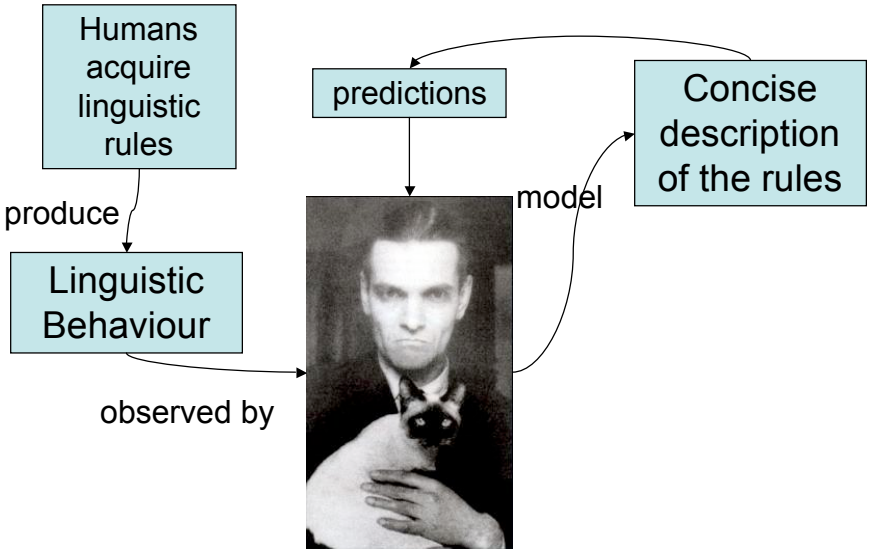
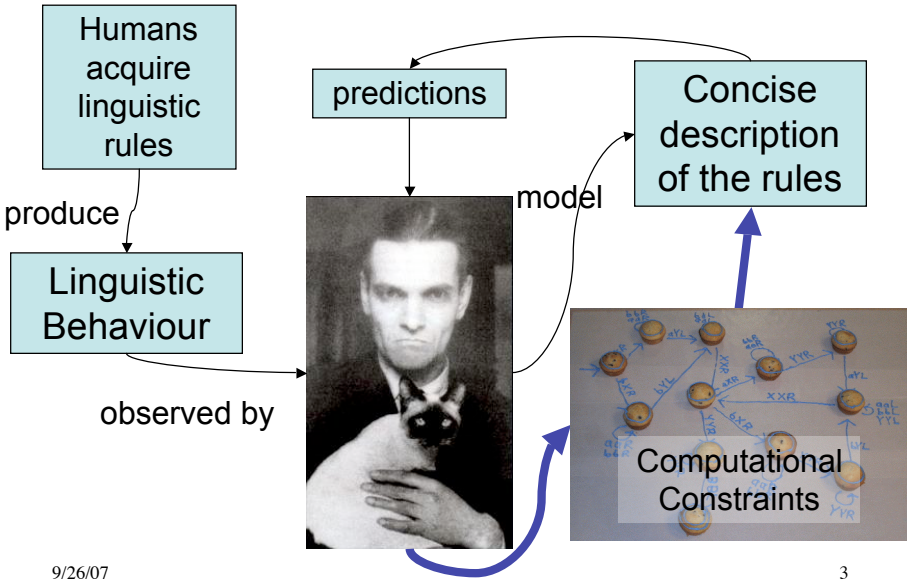
Computational Constraints on Linguistic Descriptions

Anoop Sarkar
Computing Science, SFU
anoop@cs.sfu.ca

Life-cycle of a Linguist



Life-cycle of a Linguist (revised)



