Natural paths in MCFGs

Edward Stabler

MCFG+, Tokyo, 2010
The interest of MCFLs

- MCFLs are defined by many independent formalisms
- MCFLs are useful for describing linguistic, biological structures

(EQ1) $\text{HL} \subseteq \text{MCFL}$?

Kracht & Michaelis, Kobele: No. (Old Georgian case; Yoruba clefts)

(EQ2) $\text{HL} \subseteq \text{MCFL}_{\text{wn}}$?

Joshi: Yes. $\text{HL} \subseteq \text{TAL}=2-\text{MCFL}_{\text{wn}}$

(EQ3) ‘Semantically appropriate’ $\text{HG} \subseteq \text{MCFG}$?

Rambow: No. (German scrambling)

These are matters of current (useful!) controversy.
Natural paths in MCFGs

- **ML = MCFL** (Michaelis’01, Harkema’01, Seki & al’91)
- In proving $\subseteq$ Michaelis’98 had already revealed a ‘strong’ equivalence.

0. **MGs provide a succinct notation for ‘strongly equivalent’ MCFGs**
1. ‘Order universals’ derivable from fixed categories and selection features
2. ‘Improper movements’ banned by fixed order of licensee features
3. In MCFG these are path restrictions with linear order consequences;
   - allows even more succinct grammars,
   - limits expressive power,
   - ‘near’ the range of Yoshinaka & Clark’s learner.
MGs

\[ \text{MG} = \langle \Sigma, \text{Cat}, \text{Ep}, \text{Lex}, M, S \rangle \text{ where} \]

\[
\begin{align*}
\Sigma &= \{ \text{John, Mary, who, criticize, praise, -s, -ed, \ldots} \} \quad \text{(vocabulary)} \\
\text{Cat} &= \{ \text{N, V, A, P, \ldots} \} \quad \text{(categories)} \\
\text{Sel} &= \{ =f \mid f \in \text{Cat} \} \quad \text{(selectors)} \\
\text{Ep} &= \{ +\text{case}, +\text{wh}, +\text{q}, +\text{foc}, +\text{top}, \ldots \} \quad \text{(licensors)} \\
\text{Lic} &= \{ -f \mid +f \in \text{Ep} \} \quad \text{(licensees)} \\
F &= \text{Cat} \cup \text{Sel} \cup \text{Ep} \cup \text{Lic} \\
\text{Lex} &\subseteq \Sigma^{\epsilon} \times F^*, \text{ finite} \quad \text{(lexicon)} \\
M &= \text{merge rules.} \ldots \\
S &= \epsilon \text{ Cat} \quad \text{(start)}
\end{align*}
\]

\(\Sigma, \text{Cat}, \text{Sel}, \text{Ep}, \text{Lic}, \text{Lex} \text{ finite, non-empty, pairwise disjoint.}\)
‘naive Zapotec’ VSO: *praised the students the idea*

<table>
<thead>
<tr>
<th>the</th>
<th>=N D -ep</th>
</tr>
</thead>
<tbody>
<tr>
<td>students</td>
<td>N</td>
</tr>
<tr>
<td>idea</td>
<td>N</td>
</tr>
<tr>
<td>praised</td>
<td>=D V -v</td>
</tr>
<tr>
<td>ε</td>
<td>=V +ep =D v</td>
</tr>
<tr>
<td>ε</td>
<td>=v +ep T</td>
</tr>
<tr>
<td>ε</td>
<td>=T +v C</td>
</tr>
</tbody>
</table>

Each MG names a ‘strongly equivalent’ $k$-MCFG, $k = |Ep| + 1$. 
Given $MG = \langle \Sigma, \text{Cat}, \text{Ep}, \text{Lex}, M, S \rangle$, we define an MCFG with two start categories $MG = \langle \Sigma, N, P, \{\langle 0, S \rangle, \langle 1, S \rangle \} \rangle$, defining the language

$$N = \{ \langle x, \delta_0, \delta_1, \ldots, \delta_j \rangle \mid x \in \{0, 1\}, \quad 0 \leq j \leq |\text{Ep}|, \quad \text{all } \delta_i \in \text{suffix}(\pi_2(\text{Lex})) \},$$

where each nonterminal $\langle x, \delta_0, \ldots, \delta_j \rangle$ has rank $j + 1$. 
MGs as MCFGs

For $0 \leq i, j \leq |Ep|$, $\beta \neq \epsilon$, $x, y \in \{0, 1\}$:

