Starting Parsing with NLTK

Adam Meyers
Montclair State University
Preliminary

• Python and NLTK should work with any of the Language Lab machines
• If it does not, ask for help
• I spoke with Turker and he said if the monitors couldn't help, they would get the techies
Load nltk and load the grammar

• Import nltk
• Look at nltk book online chapter 7
• The Groucho grammar
  - groucho_grammar = nltk.CFG.fromstring(""
    S -> NP VP
    PP -> P NP
    NP -> Det N | Det N PP | 'I'
    VP -> V NP | VP PP
    Det -> 'an' | 'my'
    N -> 'elephant' | 'pajamas'
    V -> 'shot'
    P -> 'in'
  """
)
Parsing

- sent = ['I', 'shot', 'an', 'elephant', 'in', 'my', 'pajamas']
- parser = nltk.ChartParser(groucho_grammar)
- For tree in parser.parse(sent):
  print(tree)
- Assuming you get this to work and you want to modify the grammar keep in mind
  - Each rule is on a separate line
  - The vertical bar means “or”