Phonological Alternations

- Newton
- Paris
- practical
- feasible

- Newtonian
- Parisian
- impractical
- infeasible

Finnish Harmony

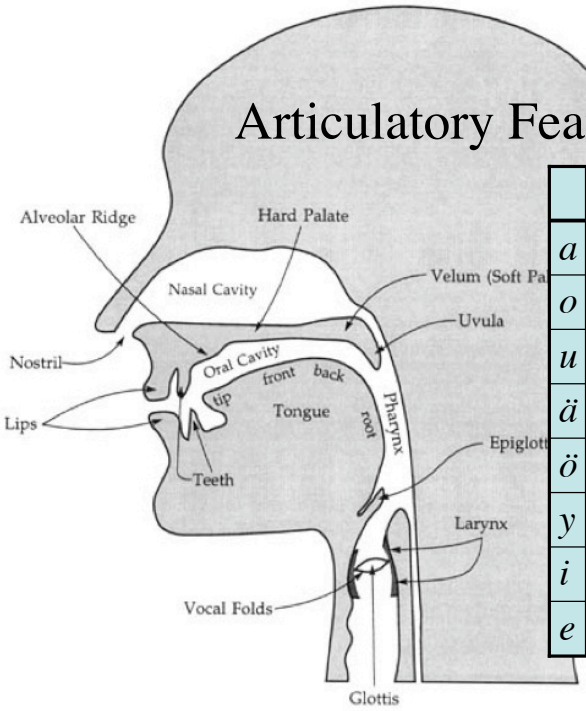
Gloss	Nominative	Partitive
• sky	• taivas	• taivas+ta
• telephone	• puhelin	• puhelin+ta
• plain	• lakeus	• lakeut+ta
• reason	• syy	• syy+tä
• short	• lyhyt	• lyhyt+tä
• friendly	• ystävällinen	• ystävällinen+tä

i,*e* are neutral wrt harmony

talossansakaanko ‘not in his house either?’
kynässänsäkäänkö ‘not in his pen either?’

Rewrite Rules (Chomsky & Halle, 1968)
a → ä / [ä,ö,y] C* ([i,e] C*)* ___
o → ö / [ä,ö,y] C* ([i,e] C*)* ___

Articulatory Features



	high	low	back	round
a	–	+	+	–
o	–	–	+	+
u	+	–	+	+
ä	–	+	–	–
ö	–	–	–	+
y	+	–	–	+
i	+	–	–	–
e	–	–	–	–

Articulatory Features

Finnish Partitive
- taivas+ta
- puhelin+ta
- lakeut+ta
- syy+tä
- lyhyt+tä
- ystävällinen+tä

	high	low	back	round
a	–	+	+	–
o	–	–	+	+
u	+	–	+	+
ä	–	+	–	–
ö	–	–	–	+
y	+	–	–	+
i	+	–	–	–
e	–	–	–	–

Rewrite Rule

[V,+back] → [V, –back] / [V, –back] C* ([i,e] C*)* ___

Rewrite Rules

- Rewrite rules specify a relation between base-forms and observed forms, e.g. google+ing → googling
- Assume base-forms come from (C*(u|i))+ C* (C*V)* C*

V → i / i C* __

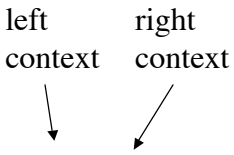
V → u / u C* __

kikVkVkV
kikVkVkV
kikikVkV
kikikVkV
kikikikV
kikikikV
kikikiki

output of one application *feeds* next application

left to right application

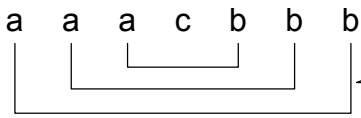
Rewrite Rules



- Context dependent rewrite rules: $\alpha \rightarrow \beta / \lambda \_\_ \rho$
  - ($\lambda \alpha \rho \rightarrow \lambda \beta \rho$; that is α becomes β in context $\lambda \_\_ \rho$)
- How to apply rewrite rules:
  - Consider rewrite rule: a → b / ab __ ba
  - Apply rule on string *abababababa*
  - Three different outcomes are possible:
    - *abbbabbbaba* (left to right, iterative)
    - *ababbbabbba* (right to left, iterative)
    - *abbbbbbbbba* (simultaneous)

Rewrite Rules

- Context dependent rewrite rules: $\alpha \rightarrow \beta / \lambda \_\_ \rho$
- Can express **context sensitive** rules or **regular** relations
- Computational constraints on rewrite rules:
  - Consider rewrite rule: c → acb / a __ b
  - Apply left to right iteratively on base-form *c*
  - Produces a sequence of strings:



Do we need such long-distance effects in morpho-phonological rules?

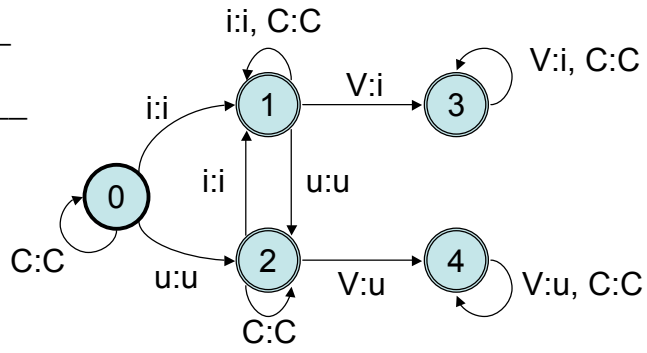
Computational Constraints on Rewrite Rules

