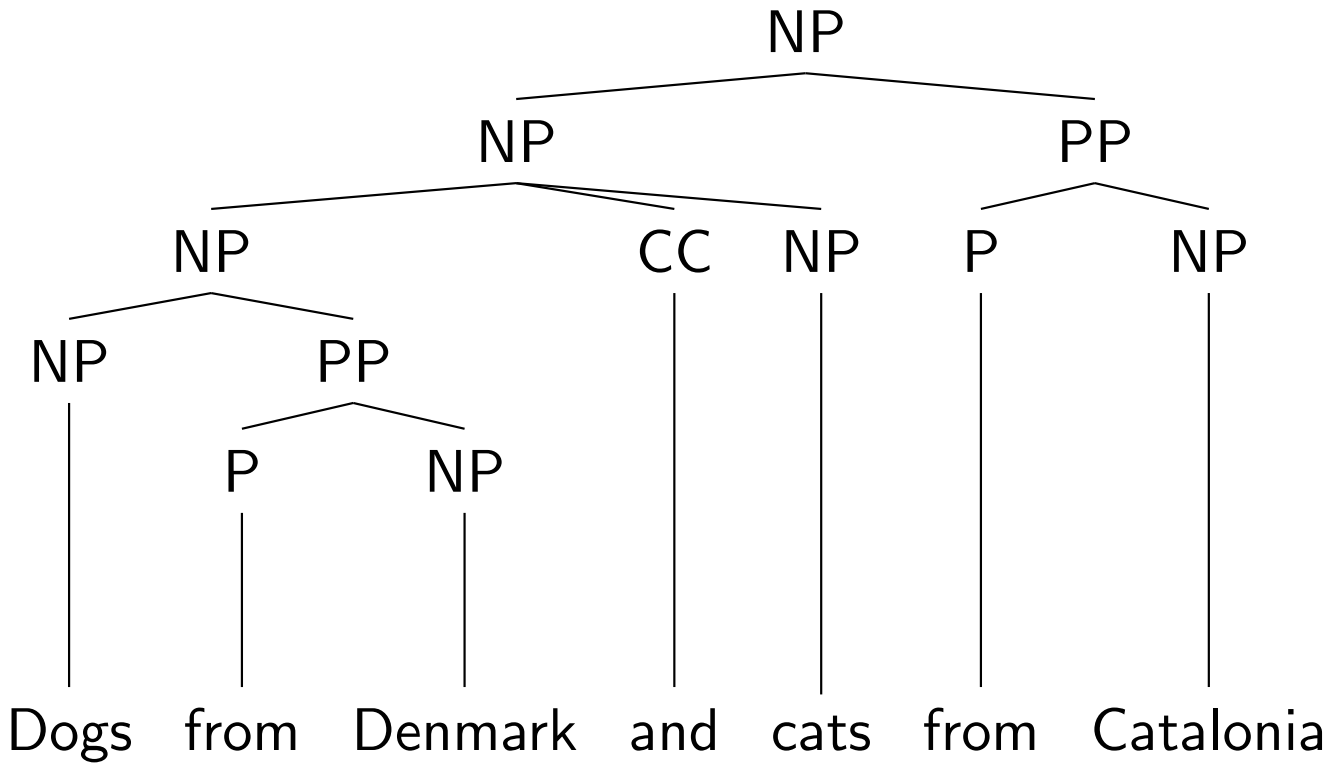
Equivalent parses under PCFGs



$\mathrm{NP} \rightarrow \mathrm{NP\ CC\ NP}$	1
$\mathrm{NP} \rightarrow \mathrm{NP\ PP}$	2
$\mathrm{PP} \rightarrow \mathrm{P\ NP}$	2

All parses get the same probability, regardless of the training data.

