Patterns and factors in natural systems

E. Stabler, EALING, ENS Paris, 2012

1. Human languages: what they are, how we can study them
   • from the Chomsky hierarchy to linguistic theory
2. Factored grammars and models of language recognition
   • derivation, spellout, agreement
3. Beyond MCS ⇐
   • HLs and birdsong as non-FS, non-CF, non-MCS
Kinds of recursive systems, areas of interest refined

Reg: $A \rightarrow aB, A \rightarrow \epsilon$

CFG: $A \rightarrow X$

TAG: tree adjoining grammar

MG: ck+spellout

CMG: MG with copying

E. Stabler, EALING, ENS Paris, 2012
Patterns and factors in natural systems
First argument:

- $0 \rightarrow \text{john } 1$
- $1 \rightarrow \text{and } 0$
- $1 \rightarrow \text{criticized } 1$
- $2 \rightarrow \text{john } 3$
- $3 \rightarrow \epsilon$
- $3 \rightarrow \text{and } 1$

- $0 \rightarrow \text{mary } 1$
- $1 \rightarrow \text{praised } 1$
- $2 \rightarrow \text{and } 1$
- $3 \rightarrow \text{and } 2$
- $3 \rightarrow \text{and } 0$

Diagram:

- Node 0 with arrows to john, and, and, and.
- Node 1 with arrows to john, criticized, and, and.
- Node 2 with arrows to john, mary, and, and, and.
- Node 3 with arrows to and, and, and.
Review
G(HL) \not\in \text{Reg}
G(HL) \not\in \text{MG+A: copying}
G(HL) \not\in \text{(M)CF: general formulation of move vs. merge}
G(HL) \not\in \text{MG: move and merge vs. agreement}

TP \rightarrow \text{DP VP}
VP \rightarrow \text{V DP}
V \rightarrow \text{criticized}
DP \rightarrow \text{DP D'}
D' \rightarrow \text{and DP}
DP \rightarrow \text{john}
DP \rightarrow \text{mary}

\begin{itemize}
  \item same DP\text{s} in diff positions ('simpler')
  \item new name \Rightarrow both positions
  \item boundary effects (e.g. click)
  \item semantic compositionality
\end{itemize}

Then: \(L(\text{Eng}) \not\subseteq L(\text{Reg})\) not shown by 'mastery of \(A^nB^n\) but by evidence that embedded TP has same structure as matrix TP, and evidence that factors relevant to acceptable depth \(\neq\) factors determining syntax.

E. Stabler, EALING, ENS Paris, 2012
Patterns and factors in natural systems
Second argument:

\[ \epsilon ::= V +\text{wh} C \]

\[ \text{race ::= D } = D V \]

\[ \text{which ::= N } D -\text{wh} \text{ horse ::= N} \]

\[ \langle 0, C \rangle \text{ (which horse they race) } \]

\[ \langle 0, +\text{wh} C, -\text{wh} \rangle \text{ (they race, which horse) } \]

\[ \langle 1, = V +\text{wh} C \rangle (\epsilon) \]

\[ \langle 0, = D V, -\text{wh} \rangle \text{ (race, which horse) } \]

\[ \langle 1, = D = D V \rangle \text{ (race) } \]

\[ \langle 1, = N D -\text{wh} \rangle \text{ (which) } \]

\[ \langle 1, N \rangle \text{ (horse) } \]
• MG treats movement configurations [+fα] . . . [−fβ] . . . alike, but MCFG needs a separate rule for every instance.

• MCFGs miss this important generalization about merge/move configurations, allowing us to prove that MGs can be exponentially smaller than strongly equivalent MCFGs.

Then: \( L(\text{Eng}) \not\subset L(\text{MCF}) \) not shown by ‘mastery’ of ‘X or no X’ constructions, for example (Pullum and Rawlins, 2007; Kobele, 2007a), but by considering best factored (eg. MG + memory restrictions) account.
ck+spellout: simplest perspective on merge/move grammar
(Kobele, Retoré, and Salvati, 2007)

0. form derivation (trivial!)
1. check derivation (at interfaces?) (FS dbutt)
2. map to PF/LF (FS dmbutt)
Third argument (in notes): adding $\phi$-agr to MGs

\[
\epsilon :: V + wh \ C
\]

\[
\text{race-}a_1 - a_2 :: = D = D V
\]

\[
\text{which-}a_1 :: = N \ D - wh \ \text{horse} :: N
\]

