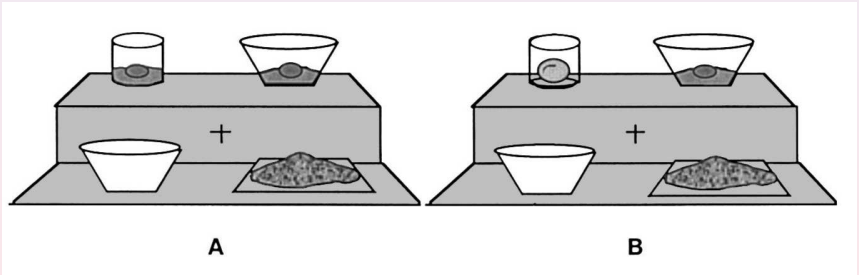


# Grammar in Performance and Acquisition

E Stabler, UCLA

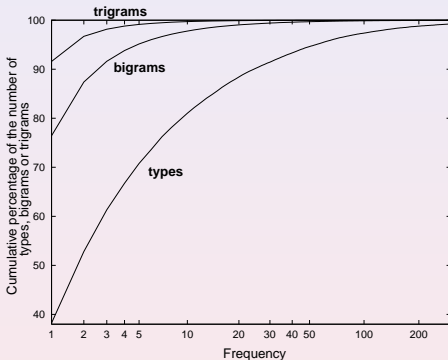
ENS Paris • 2008

(Chambers et al., 2004)



'Pour the egg (that's) in the bowl over the flour'

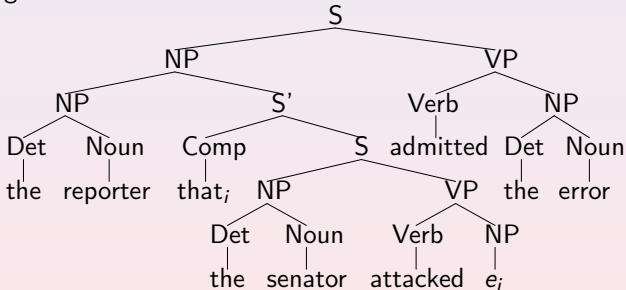
**Q1** How are utterances interpreted 'incrementally'?



tb2:  $\approx 40\%$  words unique, 75% bigrams, 90% trigrams, 99.7% sentences  
 $\Rightarrow$  most sentences heard only once

**Q2** How are linguistic abilities acquired from available evidence?

If words are encountered that necessitate other syntactic heads to form a grammatical sentence, then these categories are also predicted, and an additional memory load is incurred. For example, ... at the point of processing the second occurrence of the word “the”...



there are four obligatory syntactic predictions: (1) a verb for the matrix clause, (2) a verb for the embedded clause, (3) a subject noun for the embedded clause, and (4) an empty category NP for the wh-pronoun “who.” (Gibson, 1998, pp.13-14)



we get evidence of recognition mechanisms, and of how learners generalize, from *what we find in languages*

- we don't need to start from zero

*... processing can be seen as the rapid incremental satisfaction of grammatical constraints. . . which are needed independently (Weinberg'00)*

- seek broad solutions with convergent evidence

*Rational arguments about two theories' comparative success. . . depend on a broad assessment of their properties; lacking that, such discussions not infrequently descend into the cherry-picking of isolated favorable and unfavorable instances. (Prince'07)*

- (Greenberg 1966; cf Cinque 2005, Abels&Neeleman 2006)

D Num A N      1234      4123      4321      \*4213

- (Koopman & Szabolcsi 2000) Verbal complexes in Hungarian

V<sub>1</sub> V<sub>2</sub> M-V<sub>3</sub>      123      321      132      \*213

**Q3** Why are some constituent orders unattested across languages?

Japanese (Potts et al '07):

yomu koto wa yon-da  
 read nom part read-pst  
 'I read (but didn't necessarily understand)'

Yoruba (Kobele '06): copies of copies predicted in embedded relatives

Ri-ra adie ti Jimo ra adie  
 buying chicken rel Jimo buy chicken  
 'the fact that Jimo bought chicken'

Coll. Icelandic (Barbiers '07)

Um havð eruð þið að tala um?  
 about what are you to talk about  
 'What are you talking about?'

**Q4** What kind of grammar model makes copying a natural option?

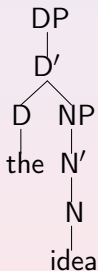
- Q1 How are utterances interpreted 'incrementally'?
- Q2 How is that ability acquired, from available evidence?
- Q3 Why are some constituent orders unattested across languages?
- Q4 What kind of grammar makes copying a natural option?
- we don't need to start from zero
  - frame explanations supported by convergent evidence

(instead of starting from zero, let's start from a family of grammars)

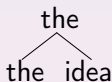


## first 'minimalist' grammars (MG)

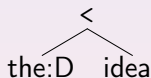
NOT:



NOT:



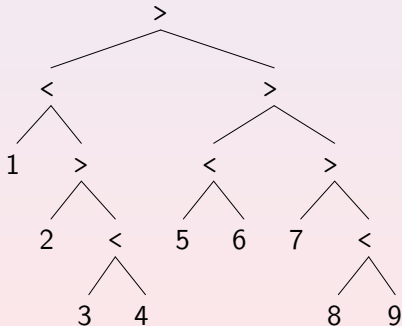
But:



The < “points toward” the **head** of the phrase.

The largest subtree with a given head is a **maximal** projection.