**lex:** $\langle 1, \alpha \rangle(s)$

**em1:** $\langle 0, \alpha, \delta_1, \ldots, \delta_j \rangle(s_0 t_0, t_1, \ldots, t_j)$:

$\langle 1, =f \alpha \rangle(s_0)$,

$\langle x, f, \delta_1, \ldots, \delta_j \rangle(t_0, \ldots, t_j)$

**em2:** $\langle 0, \alpha, \delta_1, \ldots, \delta_i, \gamma_1, \ldots, \gamma_j \rangle(t_0 s_0, s_1, \ldots, s_i, t_1, \ldots, t_j)$:

$\langle 0, =f \alpha, \delta_1, \ldots, \delta_i, \rangle(s_0, \ldots, s_i)$,

$\langle x, f, \gamma_1, \ldots, \gamma_j \rangle(t_0, \ldots, t_j)$

**em3:** $\langle 0, \alpha, \beta, \delta_1, \ldots, \delta_i, \gamma_1, \ldots, \gamma_j \rangle(t_0 s_0, s_1, \ldots, s_i, t_1, \ldots, t_j)$:

$\langle x, =f \alpha, \delta_1, \ldots, \delta_i, \rangle(s_0, \ldots, s_i)$,

$\langle y, f \beta, \gamma_1, \ldots, \gamma_j \rangle(t_0, \ldots, t_j)$

**im1:** $\langle 0, \alpha, \delta_1, \ldots, \delta_{i-1}, \delta_{i+1}, \ldots, \delta_j \rangle(s_i s_0, s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_j)$:

$\langle 0, +f \alpha, \delta_1, \ldots, \delta_j \rangle(s_0, \ldots, s_j)$

**im2:** $\langle 0, \alpha, \delta_1, \ldots, \delta_{i-1}, \beta, \delta_{i+1}, \ldots, \delta_j \rangle(s_0, \ldots, s_i)$:

$\langle 0, +f \alpha, \delta_1, \ldots, \delta_j \rangle(s_0, \ldots, s_i)$

SMC: $\delta_1, \ldots, \delta_{i-1}, \delta_{i+1}, \ldots, \delta_j$ do not begin with $-f$. 

Edward Stabler (MCFG+, Tokyo, 2010)  
Natural paths in MCFGs  
7 / 42
‘naive Zapotec’ VSO: praised the students the idea

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>the</td>
<td>=N D -ep</td>
<td></td>
</tr>
<tr>
<td>students</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>idea</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>praised</td>
<td>=D V -v</td>
<td></td>
</tr>
<tr>
<td>ϵ</td>
<td>=V +ep =D v</td>
<td></td>
</tr>
<tr>
<td>ϵ</td>
<td>=v +ep T</td>
<td></td>
</tr>
<tr>
<td>ϵ</td>
<td>=T +v C</td>
<td></td>
</tr>
</tbody>
</table>
‘naive Zapotec’ VSO: praised the students the idea

\[
\langle 0, C \rangle \text{(praised the students the idea)} \\
\langle 0, +vC, -v \rangle \text{(the students the idea, praised)} \\
\langle 1, =T+vC \rangle(\epsilon) \quad \langle 0, T, -v \rangle \text{(the students the idea, praised)} \\
\langle 0, +epT, -v, -ep \rangle \text{(the idea, praised, the students)} \\
\langle 1, =v+epT \rangle(\epsilon) \quad \langle 0, v, -v, -ep \rangle \text{(the idea, praised, the students)} \\
\langle 0, =Dv, -v \rangle \text{(the idea, praised)} \quad \langle 0, D-ep \rangle \text{(the students)} \\
\langle 0, +ep=Dv, -v, -ep \rangle(\epsilon, \text{praised, the idea}) \quad \langle 1, =ND-ep \rangle \text{(the)} \quad \langle 1, N \rangle \text{(students)} \\
\langle 1, =V+ep=Dv \rangle(\epsilon) \quad \langle 0, V-v, -ep \rangle \text{(praised, the idea)} \\
\langle 1, =DV-v \rangle \text{(praised)} \quad \langle 0, D-ep \rangle \text{(the idea)} \\
\langle 1, =ND-ep \rangle \text{(the)} \quad \langle 1, N \rangle \text{(idea)}
\]
‘naive Zapotec’ VSO: praised the students the idea
‘naive Zapotec’ VSO: praised the students the idea
‘naive Tamil’ SOV: the students the idea praised

```
the = N D -ep
students = N
idea = N
praised = D V
ε = V + ep = D V
ε = v + ep T
ε = T C
```

```
ε ::= T C
ε ::= v + ep T
ε ::= V + ep = D V
praised ::= D V
the ::= N D - ep
students ::= N
```

Edward Stabler (MCFG+, Tokyo, 2010)
Natural paths in MCFGs
‘naive Tamil’ SOV: *the students the idea praised*
‘naive English’ SVO: *the students praised the idea*

