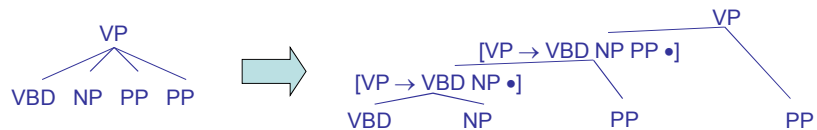


# Chomsky Normal Form

- Chomsky normal form:

- All rules of the form  $X \rightarrow YZ$  or  $X \rightarrow w$
- In principle, this is no limitation on the space of (P)CFGs
  - N-ary rules introduce new non-terminals



- Unaries / empties are “promoted”
- In practice it’s kind of a pain:
  - Reconstructing n-aries is easy
  - Reconstructing unaries is trickier
  - The straightforward transformations don’t preserve tree scores
- Makes parsing algorithms simpler!

# A Recursive Parser

```
bestScore(X,i,j,s)
  if (j = i+1)
    return tagScore(X,s[i])
  else
    return max score(X->YZ) *
               bestScore(Y,i,k) *
               bestScore(Z,k,j)
```

- Will this parser work?
- Why or why not?
- Memory requirements?