## Practice



|                            |              |
|----------------------------|--------------|
| every,some,student,...     | (vocabulary) |
| C, T, D, N, V, P,...       | (categories) |
| =C, =T, =D, =N, =V, =P,... | (selectors)  |
| +wh, +case, +focus,...     | (licensors)  |
| -wh, -case, -focus,...     | (licensees)  |

Examples:

Marie::D

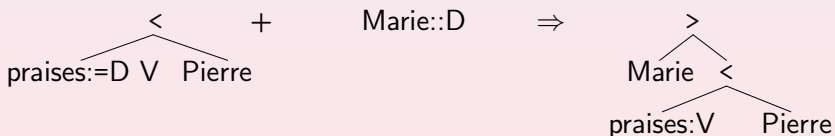
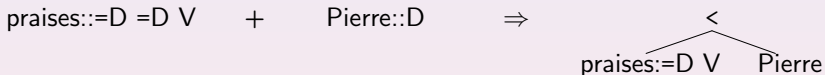
who::D -wh

praises::=D =D V

ε::=I +wh C

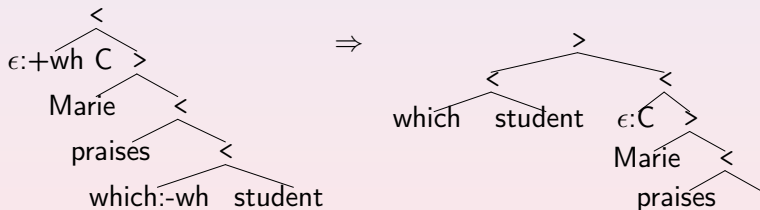
These lexical items combined by *merge*...

**External merge (em)** complements on right, additional selected elements on left



(2 features deleted, and :: in lexical items changes to : in derived structures)

**Internal merge (im)** in a tree whose head has first feature +f, move maximal -f subtree specifier position:

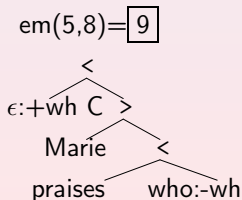
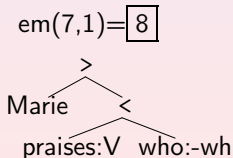
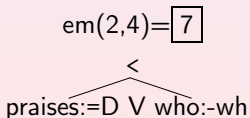


(SMC) *im* applies only when exactly 1 head has -f first feature

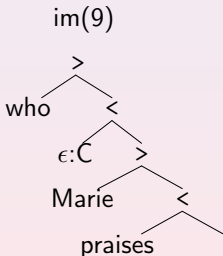
example grammar:

|   |                      |                        |   |
|---|----------------------|------------------------|---|
| 0 | Pierre::D            | who::D -wh             | 4 |
| 1 | Marie::D             | $\epsilon ::= V +wh C$ | 5 |
| 2 | $praises ::= D =D V$ | $know ::= C =D V$      | 6 |
| 3 | $\epsilon ::= V C$   |                        |   |

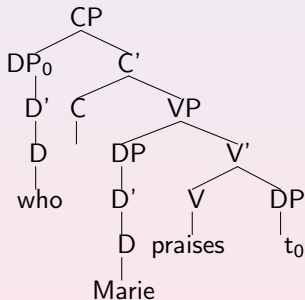
steps 1,2,3



## step 4



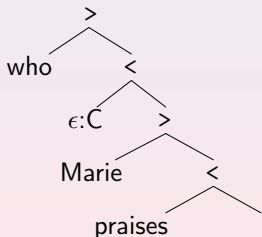
derived tree



conventional

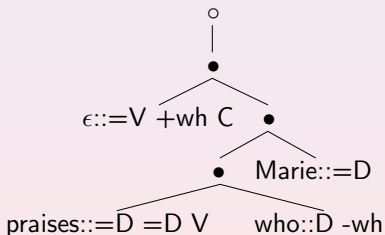
(completed derivation with 1 feature left; 8 features checked in total)

step 4



derived tree

all 4 steps

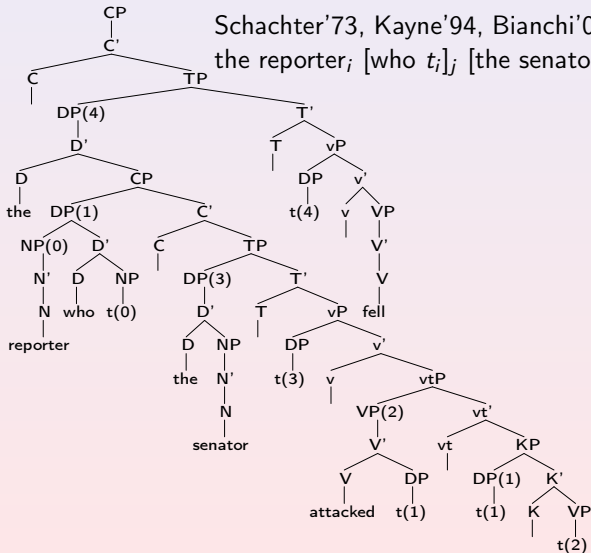


derivation tree

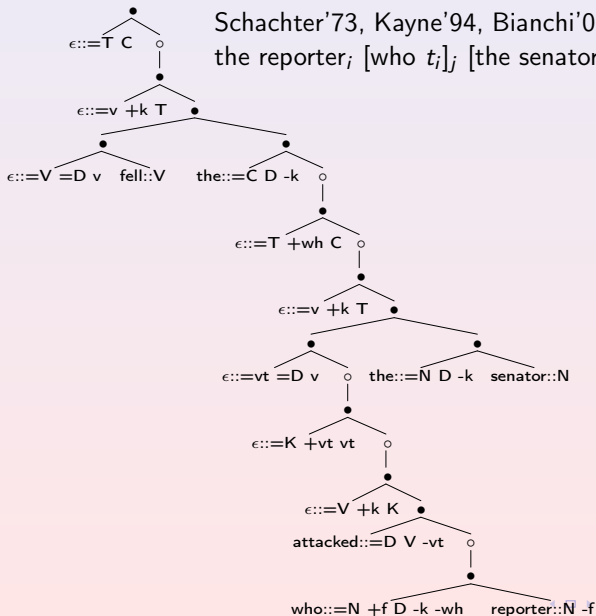


Schachter'73, Kayne'94, Bianchi'00, Bhatt'02, Harris'08:  
the reporter<sub>i</sub> [who t<sub>i</sub>]<sub>j</sub> [the senator]<sub>h</sub> [t<sub>h</sub> attacked t<sub>j</sub>]<sub>k</sub> t<sub>k</sub> fell

(contrast Gibson's structure)



Schachter'73, Kayne'94, Bianchi'00, Bhatt'02, Harris'08:  
 the reporter<sub>i</sub> [who t<sub>j</sub>]<sub>j</sub> [the senator]<sub>h</sub> [t<sub>h</sub> attacked t<sub>j</sub>]<sub>k</sub> t<sub>k</sub> fell