- the = N D
- the = N D - ep
- students = N
- idea = N
- praised = D V
- ε = V = D v
- ε = v + ep T
- ε = T C
- which = N D - wh
- which = N D - ep - wh
- teachers = N
- ε = T + wh C
- knew = D V

Diagram:

```
ε ::= T C
  /       \
 /         \
ε ::= v + ep T
  |         |
  +---------+
     ε ::= V = D v
  |         |
  |         |
ε ::= = D V
  |         |
  |         |
  |         |
  +---------+
    ε ::= N D - ep
    |         |
    |         |
    +---------+
      ε ::= N D
      |         |
      |         |
      +---------+
        idea ::= N
        |         |
        |         |
        +---------+
          students ::= N
```

Edward Stabler (MCFG+, Tokyo, 2010)
“naive English” SVO: the students praised the idea
‘naive English’ SVO

\[
\langle \text{the teachers knew which idea the students praised} \rangle : \langle 0, C \rangle
\]
‘naive English’ SVO MCFG (page 1)

\[
\begin{align*}
\langle 0, C \rangle(s_0 t_0) :&= \langle 1, = T \ C \rangle(s_0), \langle 0, T \rangle(t_0) \\
\langle 1, = T \ C \rangle(\epsilon) :&= \langle 0, + ep \ T, - ep \rangle(s_0, s_1) \\
\langle 0, + ep \ T, - ep \rangle(s_0 t_0, t_1) :&= \langle 1, = v + ep \ T \rangle(s_0), \langle 0, v, - ep \rangle(t_0, t_1) \\
\langle 1, = v + ep \ T \rangle(\epsilon) :&= \langle 0, + D \ v \rangle(s_0), \langle 0, D, - ep \rangle(t_0) \\
\langle 0, + D \ v \rangle(s_0 t_0) :&= \langle 1, = V + D \ v \rangle(s_0), \langle 0, V \rangle(t_0) \\
\langle 1, = V + D \ v \rangle(\epsilon) :&= \langle 0, + T \ C \rangle(s_0) \\
\langle 0, + T \ C \rangle(\epsilon) :&= \langle 1, = C \ V \rangle(s_0), \langle 0, C \rangle(t_0) \\
\langle 1, = C \ V \rangle(knew) &:= \langle 0, + wh \ C, - wh \rangle(s_0, s_1) \\
\langle 0, + wh \ C, - wh \rangle(s_0 t_0, t_1) :&= \langle 1, = V + wh \ C \rangle(s_0) \langle 0, V, - wh \rangle(t_0, t_1) \\
\langle 1, = V + wh \ C \rangle(\epsilon) :&= \langle 0, + wh \ C \rangle(t_0, t_1) \\
\langle 0, + wh \ C \rangle(s_0 t_0, t_1) :&= \langle 1, = C \ V \rangle(s_0), \langle 0, C, - wh \rangle(t_0, t_1) \\
\langle 0, C, - wh \rangle(s_0 t_0, t_1) :&= - em1[1, = T \ C \rangle(s_0), \langle 0, T, - wh \rangle(t_0, t_1)
\end{align*}
\]

\[
\begin{align*}
\text{the} &= N \ D \\
\text{the} &= N \ D, - ep \\
\text{students} &= N \\
\text{idea} &= N \\
\text{praised} &= D \ V \\
\epsilon &= V = D \ v \\
\epsilon &= v + ep \ T \\
\epsilon &= T \ C \\
\text{which} &= N \ D, - wh \\
\text{which} &= N \ D, - ep, - wh \\
\text{teachers} &= N \\
\epsilon &= T + wh \ C \\
\text{knew} &= C \ V
\end{align*}
\]
\[ \langle 0, T, \text{-wh} \rangle (s_2 s_0, s_1) \quad \text{:-} \quad \langle 0, \text{+ep T}, \text{-wh}, \text{-ep} \rangle (s_0, s_1, s_2) \]
\[ \langle 0, \text{+ep T}, \text{-wh}, \text{-ep} \rangle (s_0 t_0, t_1, t_2) \quad \text{:-} \quad \langle 1, \text{=v +ep T} \rangle (s_0), \langle 0, \text{v, -wh}, \text{-ep} \rangle (t_0, t_1, t_2) \]
\[ \langle 0, \text{v, -wh}, \text{-ep} \rangle (s_0, s_1, t_0) \quad \text{:-} \quad \langle 0, \text{=D v, -wh} \rangle (s_0, s_1), \langle 0, \text{D -ep} \rangle (t_0) \]
\[ \langle 0, \text{=D v, -wh} \rangle (s_0 t_0, t_1) \quad \text{:-} \quad \langle 1, \text{=V =D v} \rangle (s_0), \langle 0, \text{V, -wh} \rangle (t_0, t_1) \]
\[ \langle 0, \text{V, -wh} \rangle (s_0, t_0) \quad \text{:-} \quad \langle 1, \text{=D V} \rangle (s_0), \langle 0, \text{D -wh} \rangle (t_0) \]
\[ \langle 1, \text{=D V} \rangle \text{(praised)} \]
\[ \langle 0, \text{D -wh} \rangle (s_0 t_0) \quad \text{:-} \quad \langle 1, \text{=N D -wh} \rangle (s_0), \langle 1, \text{N} \rangle (t_0) \]
\[ \langle 1, \text{=N D -wh} \rangle \text{(which)} \]
\[ \langle 1, \text{N} \rangle \text{(teachers)} \]
\[ \langle 1, \text{N} \rangle \text{(idea)} \]
\[ \langle 1, \text{N} \rangle \text{(students)} \]
\[ \langle 0, \text{D -ep} \rangle (s_0 t_0) \quad \text{:-} \quad \langle 1, \text{=N D -ep} \rangle (s_0), \langle 1, \text{N} \rangle (t_0) \]
\[ \langle 1, \text{=N D -ep} \rangle \text{(the)} \]
\[ \langle 0, \text{T, -wh} \rangle (s_1 s_0, s_2) \quad \text{:-} \quad \langle 0, \text{+ep T, -ep, -wh} \rangle (s_0, s_1, s_2) \]
\[ \langle 0, \text{+ep T, -ep, -wh} \rangle (s_0 t_0, t_1, t_2) \quad \text{:-} \quad \langle 1, \text{=v +ep T} \rangle (s_0), \langle 0, \text{v, -ep, -wh} \rangle (t_0, t_1, t_2) \]
Edward Stabler (MCFG+, Tokyo, 2010)
Constituent order

