Statistical NLP
Spring 2007
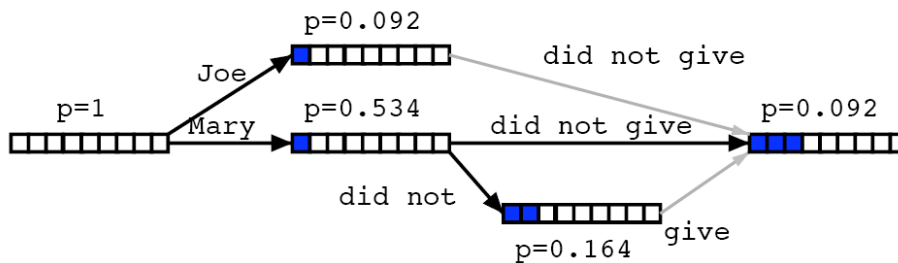
University of California
C A L
N L P
Berkeley

Lecture 13: Syntax

Dan Klein – UC Berkeley

Hypotheis Lattices

Maria	no	dio	una	bofetada	a	la	bruja	verde

Mary   not   give   a   slap   to   the   witch   green
did not   a slap   by   green witch
no   slap   to the
did not give   to
             the
slap   the witch

$p=0.092$

Joe

$p=1$

$p=0.534$ Mary   did not give

did not give

did not

$p=0.092$

give

$p=0.164$

Pruning

```
Maria no      dio una bofetada     a la      bruja verde
```
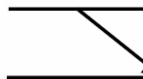
```
e: Mary did not
f: **-------
p: 0.154
```

```
e: the
f: -----**--
p: 0.354
```

better partial translation

covers easier part --> lower cost

- Problem: easy partial analyses are cheaper
  - Solution 1: use beams per foreign subset
  - Solution 2: estimate forward costs (A*-like)


WSD?

- Remember when we discussed WSD?
  - Word-based MT systems rarely have a WSD step
  - Why not?

What's Next?

- Modeling syntax
  - PCFGs and phrase structure
  - Syntactic parsing
  - Grammar induction
 - Syntactic language and translation models

Phrase Structure Parsing

- Phrase structure parsing organizes syntax into *constituents* or *brackets*
- In general, this involves nested trees
- Linguists can, and do, argue about details
- Lots of ambiguity

- Not the only kind of syntax…
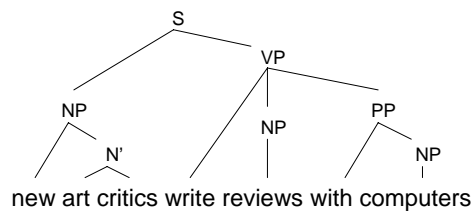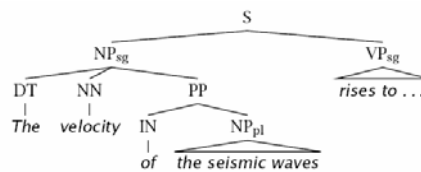
Constituency Tests

- How do we know what nodes go in the tree?

- Classic constituency tests:
  - Substitution by *proform*

  - Question answers

  - Semantic reference

  - Dislocation

- Cross-linguistic arguments, too
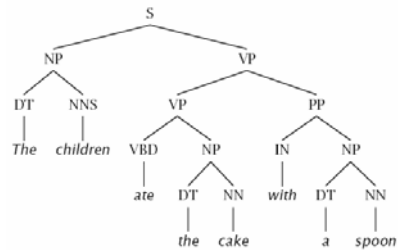
Conflicting Tests

- Constituency isn't always clear
  - Units of transfer:
    - think about ~ penser à
    - talk about ~ hablar de

  - Phonological reduction:
    - I will go → I'll go
    - I want to go → I wanna go
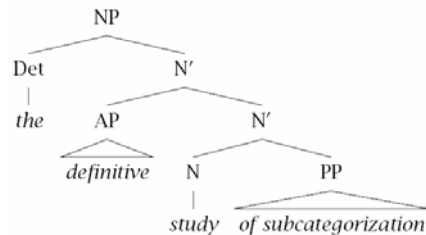   - a le centre → au centre
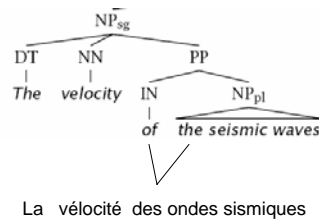
Non-Local Phenomena

- Dislocation / gapping
  - Why did the postman think that the neighbors were home?
  - A debate arose which continued until the election.