- Rewrite rules express a context-sensitive grammar: $\lambda\ \alpha\ \rho \rightarrow \lambda\ \beta\ \rho$, cf. $\alpha \rightarrow \beta\ /\ \lambda\ \_\_\ \rho$
- CSGs are very powerful: they can generate languages like $\{ 1^p : p \text{ is prime} \}$
- Kaplan and Kay:
  - Impose a simple constraint on how rewrite rules are applied: output cannot be re-written
    e.g. c → acb / a __ b
  - Constraint ensures rewrite rules are equivalent to regular relations
  - Naturally expresses the **local** nature of morpho-phonemic properties

Computational Constraints on Rewrite Rules

V → i / i C* __

V → u / u C* __



*kikukuku
√kikikikiki

Finite state transducer
•each edge maps *input*:*output*
•defines a **regular** relation

Constraint-based approach

- Instead of explicit rules assume all variants are generated and surface constraints filter out illegal variants
  - Finnish: generate alternatives with both -ta and -tä
  - -ta is filtered out after [–back] vowels
  - -tä is filtered out after [+back] vowels
- Optimality Theory (Prince & Smolensky)
  - GEN produces all possible forms
  - rank-ordered violable constraints used to assign violations
  - form with least number of violations is produced

Constraint-based approach

- Karttunen:
  - No computational difference between traditional ordered rules and OT
  - Traditional ordered rules can be "compiled into" a regular relation with composed transducers
  - OT is "compiled into" a regular relation with *leniently composed* transducers
  - This notion of lenient composition captures the linguistic intuition: surface form has least violations
  - OT constraints must obey computational constraints (cf. Jason Eisner)

Structural Constraints in Syntax

wh- questions in English

did Alice read which book

which book$_i$ did Alice read t_i

has Leona said that Alice would read which book

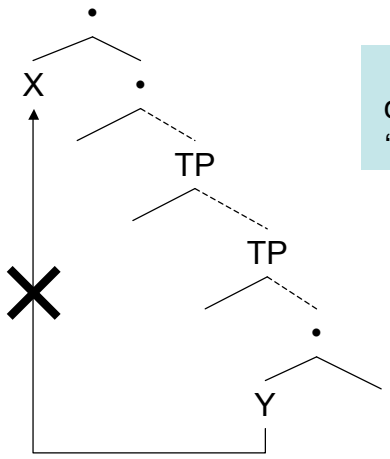
which book$_i$ has Leona said that Alice would read t_i

Mark asked to whom you had lent War & Peace

did Mark ask to whom you had lent what book

*what book$_i$ did Mark ask to whom you had lent t_i

The Subjacency Condition



non-local condition on "movement"

The Subjacency Condition

which book$_i$ did [TP Alice read t_i]

which book$_i$

has [TP Leona said [CP t_i that [TP Alice would read t_i]]]

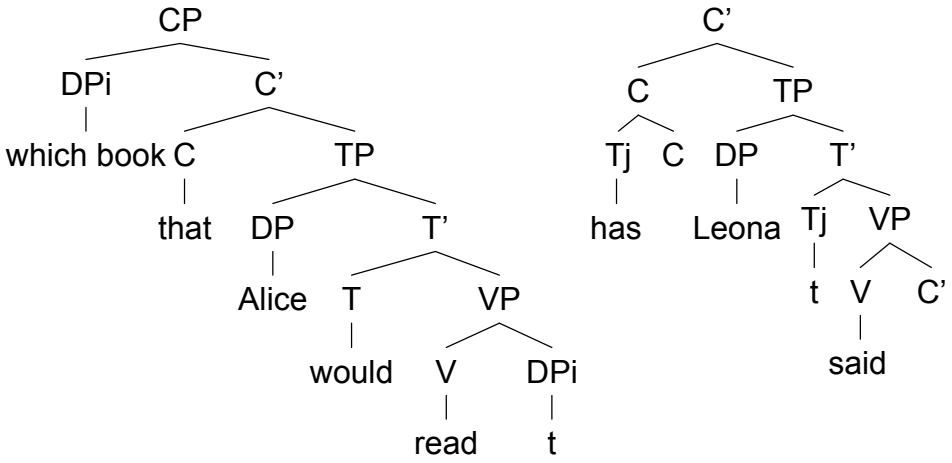
*what book$_i$ did [TP Mark ask to whom [TP you had lent t_i]]