- MG builds spec-head-comp, which can then be distorted by movements.

\[ =D =D V \text{ is analagous to } (D \backslash V)/D. \]

MG has no analog of changing slash direction. Alternative orders by movement to the left, introducing asymmetries...

- Cinque’05, Greenberg’63 Universal 20: not all orders of [Dem Num Adj N] are attested

- Given [1 [2 [3 [4]]]], what orders by adding ‘licensing’?
**Constituent order**

<table>
<thead>
<tr>
<th>order</th>
<th>C</th>
<th>MG</th>
<th>order</th>
<th>C</th>
<th>MG</th>
</tr>
</thead>
<tbody>
<tr>
<td>1234</td>
<td>4</td>
<td>4</td>
<td>1324</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1243</td>
<td>3</td>
<td>3</td>
<td>1342</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>1423</td>
<td>1</td>
<td>3</td>
<td>1432</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>4123</td>
<td>2</td>
<td>3</td>
<td>4132</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2134</td>
<td>0</td>
<td>0</td>
<td>2314</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2143</td>
<td>0</td>
<td>0</td>
<td>2341</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>2413</td>
<td>0</td>
<td>0</td>
<td>2431</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>4213</td>
<td>0</td>
<td>0</td>
<td>4231</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3124</td>
<td>0</td>
<td>0</td>
<td>3214</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3142</td>
<td>0</td>
<td>2</td>
<td>3241</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3412</td>
<td>1</td>
<td>3</td>
<td>3421</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>4312</td>
<td>2</td>
<td>2</td>
<td>4321</td>
<td>4</td>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cinque</th>
</tr>
</thead>
<tbody>
<tr>
<td>0=</td>
</tr>
<tr>
<td>1=</td>
</tr>
<tr>
<td>2=</td>
</tr>
<tr>
<td>3=</td>
</tr>
<tr>
<td>4=</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MG</th>
</tr>
</thead>
<tbody>
<tr>
<td>unattested</td>
</tr>
<tr>
<td>very few</td>
</tr>
<tr>
<td>few</td>
</tr>
<tr>
<td>many</td>
</tr>
<tr>
<td>very many</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>4 licensors</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 licensors</td>
</tr>
<tr>
<td>2 licensors</td>
</tr>
<tr>
<td>0-1 licensors</td>
</tr>
</tbody>
</table>
## Constituent order

<table>
<thead>
<tr>
<th>order</th>
<th>C</th>
<th>MG</th>
</tr>
</thead>
<tbody>
<tr>
<td>1234</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>1243</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>1423</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>4123</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>2134</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2143</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2413</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4213</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3124</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3142</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>3412</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>4312</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>1324</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1342</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>1432</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>4132</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2314</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2341</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>2431</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>4231</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3214</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3241</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3421</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>4321</td>
<td>4</td>
<td>2</td>
</tr>
</tbody>
</table>

**Cinque MG**

0 = unattested
1 = very few
2 = few
3 = many
4 = very many

MG

underivable

4 licensors
3 licensors
2 licensors
0-1 licensors
Constituent order

- MG hypothesis explains U20 only with fixed selection
  \[ \text{Dem} < \text{Num} < \text{Adj} < \text{N} \]
  MG allows no change in slash direction; now, no change in selection..
  (Makes sense only with independent properties. 'Flexible' accounts similar.)

