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) HL $\subseteq$ MCFL?

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

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

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

(EQ3) 'Semantically appropriate' HG $\subseteq$ 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

$\mathrm{MG} = \langle\, \Sigma, \mathrm{Cat}, \mathrm{Ep}, \mathrm{Lex}, \mathrm{M}, \mathrm{S} \,\rangle$ where

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

Σ, Cat, Sel, Ep, Lic, Lex finite, non-empty, pairwise disjoint.

'naive Zapotec' VSO: *praised the students the idea*

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

Each MG names a 'strongly equivalent' k-MCFG, $k = |\,\text{Ep}\,| + 1$.

MGs as MCFGs

Given $\mathrm{MG} = \langle \Sigma, \mathrm{Cat}, \mathrm{Ep}, \mathrm{Lex}, \mathrm{M}, \mathrm{S} \rangle$, we define an MCFG with two start categories $\mathrm{MG} = \langle \Sigma, N, P, \{\langle 0, \mathrm{S} \rangle, \langle 1, \mathrm{S} \rangle\} \rangle$, defining the language

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

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

MGs as MCFGs

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

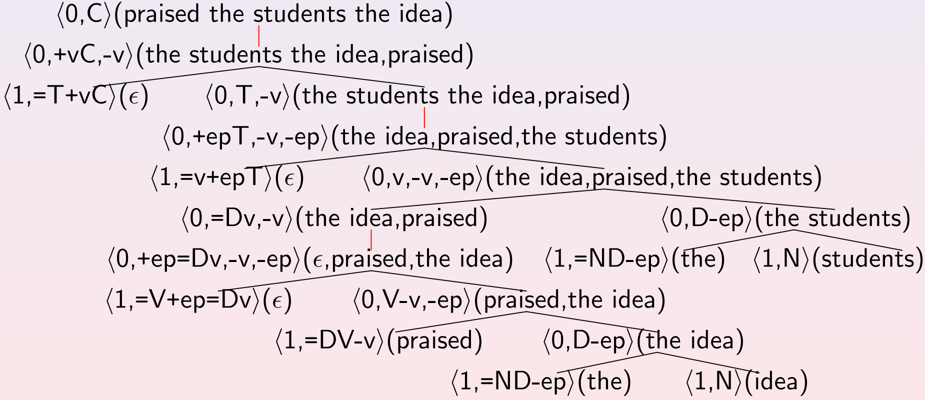
$$
\begin{array}{rll}
\text{lex:} & \langle 1, \alpha \rangle(s) & \langle s, \alpha \rangle \in \text{Lex} \\
\text{em1:} & \langle 0, \alpha, \delta_1, \ldots, \delta_j \rangle(s_0 t_0, t_1, \ldots, t_j) :\!- & \\
 & \quad \langle 1, {=}f\,\alpha \rangle(s_0), & \\
 & \quad \langle x, f, \delta_1, \ldots, \delta_j \rangle(t_0, \ldots, t_j) & \\
\text{em2:} & \langle 0, \alpha, \delta_1, \ldots, \delta_i, \gamma_1, \ldots, \gamma_j \rangle(t_0 s_0, s_1, \ldots, s_i, t_1, \ldots, t_j) :\!- & \\
 & \quad \langle 0, {=}f\,\alpha, \delta_1, \ldots, \delta_i, \rangle(s_0, \ldots, s_i), & \\
 & \quad \langle x, f, \gamma_1, \ldots, \gamma_j \rangle(t_0, \ldots, t_j) & \\
\text{em3:} & \langle 0, \alpha, \beta, \delta_1, \ldots, \delta_i, \gamma_1, \ldots, \gamma_j \rangle(s_0, t_0, s_1, \ldots, s_i, t_1, \ldots, t_j) :\!- & \\
 & \quad \langle x, {=}f\,\alpha, \delta_1, \ldots, \delta_i, \rangle(s_0, \ldots, s_i), & \\
 & \quad \langle y, f\,\beta, \gamma_1, \ldots, \gamma_j \rangle(t_0, \ldots, t_j) & \\
\text{im1:} & \langle 0, \alpha, \delta_1, \ldots, \delta_{i-1}, \delta_{i+1}, \ldots, \delta_j \rangle(s_i s_0, s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_j) :\!- & \delta_i = \text{-f}, \text{SMC} \\
 & \quad \langle 0, {+}f\,\alpha, \delta_1, \ldots, \delta_j \rangle(s_0, \ldots, s_j) & \\
\text{im2:} & \langle 0, \alpha, \delta_1, \ldots, \delta_{i-1}, \beta, \delta_{i+1}, \ldots, \delta_j \rangle(s_0, \ldots, s_i) :\!- & \delta_i = \text{-f}\beta, \text{SMC} \\
 & \quad \langle 0, {+}f\,\alpha, \delta_1, \ldots, \delta_j \rangle(s_0, \ldots, s_i) &
\end{array}
$$