## Notation:

$t[f]$  = tree with 1st feature  $f$  at its head

$t$  = remove  $f$  from  $t[f]$  and change  $::$  to  $:$

$t\{t_1 \mapsto t_2\}$  = the result of replacing subtree  $t_1$  by  $t_2$  in  $t$

$t_1^M$  = the maximal projection of the head of  $t_1$

sometimes we write  $\text{word} : \epsilon$  simply as  $\text{word}$ , and nodes with no features  $\epsilon : \epsilon$  are usually written just  $\epsilon$ , or with no label

$Lex \subseteq (\mathbf{Vocabulary} \times \{::\}) \times \mathbf{Features}$ , a finite set

$$em(t_1[=c], t_2[c]) = \begin{cases} \begin{array}{c} < \\ t_1 \quad t_2 \end{array} & \text{if } t_1 \text{ has exactly 1 node} \\ \begin{array}{c} > \\ t_2 \quad t_1 \end{array} & \text{otherwise} \end{cases}$$

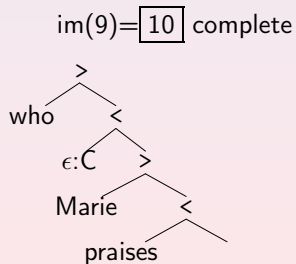
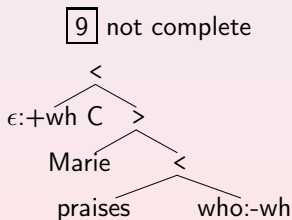
$$im(t_1[+f]) = \begin{array}{c} > \\ t_2^M \quad t_1 \{t_2[-f]^M \mapsto \epsilon\} \end{array} \quad \begin{array}{l} \text{if (SMC) only one head} \\ \text{has } -f \text{ as its first feature} \end{array}$$

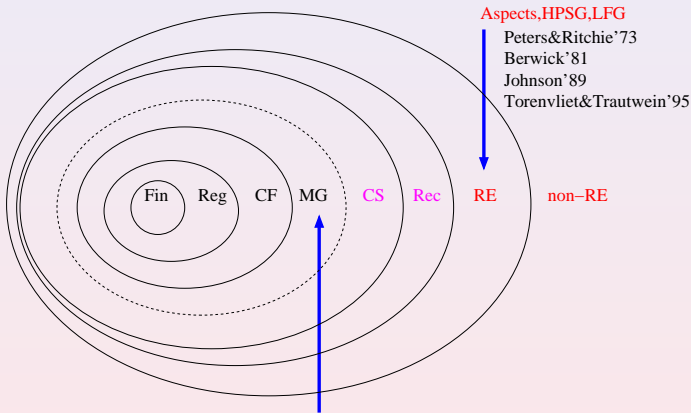
(allows 'surfing' and 'diving' paths!)

**structures**( $G$ ) = closure( $\text{Lex}, \{em, im\}$ )

**completed structures** = trees in structures( $G$ ) with exactly 1 syntactic feature, the “start” category, at its head

**sentences**  $L(G)$  = phonetic yields of completed structures





$$CF \subset \boxed{\text{TAG} \equiv \text{CCG} \equiv \text{2-MCFG} \dots} \subset \boxed{\text{MCFG} \equiv \text{LCFRS} \equiv \text{MG} \dots} \subset CS$$

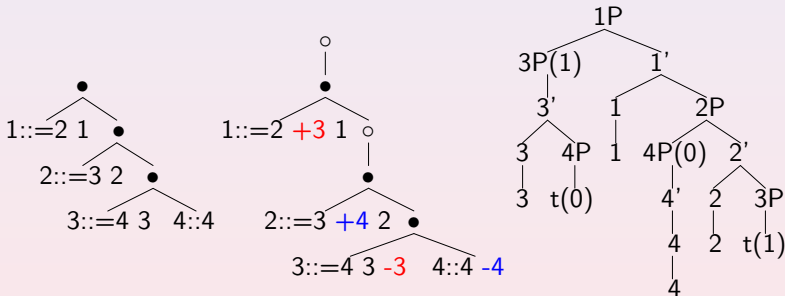
Joshi 1985: Human languages are 'mildly context sensitive'

(linear asym? no!)

(Cinque 1996, 2005): 14/24 [Dem Num Adj N] orders attested

| unattested | 0 |      |   |
|------------|---|------|---|
| attested   | 1 |      |   |
| 1234       | 1 | 1324 | 0 |
| 1243       | 1 | 1342 | 1 |
| 1423       | 1 | 1432 | 1 |
| 4123       | 1 | 4132 | 1 |
| 2134       | 0 | 2314 | 0 |
| 2143       | 0 | 2341 | 1 |
| 2413       | 0 | 2431 | 1 |
| 4213       | 0 | 4231 | 1 |
| 3124       | 0 | 3214 | 0 |
| 3142       | 0 | 3241 | 0 |
| 3412       | 1 | 3421 | 1 |
| 4312       | 1 | 4321 | 1 |

In MGs with just 4 heads selecting each other in the order 1234: 3142, \*2134





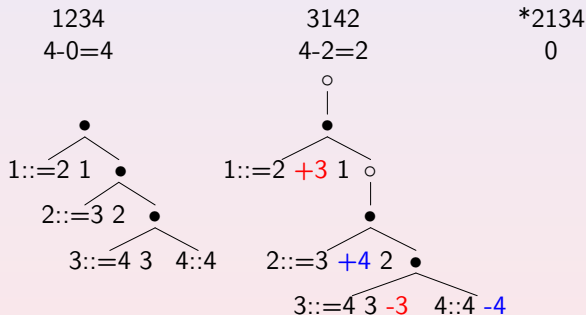
|      | Cinque | MG |      | Cinque | MG |
|------|--------|----|------|--------|----|
| 1234 | 1      | 1  | 1324 | 0      | 0  |
| 1243 | 1      | 1  | 1342 | 1      | 1  |
| 1423 | 1      | 1  | 1432 | 1      | 1  |
| 4123 | 1      | 1  | 4132 | 1      | 1  |
| 2134 | 0      | 0  | 2314 | 0      | 0  |
| 2143 | 0      | 0  | 2341 | 1      | 1  |
| 2413 | 0      | 0  | 2431 | 1      | 1  |
| 4213 | 0      | 0  | 4231 | 1      | 1  |
| 3124 | 0      | 0  | 3214 | 0      | 0  |
| 3142 | 0      | 1  | 3241 | 0      | 0  |
| 3412 | 1      | 1  | 3421 | 1      | 1  |
| 4312 | 1      | 1  | 4321 | 1      | 1  |