- Our VSO, SOV, VSO examples:
  \[ \text{C} < \text{T} < \nu < \text{V} < \text{D} < \text{N} \]

- Rizzi’04 ‘left periphery’
  \[ \text{Force} < \text{Top1} < \text{Foc} < \text{Top2} < \text{Fin} < \text{Infl} \]

- Cinque’99 adverbials
  \[ \text{Mood} < \text{Evidential} < \text{Epistemic} < \text{Habitual} < \text{Inceptive} \]
  frankly  allegedly  probably  usually  suddenly

- Manzini&Savoia’04 clitic positions
  \[ \text{Def(uninfl)} < \text{Quant}(3) < \text{N(3pl)} < \text{P(1,2)} < \text{Origin} < \text{Loc} < \text{Measure} \]
Constituent order and paths

- **(Cartographic hypothesis)** (informal)
  Categories and order of selection in human languages is universal.
  ‘Clausal hierarchy is fixed.’

- ‘C-selection’ looks like it might be semantically motivated.
  Could ‘c-selection’ be completely reduced to ‘s-selection’?
  Chomsky’95 speculates: “there is a syntactic residue” (p.33)

- For any CFG, path language regular (Thatcher’67)
Paths in naive English

Naive English is slightly complicated, so consider this ‘tiny English’...
Tiny English

Edward Stabler (MCFG+, Tokyo, 2010)
Path sets strictly 2-local (2-SL), as for any CFG

- Last symbol a terminal category, which never dominates anything.
- ‘spine”: at binary branches, the right sister; otherwise left sister.
Spinal paths
Spinal paths

tiny English: C, V, D:

Naive English: C, T, v, V, D, N:
• (Cartographic hypothesis CH1) UG fixes $\text{Cat}_\leq$ so that
  
  If $fg$ appears in a spinal path
  and $f \neq g$, then either $f$ covers $\leq g$ or $g$ minimal $\leq$.

  If $fg$ appears in a non-spinal path initiated at a spec position, and
  $f \neq g$, then either $f$ covers $\leq g$ or $g$ minimal $\leq$. 

Edward Stabler (MCFG+, Tokyo, 2010)
Movement order universals

- Naive English *which idea* moves twice, with –ep –wh
  
  the teachers knew which idea; the students ti praised ti

(BOIM) Other orders are ‘improper’, yielding ungrammaticality

  * Who; seems ti will ti leave?
  * Who; seems it is likely ti to ti leave?

- ep ◁ scrambling ◁ wh ◁ top  (Abels’07)
Movement order universals: naive English –wh and –ep–wh

MGs would allow –ep to –wh–ep equally easily, but that never happens!
Path restrictions 2

(Cartographic hypothesis CH2) Movement features $E_p \triangleleft$ are ordered so that if a component moves for $g$ and then for $f$, $f \triangleleft g$

MG allows Lex to generate paths violating these conditions, but we can block that...
Ordered MGs

\[ OMG = \langle \Sigma, \text{Cat}_\leq, \text{Ep}_\triangleleft, Df, \mu, \text{Lex}, M, S \rangle \]

\( \Sigma = \{ \text{John, Mary, who, criticize, praise, -s, -ed, } \ldots \} \) (vocabulary)
\( \text{Cat} = \langle \{ N, V, A, P, \ldots \}, \leq \rangle \) (categories)
\( \text{Ep} = \langle \{ \text{case, wh, q, foc, top, } \ldots \}, \triangleleft \rangle \) (licensors)
\( Df : \text{Cat} \to \text{Cat} \cup \text{Ep} \) (deficiencies)
\( \mu = \) meanings, e.g. \( \text{TH}(E, 2, S) \) (extends to \([\cdot]\))
\( \text{Lex} \subseteq \Sigma^e \times \text{Cat} \times \wp(\text{Ep}) \times \mu \) (lexicon)
\( M = \text{merge, the union of EM and IM rules} \)
\( S \in \text{Cat} \) (start)

- \( a \) selects \( b \) iff \( a \) covers\( \leq \) \( b \) or \( b \) minimal\( \triangleleft \)
- Semantic type may restrict selection and deficiency
- \( Df \) maps each \( \text{Cat} \) to at most one value – at most one specifier.