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

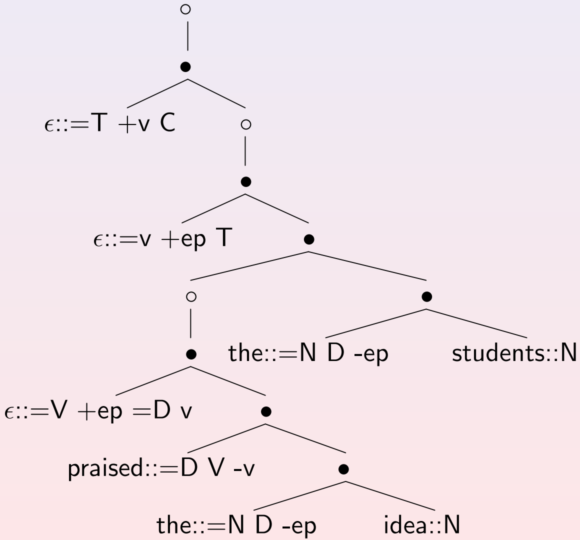
'naive Zapotec' VSO: *praised the students the idea*

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

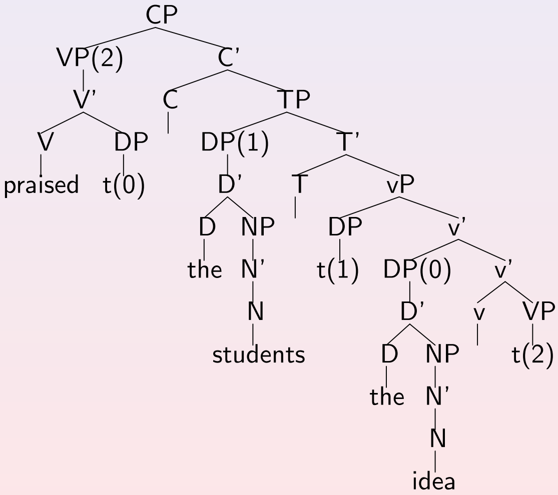
‘naive Zapotec’ VSO: *praised the students the idea*