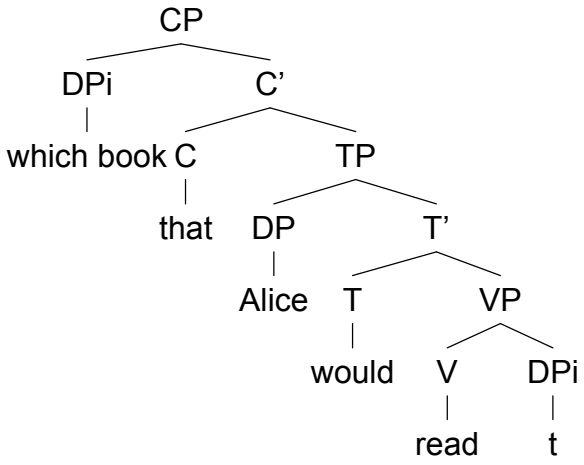
Tree-Adjoining Grammars

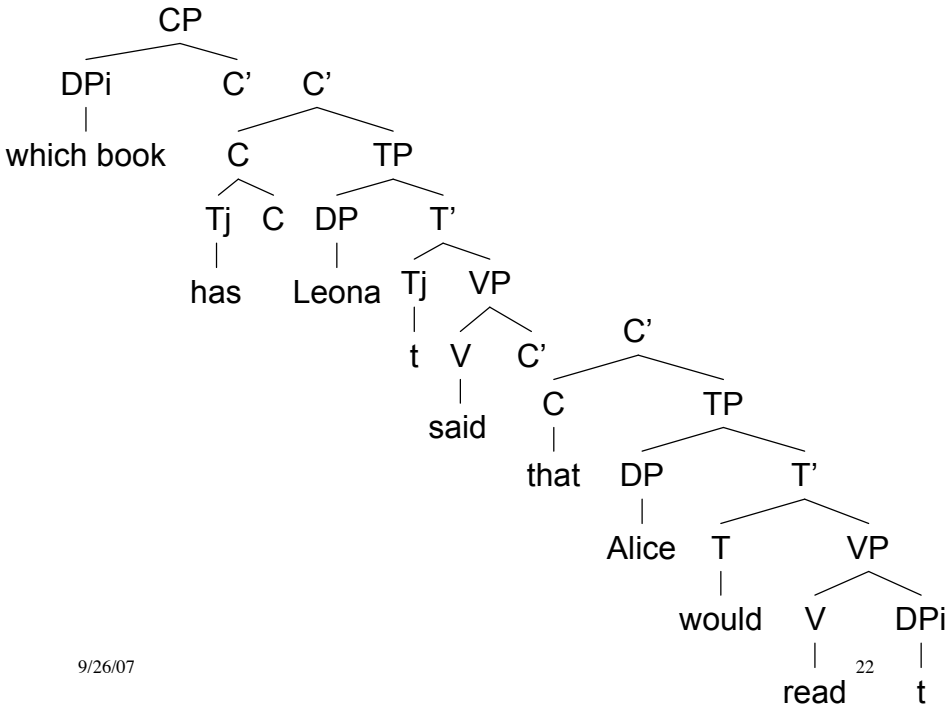
- Construct a tree set out of tree fragments
- Each fragment contains only the structure needed to express the locality of various CSG predicates
- Each tree fragment is called an elementary tree
- In general we need to expand even those non-terminals that are not leaf nodes: leads to the notion of adjunction