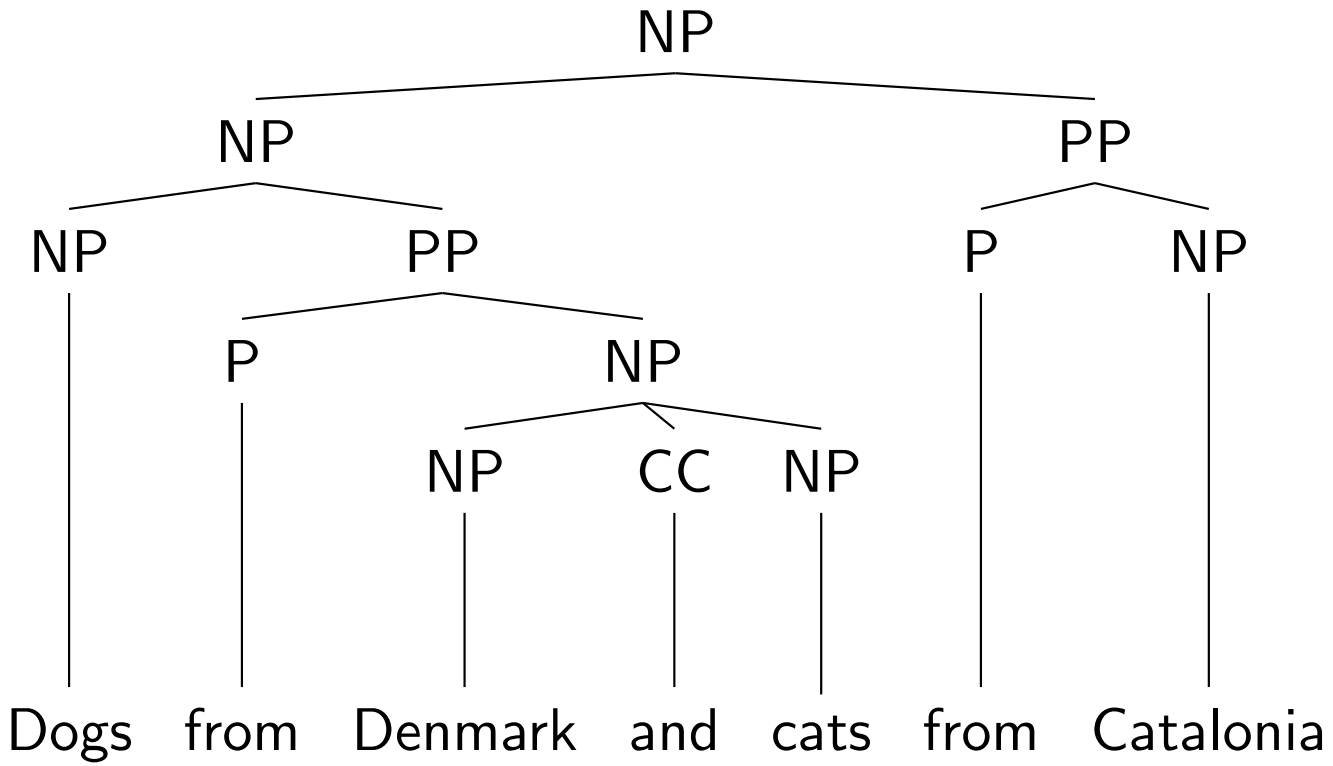
Equivalent parses under PCFGs



$\mathrm{NP} \rightarrow \mathrm{NP\ CC\ NP}$	1
$\mathrm{NP} \rightarrow \mathrm{NP\ PP}$	2
$\mathrm{PP} \rightarrow \mathrm{P\ NP}$	2

All parses get the same probability, regardless of the training data.

Equivalent parses under PCFGs



$\mathrm{NP} \rightarrow \mathrm{NP\ CC\ NP}$	1
$\mathrm{NP} \rightarrow \mathrm{NP\ PP}$	2
$\mathrm{PP} \rightarrow \mathrm{P\ NP}$	2

All parses get the same probability, regardless of the training data.