(better than a 1 0, but remember that with additional heads, all orders possible)

(Cinque 2005):

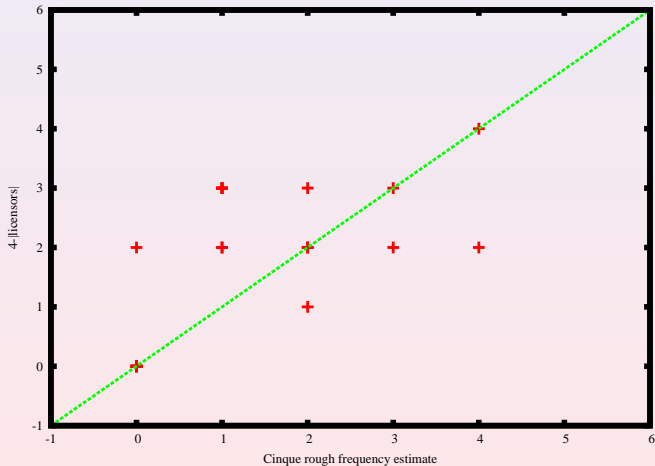
| unattested | 0 | very few | 1 | many      | 3 |
|------------|---|----------|---|-----------|---|
|            |   | few      | 2 | very many | 4 |
|            |   | 1234     | 4 | 1324      | 0 |
|            |   | 1243     | 3 | 1342      | 1 |
|            |   | 1423     | 1 | 1432      | 3 |
|            |   | 4123     | 2 | 4132      | 1 |
|            |   | 2134     | 0 | 2314      | 0 |
|            |   | 2143     | 0 | 2341      | 1 |
|            |   | 2413     | 0 | 2431      | 2 |
|            |   | 4213     | 0 | 4231      | 2 |
|            |   | 3124     | 0 | 3214      | 0 |
|            |   | 3142     | 0 | 3241      | 0 |
|            |   | 3412     | 1 | 3421      | 1 |
|            |   | 4312     | 2 | 4321      | 4 |

MGs: rank structures by  $4 - \mu|\text{licensors}|$ , with 0=impossible



(so then if derivation complexity  $\propto$  frequency, 4=frequent, 0=unattested)

#licensors required correlates with Cinque's frequency estimates



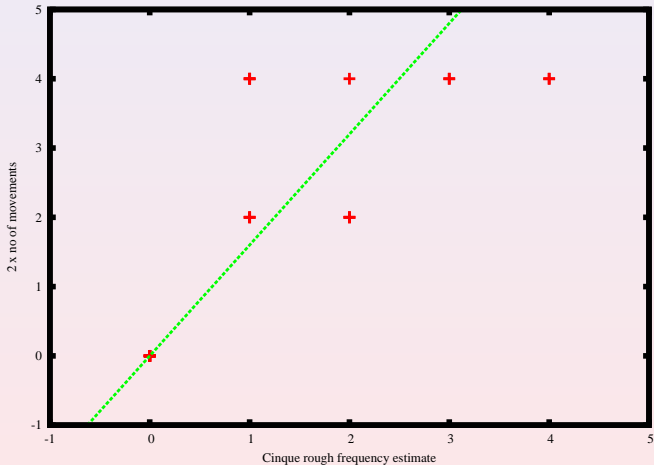
$r=0.68513$

Abels & Neeleman'06: (using  $\prec$  for c-commands)

- a. Underlyingly:  $\text{Dem} \prec \text{Num} \prec \text{A} \prec \text{N}$
  - b. All (relevant) movements move a subtree containing N
  - c. All movements target a c-commanding position
  - d. All movements are to the left
- 
- With free linear order in underlying structure, 8 orders available with no movement, remaining 6 by 1 movement

remnants (at the convergence point)

MGs  
Other assumptions  
Restricting *im*  
cross-serial dependencies



$r=0.80549$

- 0 \* who Pierre knows who \_\_\_ [ \_\_\_ criticizes \_\_\_ ]

SMC provides a 'relativized minimality' effect, but we need an appropriate classification of domains (Rizzi'02).

- 1 \*Combien a-t-il beaucoup consultés \_\_\_ de livres?  
'How many has he a lot consulted of books?'

Criterion freezing (Rizzi'07)

- 2 \*Which candidate does Bill wonder \_\_\_ you voted for \_\_\_ ?

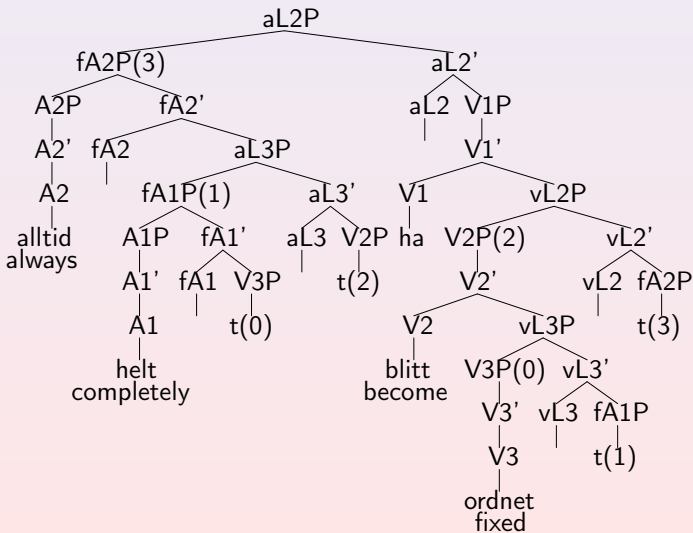
Specifier island condition (Koopman&Szabolcsi'00, Michaelis'01)  
GenPIM (Abels'07) \* $im(t[-x])$  if  $\exists y$  in  $t$  where  $y \ll x$