Simplest perspective: ck + spellout + agr

Not multiplying out all possibilities, separate probes for each feature (again avoids potentially exp blowup)
Copying and repetition

(Ouattara et al., 2009): “In nonpredation contexts, we recorded three distinct call sequence types, (i) a pair of ‘boom’ calls (B) given alone, (ii) a pair of boom calls followed by a series of krak-oo (Kₚ), and (iii) a pair of boom calls, followed by a series of Kₚ calls, with one to several hok-oo (Hₚ) calls interspersed.” p.22027

Apparently repetition of fixed material, not productive copying? (not clear, given reported geographical variation!)
(Knuth, 1984) on ‘Alouette’,..., ‘That’s the way I like it’

S → that’s the way U I like it U S
U → A A
A → uh huh

obviously we want to do still better, but for present purposes: apparently repetition of fixed material, not productive copying (?)
‘Verbal clefts’ in Vata (Koopman’83)

ngōnū ǹ wà ǹā ǹ kā ngōnū á
sleep you want NA you FUT-A sleep Q
‘Do you want to sleep?’ (p154)

* tākā ǹ wà fòtò mōmū ǹ tákā bò ̀bà
show you like picture ITIT you showed REL Aba
‘It’s show that you like the picture you showed Aba’ (p159)

Many questions: why copying in these sorts of constructions and not others?

Cf. (Lefebvre, 1992; Koopman, 1997; Kobele, 2007b; Johnson, 2010, . . . )
Case/concord:

(one-one) \( T[\text{uCase:N}][\_P \text{he}[\text{Case:N}] \_V[\text{uCase:A}] \text{ sees her}[\text{Case:A}]] \)

(concord)

- der mutmaßliche Täter (German)
  \( \text{the.N presumed.N perpetrator.N} \)
- des mutmaßlichen Täters
  \( \text{the.G presumed.G perpetrator.G} \)

(stacking)

- thabuju-karra-nguni mijil-nguni (Kayardild, Round’10)
  \( \text{brother-G-I net-I} \)
  ‘with brother’s net’

Cf Old Georgian (Michaelis&Kracht’10), and perhaps ‘hidden stacking’ explains default case effects, etc
(Moravcsik’95;Svenonius’05;Richards’07;Matushansky’11;Brattico’11)
What copying is not:

- use of a queue instead of a stack: **too strong**
  
  (see comp.sci. text, or Li & al'93 “The power of the queue”)

- Effect of MG movements: **too weak**
MGs can copy, missing generalizations (Stabler'04; Kobele'06,'07)

\[
\epsilon::T -r -l \quad a::=T +r A -r \quad b::=T +r B -r \\
\epsilon::=T +r +l T \quad a::=A +l T -l \quad b::=B +l T -l
\]

MG copying need a distinct category for every copiable word, not automatically available for new words

E. Stabler, EALING, ENS Paris, 2012

Patterns and factors in natural systems
Proposal: \( ck + \text{spellout(with copy)} \)

\[
\begin{align*}
\text{nguni} &::= N \text{Ins} \\
\text{karra} &::= N \text{Gen} \\
\text{net} &::= N
\end{align*}
\]

\[
\begin{align*}
\text{brother} &::= N
\end{align*}
\]

\( \Rightarrow \)

\[
\begin{align*}
\text{InsP} &\quad \text{N} \\
\text{GenP} &\quad \epsilon \\
\text{Gen} &\quad \text{N} \\
\epsilon &\quad \text{N} \\
\text{N} &\quad \text{Ins}
\end{align*}
\]

\[
\begin{align*}
\text{net} &\quad \text{NGen} \\
\text{nguni} &\quad \text{N}
\end{align*}
\]

\[
\begin{align*}
\text{brother} &\quad \text{N}
\end{align*}
\]

\[
\begin{align*}
\text{karra} &\quad \text{N}
\end{align*}
\]

\( \triangle \) signals “complete” spellout = adjoin to all overt heads in complement. Cf. affix hopping

E. Stabler, EALING, ENS Paris, 2012 Patterns and factors in natural systems
Birdsong:

- Finch, formal models
- California Thrasher, formal models

(cf Stabler, Taylor, Cody, forthcoming...)


Kobele, Gregory M. 2007a. Argument! a reply to Pullum and Rawlins. UCLA manuscript.


Pesetsky, David. 2010. Russian case morphology and the syntactic categories. MIT.


Richards, Norvin. 2007. Lardil case stacking and the structural/inherent case distinction. [http://ling.auf.net/LingBuzz/000405](http://ling.auf.net/LingBuzz/000405).