Many linguists assume \( UG = \langle \text{Cat}_\leq, \text{Ep}_\triangleleft, Df, M, S \rangle \); all variation in \( \text{Lex} \).
Ordered MGs

<table>
<thead>
<tr>
<th>the</th>
<th>=N D</th>
</tr>
</thead>
<tbody>
<tr>
<td>the</td>
<td>=N D -ep</td>
</tr>
<tr>
<td>students</td>
<td>N</td>
</tr>
<tr>
<td>idea</td>
<td>N</td>
</tr>
<tr>
<td>praised</td>
<td>=D V</td>
</tr>
<tr>
<td>$\epsilon$</td>
<td>=V =D $\epsilon$</td>
</tr>
<tr>
<td>$\epsilon$</td>
<td>=v +ep T</td>
</tr>
<tr>
<td>$\epsilon$</td>
<td>=T C</td>
</tr>
<tr>
<td>which</td>
<td>=N D -wh</td>
</tr>
<tr>
<td>which</td>
<td>=N D -ep -wh</td>
</tr>
<tr>
<td>teachers</td>
<td>N</td>
</tr>
<tr>
<td>$\epsilon$</td>
<td>=T +wh C</td>
</tr>
<tr>
<td>knew</td>
<td>=C V</td>
</tr>
</tbody>
</table>

Df: $C \mapsto \text{wh}, \ T \mapsto \text{ep}, \ v \mapsto \text{ep}$

<table>
<thead>
<tr>
<th>the</th>
<th>=N D</th>
</tr>
</thead>
<tbody>
<tr>
<td>the</td>
<td>=N D {ep}</td>
</tr>
<tr>
<td>students</td>
<td>N</td>
</tr>
<tr>
<td>idea</td>
<td>N</td>
</tr>
<tr>
<td>praised</td>
<td>V</td>
</tr>
<tr>
<td>$\epsilon$</td>
<td>v</td>
</tr>
<tr>
<td>$\epsilon$</td>
<td>T</td>
</tr>
<tr>
<td>$\epsilon$</td>
<td>C</td>
</tr>
<tr>
<td>which</td>
<td>=D {wh}</td>
</tr>
<tr>
<td>which</td>
<td>=D {ep,wh}</td>
</tr>
<tr>
<td>teachers</td>
<td>N</td>
</tr>
<tr>
<td>$\epsilon$</td>
<td>C</td>
</tr>
<tr>
<td>knew</td>
<td>V</td>
</tr>
</tbody>
</table>
Ordered MGs

Expressive power

- \( \forall \text{UG}, \text{OML}_{\text{UG}} \subsetneq (|\text{Ep}|+1)\text{-MCFL}(2) \subsetneq \text{MCFL} \)
  (Seki et al’91; Rambow & Satta’99)

Learnability

- \( \forall \text{UG}, \text{OML}_{\text{UG}} \) finite, hence identifiable from positive text
- \( p \)-congruential MCFLs are identifiable from membership queries (yes or no to \( x \in L_* \)) and equivalence queries (yes or counterexample to \( L = L_* \)).
  (Yoshinaka & Clark’10)

\[ G \in p\text{-MCFG}(r) \text{ is } p\text{-congruential iff} \]
(i) string functions linear and non-permuting, and
(ii) for every nonterminal \( A \), and any \( u, v \in L(G, A) \), \( L(G)/u = L(G)/v \).
Non-permuting OMGs

MGs, OMGs have this permuting rule, when $\delta_i = -f$, SMC:

$$im1: \langle 0, \alpha, \delta_1, \ldots, \delta_{i-1}, \delta_{i+1}, \ldots, \delta_j \rangle (s_is_0, s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_j) :-$$

$$\langle 0, +f \alpha, \delta_0, \ldots, \delta_j \rangle (s_0, \ldots, s_j)$$

? Can we change this to something like:

$$im1??: \langle 0, \delta_1, \ldots, \delta_{i-1}, \alpha \rangle (s_1, \ldots, s_{i-1}, si, s_0) :-$$

$$\langle 0, \delta_1, \ldots, \delta_i, +f \alpha \rangle (s_1, \ldots, s_i, s0)$$

Yes, for certain OMGs we can formulate (strongly equivalent) non-permuting rules. E.g. suppose $\leq, \triangleleft$ total, Df total and 1-1...
Non-permuting OMGs

To form a constituent of category $f$ with the components

$$\alpha f \beta, \delta_1, \ldots, \delta_i, \gamma_1, \ldots, \gamma_j.$$ 

Is it possible to determine the surface order of these elements? Yes.