- 3a Max asked [how likely \_\_\_ to win Oscar was]  
3b \*Oscar was asked [how likely \_\_\_ to win it was]

Remnant movement possible only when  $gap_i$  is  $pro_i$  (Collins&Sabel'07)

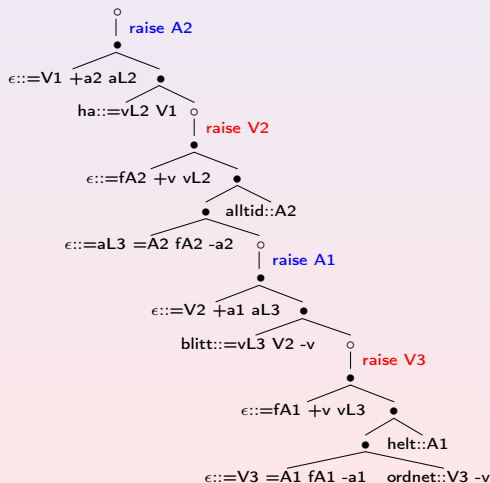
(all these proposals have the simplicity and generality to warrant formal study)

## Cross-serial by remnant movement (Abels'07, Nilsen'03, Bentzen'05):



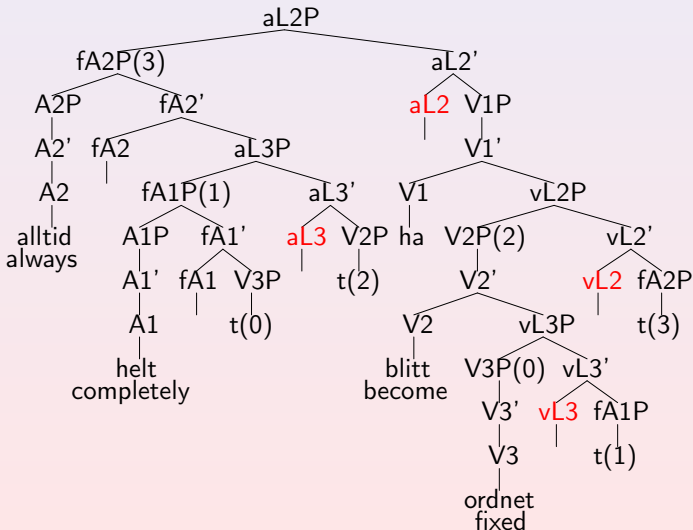


## Cross-serial by remnant movement (Abels'07, Nilsen'03, Bentzen'05):

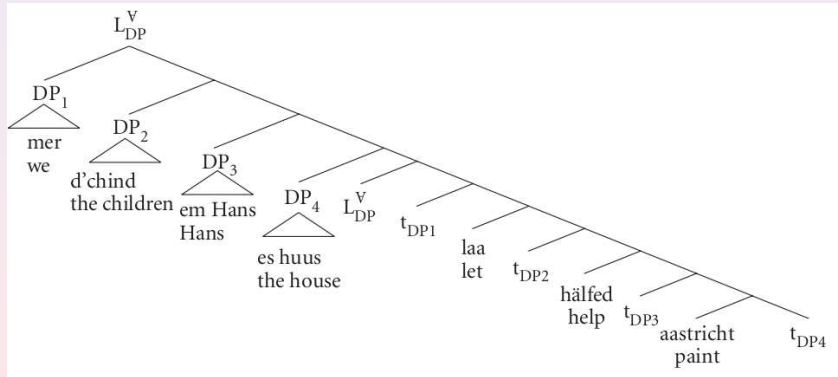


- These interleaved movements cannot be ordered

'Inverse' cross-serial generated, unattested:



## Abels'07 alternative:

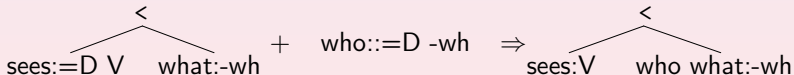


- 'move-all' blocked by SMC; ordering constraint needed

## Bošković, Rudin, et al.:

- 1 Koj kakvo vižda? (Bulgarian)  
who what sees
- 2 Ko šta vidi? (Serbo-Croatian)  
who what sees

absorption:



Bambara (Culy, 1985), Buli (Hiraiwa, 2005), Chinese (Radzinski, 1990; Huang, 1991; Stabler, 2004; Fang, 2006), English (Ghomeshi et al., 2004; Pullum, 2006), Hebrew (Landau, 2006), Italian (Gulli, 2003), Japanese (Potts et al., 1997), Korean (Cho and Nishiyama, 2000) Krio (Nylander, 1985), Vata (Koopman 1983, 1997), Russian (Abels, 2001), Yiddish (Landau, 2007), ...

Yoruba (Kobele '06)

|        |         |      |      |        |          |
|--------|---------|------|------|--------|----------|
| Ri-ra  | adie    | ti   | Jimo | ra     | adie     |
| buying | chicken | rel  | Jimo | buy    | chicken  |
| 'the   | fact    | that | Jimo | bought | chicken' |

Coll. Icelandic (Barbiers '07)

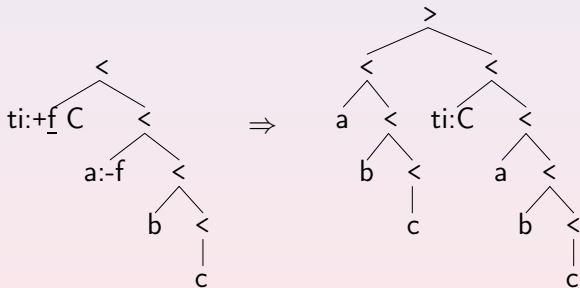
|       |       |      |     |         |         |       |
|-------|-------|------|-----|---------|---------|-------|
| Um    | havð  | eruð | þið | að      | tala    | um?   |
| about | what  | are  | you | to      | talk    | about |
|       | 'What | are  | you | talking | about?' |       |

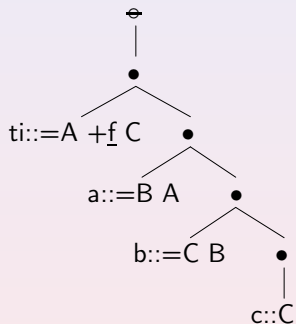
Q4 What kind of grammar makes copying a natural option?