- Binding
  - Reference
    - The IRS audits itself
  - Control
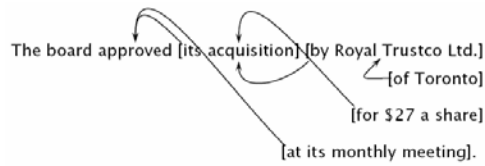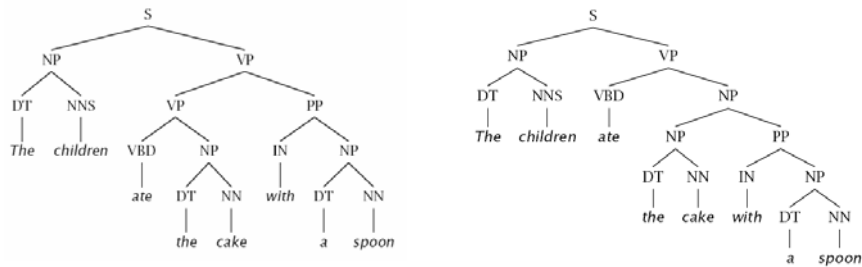   - I want to go
    - I want you to go
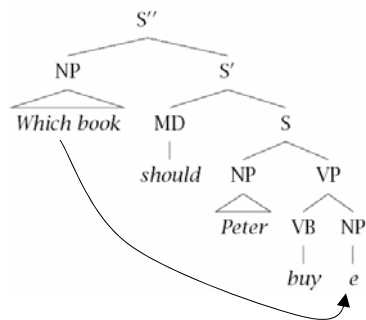
Regularity of Rules

- Argumentation
- Adjunction
- Coordination
- X' Theory

PP Attachment

PP Attachment

```
V             N1           P      N2         Attachment
--------------------------------------------------
join          board        as     director   V
is            chairman     of     N.V.       N
using         crocidolite  in     filters    V
bring         attention    to     problem    V
is            asbestos     in     products   N
making        paper        for    filters    N
including     three        with   cancer     N
```

Method	Accuracy
Always noun attachment	59.0
Most likely for each preposition	72.2
Average Human (4 head words only)	88.2
Average Human (whole sentence)	93.2

Attachment is a Simplification

- I cleaned the dishes from dinner

- I cleaned the dishes with detergent

- I cleaned the dishes in the sink

Syntactic Ambiguities I

- Prepositional phrases:
  They cooked the beans in the pot on the stove with handles.

- Particle vs. preposition:
  A good pharmacist dispenses with accuracy.
  The puppy tore up the staircase.

- Complement structures
  The tourists objected to the guide that they couldn't hear.
  She knows you like the back of her hand.

- Gerund vs. participial adjective
  Visiting relatives can be boring.
  Changing schedules frequently confused passengers.

Syntactic Ambiguities II

- Modifier scope within NPs
  impractical design requirements
  plastic cup holder

- Multiple gap constructions
  The chicken is ready to eat.
  The contractors are rich enough to sue.

- Coordination scope:
  Small rats and mice can squeeze into holes or cracks in the wall.

Treebank Sentences

```
( (S (NP-SBJ The move)
     (VP followed
         (NP (NP a round)
             (PP of
                 (NP (NP similar increases)
                     (PP by
                         (NP other lenders))
                     (PP against
                         (NP Arizona real estate loans)))))
         ,
         (S-ADV (NP-SBJ *)
                (VP reflecting
                    (NP (NP a continuing decline)
                        (PP-LOC in
                                (NP that market))))))
     .))
```

Human Processing

- **Garden pathing:**

  the man who hunts ducks out on weekends

  the cotton shirts are made from grows in Mississippi

  the daughter of the king's son loves himself

- **Ambiguity maintenance**

  Have the police ... eaten their supper?
                      come in and look around.
                      taken out and shot.