Suppose $Cat = \{1, \ldots, n\}$ where $\forall c \in cat, c \mapsto +c$ and $\beta = -i-j-k$:

If $f \leq i$, then $-i$ checked at $i$, $(i - f)$ steps away.
Then if $j \leq i$, $-j$ checked at $j$, $(j - i) + (i - f)$ steps away.
Then if $k \leq j$, $-k$ checked at $j$, $(k - j) + (j - i) + (i - f)$ steps away.
Other cases similarly.
OMGs as MCFGs

em1: \( \langle 0, \delta_0, \ldots, \delta_{i-1}, \alpha, \delta_{i+1}, \ldots, \delta_j \rangle(t_0, \ldots, t_{i-1}, s_0, t_i, t_{i+1}, \ldots, t_j) : - \langle 1, =f\alpha \rangle(s_0), \) 
\( \langle x, \delta_0, \ldots, \delta_{i-1}, f, \delta_{i+1}, \ldots, \delta_j \rangle(t_0, \ldots, t_j) \)

em2: \( \langle 0, \eta_0, \ldots, \eta_{j+\ell+1} \rangle(u_0, \ldots, u_{j+\ell+1}) : - \langle 0, \delta_0, \ldots, \delta_{i-1}, =f\alpha, \delta_{i+1}, \ldots, \delta_j, \rangle(s_0, \ldots, s_j), \) 
\( \langle x, \gamma_0, \ldots, \gamma_{j-1}, f, \gamma_{j+1}, \ldots, \gamma_{\ell} \rangle(t_0, \ldots, t_{\ell}) \)

where \( \eta_0, \ldots, \eta_{j+\ell+1} \) is the sort of \( \delta_0, \ldots, \delta_{i-1}, \alpha, \delta_{i+1}, \ldots, \delta_j, \gamma_0, \ldots, \gamma_{k-1}, \gamma_{k+1}, \ldots, \gamma_{\ell}, \) and \( u_0, \ldots, u_{j+\ell+1} \) the corresponding sort of \( s_0, \ldots, s_{i-1}, t_k s_i, s_{i+1}, \ldots, s_j, t_0, \ldots, t_{k-1}, t_{k+1}, \ldots, t_{\ell}. \)

em3: \( \langle 0, \eta_0, \ldots, \eta_{j+\ell+2} \rangle(u_0, \ldots, u_{j+\ell+2}) : - \langle x, \delta_0, \ldots, \delta_{i-1}, =f\alpha, \delta_{i+1}, \ldots, \delta_j, \rangle(s_0, \ldots, s_j), \) 
\( \langle y, \gamma_0, \ldots, \gamma_{j-1}, f\beta, \gamma_{j+1}, \ldots, \gamma_{\ell} \rangle(t_0, \ldots, t_{\ell}) \)

where \( \eta_0, \ldots, \eta_{j+\ell+2} \) is the sort of \( \delta_0, \ldots, \delta_{i-1}, \alpha, \delta_{i+1}, \ldots, \delta_j, \gamma_0, \ldots, \gamma_{k-1}, \beta, \gamma_{k+1}, \ldots, \gamma_{\ell}, \) and \( u_0, \ldots, u_{j+\ell+2} \) the corresponding sort of \( s_0, \ldots, s_j, t_0, \ldots, t_{\ell}. \)
non-permuting ‘naive Zapotec’

\[ \langle 0, C \rangle \text{ (praised the students the idea)} \]
\[ \langle 0, -v, +vC \rangle \text{ (praised, the students the idea)} \]
\[ \langle 1, =T + vC \rangle (\varepsilon) \quad \langle 0, -v, T \rangle \text{ (praised, the students the idea)} \]
\[ \langle 0, -v, -ep, +epT \rangle \text{ (praised, the students, the idea)} \]
\[ \langle 1, =v + epT \rangle (\varepsilon) \quad \langle 0, -v, -ep, v \rangle \text{ (praised, the students, the idea)} \]
\[ \langle 0, -v, =Dv \rangle \text{ (praised, the idea)} \quad \langle 0, D-ep \rangle \text{ (the students)} \]
\[ \langle 0, -v, -ep, +ep = Dv \rangle \text{ (praised, the idea, } \varepsilon \rangle \quad \langle 1, =ND-ep \rangle \text{ (the)} \quad \langle 1, N \rangle \text{ (students)} \]
\[ \langle 1, =V + ep = Dv \rangle (\varepsilon) \quad \langle 0, V-v, -ep \rangle \text{ (praised, the idea)} \]
\[ \langle 1, =DV-v \rangle \text{ (praised)} \quad \langle 0, D-ep \rangle \text{ (the idea)} \]
\[ \langle 1, =ND-ep \rangle \text{ (the)} \quad \langle 1, N \rangle \text{ (idea)} \]

- Non-permuting, and well-nested!
- Unlike many Ep+1-MCFG hypotheses, ‘semantically coherent’
SpIC-violating non-permuting well-nested OMG