Define an operation which applies to a subtree  $t$ , deleting some of its features to leave  $\neq$ . Then we can extend  $im$ , perhaps conditioned by some property  $+f$ :

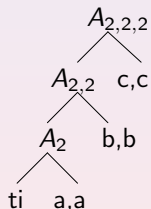
$$im(t_1[+f]) = \begin{array}{c} > \\ \swarrow \quad \searrow \\ t_2^M \quad t_1\{t_2[-f]^M \mapsto t_2[\cancel{-f}]^M\} \end{array}$$

Here, let  $\neq$  leave all and only phonetic features of  $t$ .





copy-move (MGC)



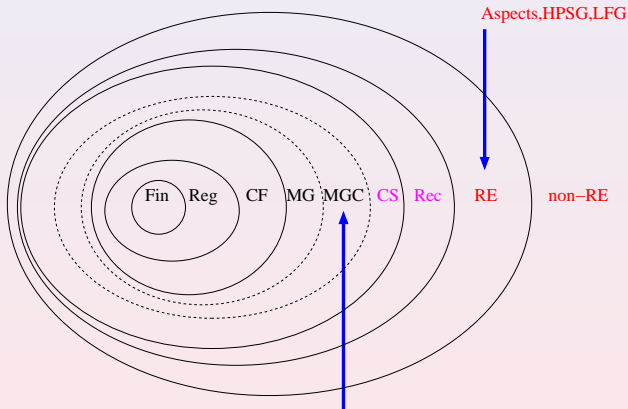
TAG

...

N/A

On N/A, e.g. Pullum'06 "It is the semantics that holds the key..." but then "synonymous lexical items never seem to be synonymous enough" Cf. Stabler'04, Kobele & Stabler'07, Kobele'06, Chen-Main'06, . . .





$$CF \subset \boxed{\text{TAG} \equiv \text{CCG}} \subset \boxed{\text{MCFG} \equiv \text{MG}} \subset \boxed{\text{MGC} \subseteq \text{PMCFG}} \subset \text{CS}$$

- simple formalisms can model many linguistic proposals!
  - CCG, (MC)TAG, MGs converge on a class of MCS languages
  - MGs (defined in  $\approx 5$  lines) empirically threatened at 2 interesting points
    - Removing (SMC) and adding “move all” weakens the theory very considerably, but some version of (RM),(GenPIM)...
    - Adding “copy+move” variants seems required but definitely breaks with convergence
- Q1 What performance models allow incremental interpretation (and remnant movement, doubling constructions)?

$\text{Lex} \subseteq (\text{Vocabulary} \times \{::\} \times \text{Features})$ , a finite set

$$\text{em}(t_1[=c], t_2[c]) = \begin{cases} \begin{array}{c} < \\ t_1 \quad t_2 \end{array} & \text{if } t_1 \text{ has exactly 1 node} \\ \begin{array}{c} > \\ t_2 \quad t_1 \end{array} & \text{otherwise} \end{cases}$$

(replacing these 2 cases with selection on right =c and left c= will not have significant effects)

$$\text{im}(t_1[x]) = \begin{cases} \begin{array}{c} > \\ t_2^M \quad t_1 \{t_2[-f]^M \mapsto \epsilon\} \end{array} & \text{if (SMC) \& } x = +f \\ \begin{array}{c} > \\ t_2^M \quad t_1 \{t_2[-f]^M \mapsto \underline{t_2[-f]^M}\} \end{array} & \text{if (SMC) \& } x = +\underline{f} \end{cases}$$

(replace (SMC) with (RM),(GenPim) etc, but carefully! – cf Gärtner&Michaelis'07)

LCFRS  $\equiv$  MCFG  $\equiv$  MG  $\equiv$  ...

MG+head movement  $\approx$  MG

MG+LF movement  $\approx$  MG

MG+sideways movement  $\approx$  MG

MG+feature percolation, all-powerful

MG+copying  $\not\approx$  MG, but tractable

alternatives to SMC matter!

not clear how to get scrambling

Seki&al.'91; Vijay-Shanker&Weir'94;

Harkema'01; Michaelis'01

Stabler'97,'01; Michaelis'01,'02

Stabler'97; Michaelis'01

Stabler'06,'07

Kobele'05, Kobele&Michaelis'05

Kobele'06

Michaelis'01,'05, Gärtner&Michaelis'07

Rambow'94; Chen-Main&Joshi'08

Abels, Klaus. 2001. The predicate cleft construction in Russian. In Steven Franks and Michael Yadroff, editors, Formal Approaches to Slavic Linguistics 9. Michigan Slavic Publications, Bloomington, Indiana, pages 1–19.

Abels, Klaus. 2007. Towards a restrictive theory of (remnant) movement: Improper movement, remnant movement, and a linear asymmetry. Linguistic Variation Yearbook 2007, 7:53–120.

Abels, Klaus and Ad Neeleman. 2006. Universal 20 without the LCA. Ms., University College, London.

Barbiers, Sjef. 2007. Microvariation in syntactic doubling: An introduction. In Forthcoming.

Berwick, Robert C. 1981. Computational complexity of lexical functional grammar. In Proceedings of the 19th Annual Meeting of the Association for Computational Linguistics, ACL'81, pages 7–12.

Berwick, Robert C. and Amy S. Weinberg. 1984. The Grammatical Basis of Linguistic Performance: Language Use and Acquisition. MIT Press, Cambridge, Massachusetts.

Bhatt, Rajesh. 1999. Adjectival modifiers and the raising analysis of relative clauses. In Proceedings of the North Eastern Linguistic Society, NELS 30.

Bhatt, Rajesh. 2002. The raising analysis of relative clauses: Evidence from adjectival modification. Natural Language Semantics, 10:43–90.

Bianchi, V. 2000. The raising analysis of relative clauses: A reply to Borsley. Linguistic Inquiry, 31:123–140.