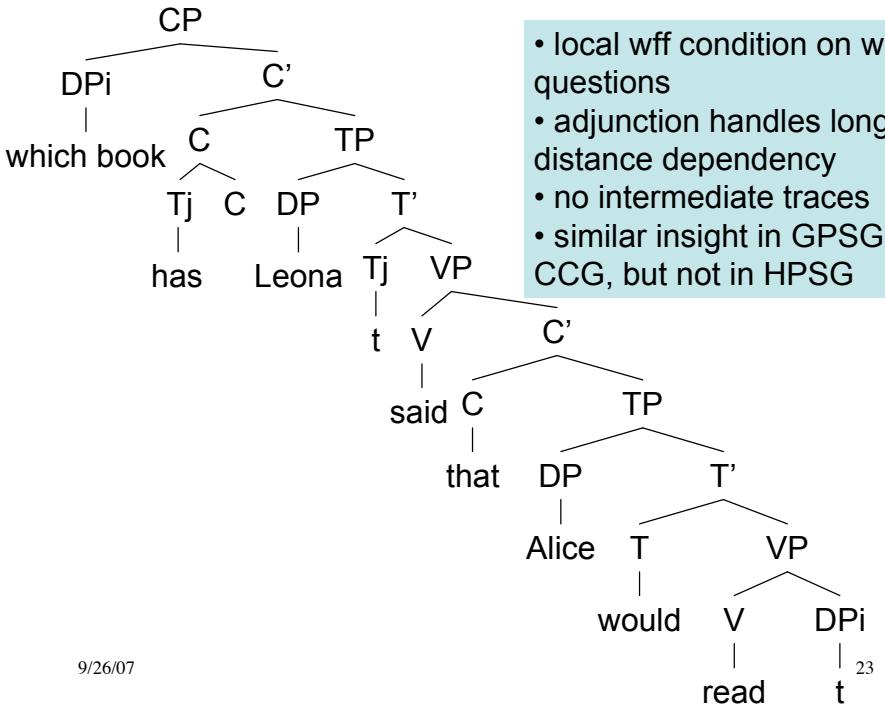
TAG Analysis

(Kroch, Frank)




CP
DPi
C'
which book
C
TP
that
DP
T'
Alice
T
VP
would
V
DPi
read
t

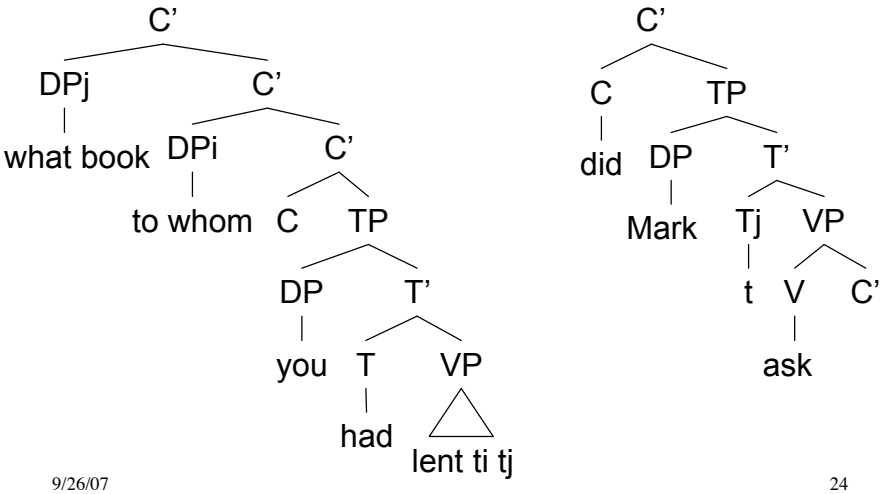

CP
DPi
C'
C'
which book
C
TP
Tj
C
DP
T'
has
Leona
Tj
VP
t
V
C'
said
C
TP
that
DP
T'
Alice
T
VP
would
V
DPi
read
t



- local wff condition on wh-questions
- adjunction handles long-distance dependency
- no intermediate traces
- similar insight in GPSG, CCG, but not in HPSG

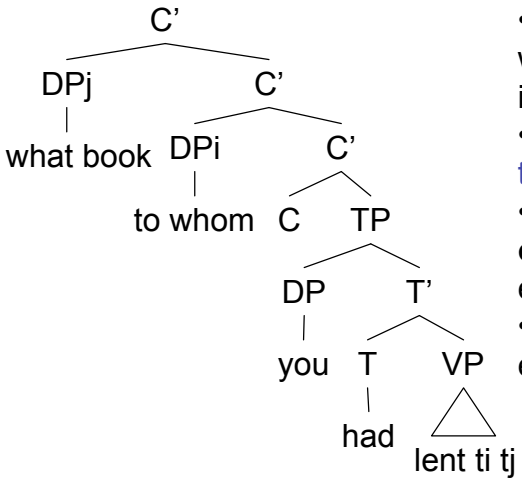
TAG Analysis

*what book$_i$ did Mark ask to whom you had lent t$_i$



TAG Analysis

*what book$_i$ did Mark ask to whom you had lent t_i



- Kroch: such a tree is not well formed in English even in a single clause
- e.g. *I wonder what book to whom Mark gave
- Linguistic rules are used to **only** construct well formed elementary trees
- Computational model enforces local constraints

What about languages that allow multiple wh- fronting? e.g. Romanian

Summary

- Linguistic descriptions have computational properties: being sensitive to these properties provides insight
- Computational complexity explains the existence of locality in applicability of linguistic rules
- It is interesting from a computational viewpoint exactly how complex are the generalizations produced by linguists
- The computational treatment allows the "compilation" of linguistic rules into automata: a constructive proof of their efficacy