- non-permuting, well-nested movements may violating SpIC – when Spec components do not move higher than components of selecting clause

\[
\text{em2: } \langle 0, \eta_0, \ldots, \eta_{j+\ell+1} \rangle (u_0, \ldots, u_{j+\ell+1}) : -
\langle 0, \delta_0, \ldots, \delta_{i-1}, =f \alpha, \delta_{i+1}, \ldots, \delta_j, \rangle (s_0, \ldots, s_j),
\langle x, \gamma_0, \ldots, \gamma_{j-1}, f, \gamma_{j+1}, \ldots, \gamma_\ell \rangle (t_0, \ldots, t_\ell)
\]

where \( \eta_0, \ldots, \eta_{j+\ell+1} \) is the sort of \( \delta_0, \ldots, \delta_{i-1}, \alpha, \delta_{i+1}, \ldots, \delta_j, \gamma_0, \ldots, \gamma_{k-1}, \gamma_{k+1}, \ldots, \gamma_\ell \).

Cf. informal discussions of conditions on “surfing” movement (Sauerland’96, Abels’07, . . . )
Conclusions

- MGs provide a succinct notation for ‘strongly equivalent’ MCFGs
- OMGs provide posets $\text{Cat}_\leq, \text{Ep}_\sqsubseteq$
  - Selection can be restricted with $\leq$ for ‘Order universals’
  - Movement can be restricted with $\sqsubseteq$ for ‘BOIM’
  - More succinct than strongly equivalent MGs, MCFGs.
- Finite class $\text{OML}_{\text{UG}} \not\subseteq (|\text{Ep}| + 1)-\text{MCFL}(2) \not\subseteq \text{MCFL}$
- Some OMGs have non-permuting strong equivalents, intersecting with the range of Yoshinaka&Clark’s learner. And some of these are well-nested.

(Q) For which UG, $\subseteq \text{OMG}_{ug}$ non-permuting?

(Q) For which UG, $\subseteq \text{OMG}_{ug}$ well-nested?

(EQ4) ‘Semantically appropriate’ HG $\subseteq \text{MCFG}_{wn}$?
Abels, K.
Towards a restrictive theory of (remnant) movement: Improper movement, remnant movement, and a linear asymmetry.

Abels, K.
Some implications of improper movement for cartography.

Abner, N.
Selection in grammar: Evidence and consequences.

Chomsky, N.
Conditions on transformations.

Chomsky, N.
*Knowledge of Language*.

Chomsky, N.
*The Minimalist Program*.

Cinque, G.

Cinque, G.
Deriving Greenberg’s Universal 20 and its exceptions.

Fukui, N.
A note on improper movement.
Greenberg, J.
Some universals of grammar with particular reference to the order of meaningful elements.

Grimshaw, J.
Complement selection and the lexicon.

Harkema, H.
A characterization of minimalist languages.

Kaji, Y., Nakanishi, R., Seki, H., and Kasami, T.
The universal recognition problems for multiple context-free grammars and for linear context-free rewriting systems.

Kobele, G. M.
Inverse linking via function composition.
*Natural Language Semantics* 18 (2010), 183–196.

Kobele, G. M., and Kandybowicz, J.
A normal form theorem for minimalist grammars.

Kuno, S., and Takami, K.
*Functional Constraints in Grammar: On the Unergative-Unaccusative Distinction.*

Levin, B., and Rappaport Hovav, M.
*Unaccusativity: At the Syntax-Lexical Semantics Interface.*

Manzini, M. R., and Savoia, L. M.
Clitics: Cooccurrence and mutual exclusion patterns.

Manzini, M. R., and Savoia, L. M.
Syncretism and suppletion in clitic systems: Underspecification, silent clitics or neither?

May, R.
Must comp-to-comp movement be stipulated.

McNaughton, R., and Papert, S.
*Counter-Free Automata*.

Michaelis, J.
Derivational minimalism is mildly context-sensitive.

Michaelis, J.
Transforming linear context free rewriting systems into minimalist grammars.

Nakajima, H.
Adverbial cognate objects.

Obata, M., and Epstein, S. D.
Deducing improper movement from phase based C-to-T Phi transfer: Feature-splitting internal merge.

Ohara, M.
*Essex Graduate Student Papers in Language and Linguistics* 1 (1997).
Pesetsky, D.
*Paths and Categories.*

Rambow, O., and Satta, G.
Independent parallelism in finite copying parallel rewriting systems.
*Theoretical Computer Science 223* (1999), 87–120.

Rizzi, L.
Locality and left periphery.

Seki, H., Matsumura, T., Fujii, M., and Kasami, T.
On multiple context-free grammars.

Stabler, E. P.
Computational perspectives on minimalism.

Thatcher, J. W.
Characterizing derivation trees of context-free grammars through a generalization of finite automata theory.

Yoshinaka, R., and Clark, A.
Polynomial time learning of some multiple context-free languages with a minimally adequate teacher.