Bošković, Željko. 2007a. On multiple feature-checking: multiple wh-fronting and multiple head-movement. forthcoming.

Bošković, Željko. 2007b. On the locality and motivation of move and agree: An even more minimal theory. Linguistic Inquiry, 38(4):589–644.

Chambers, Craig G., Michael K. Tanenhaus, Kathleen M. Eberhard, Hana Filip, and Greg N. Carlson. 2004. Actions and affordances in syntactic ambiguity resolution. Journal of Experimental Psychology: Learning, Memory and Cognition, 30(3):687–696.

Champollion, Lucas. 2007. Lexicalized non-local MCTAG with dominance links is NP-complete. In Mathematics of Language 10, UCLA. <http://kracht.humnet.ucla.edu/marcus/mol10/>.

Chen-Main, Joan. 2006. On the Generation and Linearization of Multi-Dominance Structures. Ph.D. thesis, University of Pennsylvania.

Chen-Main, Joan and Aravind K. Joshi. 2008. Flexible composition, multiple adjoining and word order variation. In Proceedings of the 9th International Workshop on Tree Adjoining Grammars and Related Frameworks, TAG+9.

Chesi, Cristiano. 2004. Phases and Cartography in Linguistic Computation: toward a Cognitively Motivated Computational Model of Linguistic Competence. Ph.D. thesis, University of Siena.

Cho, E. and K. Nishiyama. 2000. Yoruba predicate clefts from a comparative perspective. In V. Carstens and F. Parkinson, editors, Advances in African Linguistics, Volume 4. Trends in African Linguistics. Africa World Press, Trenton, New Jersey.

Chomsky, Noam. 1981. Lectures on Government and Binding. Foris, Dordrecht.

Chomsky, Noam. 1995. The Minimalist Program. MIT Press, Cambridge, Massachusetts.

Chomsky, Noam. 2005. On phases. MIT.

Chomsky, Noam. 2007. Approaching UG from below. In Uli Sauerland and Hans-Martin Gärtner, editors, Interfaces + Recursion = Language? Chomsky's Minimalism and the View from Syntax-Semantics. Mouton de Gruyter, NY, pages 1–30.

Cinque, Guglielmo. 2005. Deriving Greenberg's Universal 20 and its exceptions. Linguistic Inquiry, 36(3):315–332.

- Collins, Chris and Joachim Sabel. 2007. An LF-interface constraint on remnant movement. NYU, to appear.
- Culy, Christopher. 1985. The complexity of the vocabulary of Bambara. Linguistics and Philosophy, 8(3):345–352.
- Fang, Ji. 2006. A Formal Analysis of the Verb Copy Construction in Chinese. Ph.D. thesis, Stanford University, Stanford, California.
- Frey, Werner and Hans-Martin Gärtner. 2002. On the treatment of scrambling and adjunction in minimalist grammars. In Proceedings, Formal Grammar'02, Trento.
- Gärtner, Hans-Martin and Jens Michaelis. 2005. A note on the complexity of constraint interaction. In Logical Aspects of Computational Linguistics, LACL'05, Lecture Notes in Artificial Intelligence LNCS-3492. Springer, NY, pages 114–130.
- Gärtner, Hans-Martin and Jens Michaelis. 2007. A note on the complexity of constraint interaction. In Uli Sauerland and Hans-Martin Gärtner, editors, Interfaces + Recursion = Language? Chomsky's Minimalism and the View from Syntax-Semantics. Mouton de Gruyter, NY, pages 161–196.
- Ghomeshi, Jila, Ray Jackendoff, Nicole Rosen, and Kevin Russell. 2004. Contrastive focus reduplication in English. Natural Language and Linguistic Theory, 22(2):307–357.
- Gibson, Edward. 1998. Linguistic complexity: Locality of syntactic dependencies. Cognition, 68:1–76.
- Greenberg, Joseph. 1966. Some universals of language with particular attention to the order of meaningful elements. In Universals of Human Language. MIT Press, Cambridge, Massachusetts.
- Gulli, A. 2003. Phrasal Reduplication in Syntax. Ph.D. thesis, CUNY.
- Harris, Jesse. 2008. Interpreting raising and matching analyses of relative clauses. In 27th West Coast Conference on Formal Linguistics, WCCFL XXVII.
- Hiraiwa, Ken. 2005. Dimensions of Symmetry in Syntax: Agreement and Clausal Architecture. Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, Massachusetts.
- Huang, Cheng-Teh James. 1991. Modularity and Chinese A-not-A questions. In C. Georgopoulos and R. Ishihara, editors, Interdisciplinary Approaches to Language: Essays in Honor of S.-Y. Kuroda. Kluwer, Dordrecht.

Johnson, Mark. 1988. Attribute Value Logic and The Theory of Grammar. Number 16 in CSLI Lecture Notes Series. CSLI Publications, Chicago.

Joshi, Aravind. 1985. How much context-sensitivity is necessary for characterizing structural descriptions. In D. Dowty, L. Karttunen, and A. Zwicky, editors, Natural Language Processing: Theoretical, Computational and Psychological Perspectives. Cambridge University Press, NY, pages 206–250.

Kallmeyer, Laura. 1999. Tree Description Grammars and Underspecified Representations. Ph.D. thesis, Universität Tübingen.

Kayne, Richard S. 1994. The Antisymmetry of Syntax. MIT Press, Cambridge, Massachusetts.

Kobele, Gregory M. 2005. Features moving madly: A note on the complexity of an extension to MGs. Research on Language and Computation, 3(4):391–410.

Kobele, Gregory M. 2006. Generating Copies: An Investigation into Structural Identity in Language and Grammar. Ph.D. thesis, UCLA.

Kobele, Gregory M. and Jens Michaelis. 2005. Two type 0 variants of minimalist grammars. In Proceedings of the 10th conference on Formal Grammar and the 9th Meeting on Mathematics of Language, FGMOL05.

Kobele, Gregory M. and Edward Stabler. 2007. On copying in language and grammar. LSA Annual Meeting.

Koopman, Hilda. 1983. The Syntax of Verbs: From Verb Movement Rules in the Kru Languages to Universal Grammar. Foris, Dordrecht.

Koopman, Hilda. 1997. Unifying predicate cleft constructions. Talk presented at BLS, February 1997.

Koopman, Hilda and Anna Szabolcsi. 2000. Verbal Complexes. MIT Press, Cambridge, Massachusetts.

Landau, Idan. 2006. Chain resolution in Hebrew v(p)-fronting. Syntax, 9(1):32–66.

Landau, Idan. 2007. Constraints on partial VP-fronting. Syntax, 10(2):127–164.

Michaelis, Jens. 2001a. Observations on strict derivational minimalism. In Proceedings, Formal Grammar and Mathematics of Language.

Michaelis, Jens. 2001b. On Formal Properties of Minimalist Grammars. Ph.D. thesis, Universität Potsdam. *Linguistics in Potsdam 13*, Universitätsbibliothek, Potsdam, Germany.

Michaelis, Jens. 2002. Note on the complexity of complex heads in a minimalist grammar. In Proceedings of the 6th International Workshop on Tree Adjoining Grammars and Related Frameworks, TAG+6.

Michaelis, Jens. 2005. An additional observation on strict derivational minimalism. In Proceedings of the 10th conference on Formal Grammar and the 9th Meeting on Mathematics of Language, FGMOL05.

Michaelis, Jens, Uwe Mönnich, and Frank Morawietz. 2000. Algebraic description of derivational minimalism. In International Conference on Algebraic Methods in Language Processing, AMiLP'2000/TWLT16.

Nilsen, Oystein. 2003. Eliminating Positions: Syntax and Semantics of Sentence Modification. Ph.D. thesis, University of Utrecht.

Nylander, D. K. 1985. Factivity, presupposition, and the relativized predicate in Krio. Studies in African Linguistics, 16(3):323–336.



Obenauer, Hans-Georg. 1983. Une quantification non canonique: la 'quantification à distance'. Langue Française, 48:66–88.

Peters, P. Stanley and R. W. Ritchie. 1973. On the generative power of transformational grammar. Information Sciences, 6:49–83.

Phillips, Colin. 2006. Three benchmarks for distributional approaches to natural language. In Raffaella Zanuttini, Héctor Campos, Elena Herburger, and Paul H. Portner, editors, Cross-Linguistic Research in Syntax and Semantics: Negation, Tense and Clausal Architecture. Georgetown University Press, Washington, DC.

Potts, Christopher, Luis Alonso-Ovalle, Ash Asudeh, Rajesh Bhatt, Seth Cable, Christopher Davis, Yurie Hara, Angelika Kratzer, Eric McCready, Tom Roeper, and Martin Walkow. 1997. Expressiveness and identity conditions. Ms., University of Chicago.

Prince, Alan. 2007. The pursuit of theory. In Cambridge Handbook of Phonology. Cambridge University Press, NY, pages 33–60.

Pullum, Geoffrey K. 2006. On syntactically mandated phrase reduplication. Presented at MIT  



- Radzinski, Daniel. 1990. Unbounded syntactic copying in Mandarin Chinese. Linguistics and Philosophy, 13:113–127.
- Rambow, Owen. 1994. Formal and computational aspects of natural language syntax. Ph.D. thesis, University of Pennsylvania. Computer and Information Science Technical report MS-CIS-94-52 (LINC LAB 278).
- Rizzi, Luigi. 1990. Relativized Minimality. MIT Press, Cambridge, Massachusetts.
- Rizzi, Luigi. 2000. Reconstruction, weak island sensitivity, and agreement. Università di Siena.
- Rizzi, Luigi. 2002. Relativized minimality effects. In Mark Baltin and Chris Collins, editors, The Handbook of Contemporary Syntactic Theory. Blackwell, Oxford, pages 89–110.
- Rizzi, Luigi. 2004. Locality and left periphery. In Adriana Belletti, editor, Structures and Beyond: Cartography of Syntactic Structures, Volume 3. Oxford, NY, pages 104–131.
- Schachter, Paul. 1985. Focus and relativization. Language, 61:523–568.
- Seki, Hiroyuki, Takashi Matsumura, Mamoru Fujii, and Tadao Kasami. 1991. On multiple context-free grammars. Theoretical Computer Science, 88:191–229.
- Stabler, Edward P. 1997. Derivational minimalism. In Christian Retoré, editor, Logical Aspects of Computational Linguistics. Springer-Verlag (Lecture Notes in Computer Science 1328), NY, pages 68–95.
- Stabler, Edward P. 1999. Remnant movement and complexity. In Gosse Bouma, Erhard Hinrichs, Geert-Jan Kruijff, and Dick Oehrle, editors, Constraints and Resources in Natural Language Syntax and Semantics. CSLI Publications, Stanford, California, pages 299–326.
- Stabler, Edward P. 2001. Recognizing head movement. In Philippe de Groote, Glyn Morrill, and Christian Retoré, editors, Logical Aspects of Computational Linguistics, Lecture Notes in Artificial Intelligence, No. 2099. Springer, NY, pages 254–260.
- Stabler, Edward P. 2003. Comparing 3 perspectives on head movement. In A. Mahajan, editor, From Head Movement and Syntactic Theory, UCLA/Potsdam Working Papers in Linguistics. UCLA.
- Stabler, Edward P. 2004. Varieties of crossing dependencies: Structure dependence and mild context sensitivity. Cognitive Science, 93(5):699–720.

- Stabler, Edward P. 2006. Sideways without copying. In Paola Monachesi, Gerald Penn, Giorgio Satta, and Shuly Wintner, editors, Formal Grammar'06, Proceedings of the Conference, pages 133–146, Stanford. CSLI Publications.
- Stabler, Edward P. 2007. Tupled pregroup grammars. In Claudia Casadio and Joachim Lambek, editors, Computational Algebraic Approaches to Morphology and Syntax. Polimetrica, Milan.
- Torenvliet, Leen and Marten Trautwein. 1995. A note on the complexity of restricted attribute-value grammars. In Proceedings of Computational Linguistics In the Netherlands, CLIN5, pages 145–164.
- Weinberg, Amy. 2000. A minimalist theory of human sentence processing. In Samuel David Epstein and Norbert Hornstein, editors, Working Minimalism. MIT Press, Cambridge, Massachusetts.