'naive Zapotec' VSO: *praised the students the idea*

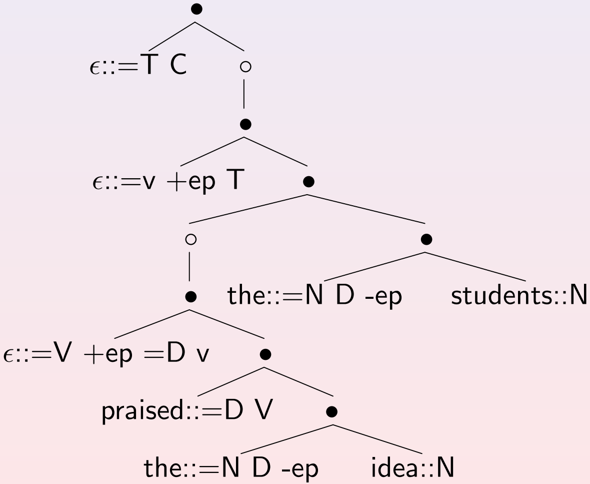


‘naive Zapotec’ VSO: *praised the students the idea*

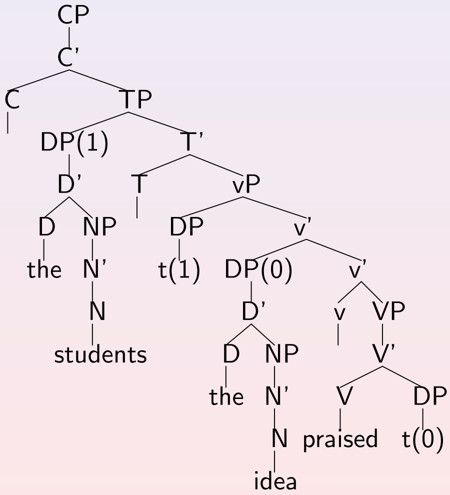


'naive Tamil' SOV: *the students the idea praised*

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

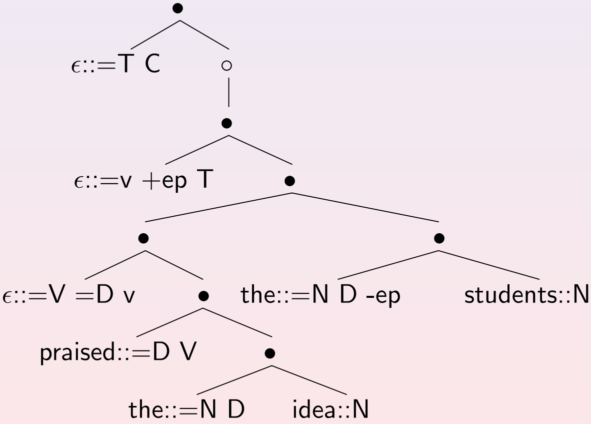


‘naive Tamil’ SOV: *the students the idea praised*

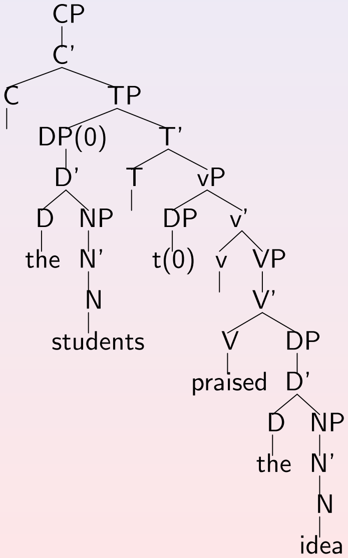


'naive English' SVO: *the students praised the idea*

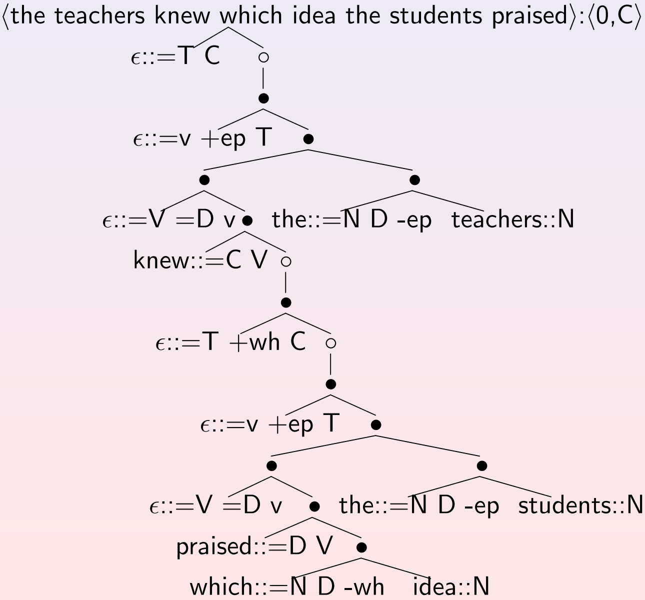
the	=N D
the	=N D −ep
students	N
idea	N
praised	=D V
ϵ	=V =D v
ϵ	=v +ep T
ϵ	=T C
which	=N D −wh
which	=N D −ep −wh
teachers	N
ϵ	=T +wh C
knew	=D V



‘naive English’ SVO: *the students praised the idea*



‘naive English’ SVO



'naive English' SVO MCFG (page 1)

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	=C V

$$\langle 0,\text{C}\rangle(s_0 t_0) \text{:-} \langle 1,=\text{T C}\rangle(s_0),\ \langle 0,\text{T}\rangle(t_0)$$

$$\langle 1,=\text{T C}\rangle(\epsilon)$$

$$\langle 0,\text{T}\rangle(s_1 s_0) \text{:-} \langle 0,+\text{ep T,-ep}\rangle(s_0, s_1)$$

$$\langle 0,+\text{ep T,-ep}\rangle(s_0 t_0, t_1) \text{:-} \langle 1,=\text{v} +\text{ep T}\rangle(s_0),\ \langle 0,\text{v,-ep}\rangle(t_0, t_1)$$

$$\langle 1,=\text{v} +\text{ep T}\rangle(\epsilon)$$

$$\langle 0,\text{v,-ep}\rangle(s_0, t_0) \text{:-} \langle 0,=\text{D v}\rangle(s_0),\ \langle 0,\text{D -ep}\rangle(t_0)$$

$$\langle 0,=\text{D v}\rangle(s_0 t_0) \text{:-} \langle 1,=\text{V} =\text{D v}\rangle(s_0),\ \langle 0,\text{V}\rangle(t_0)$$

$$\langle 1,=\text{V} =\text{D v}\rangle(\epsilon)$$

$$\langle 0,\text{V}\rangle(s_0 t_0) \text{:-} \langle 1,=\text{C V}\rangle(s_0),\ \langle 0,\text{C}\rangle(t_0)$$

$$\langle 1,=\text{C V}\rangle(\text{knew})$$

$$\langle 0,\text{C}\rangle(s_1 s_0) \text{:-} \langle 0,+\text{wh C,-wh}\rangle(s_0, s_1)$$

$$\langle 0,+\text{wh C,-wh}\rangle(s_0 t_0, t_1) \text{:-} \langle 1,=\text{V} +\text{wh C}\rangle(s_0)\ \langle 0,\text{V,-wh}\rangle(t_0, t_1$$

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

$$\langle 0,\text{V,-wh}\rangle(s_0 t_0, t_1) \text{:-} \langle 1,=\text{C V}\rangle(s_0),\ \langle 0,\text{C,-wh}\rangle(t_0, t_1)$$

$$\langle 0,\text{C,-wh}\rangle(s_0 t_0, t_1) \text{:-em1}_1[\langle 1,=\text{T C}\rangle(s_0),\ \langle 0,\text{T,-wh}\rangle(t_0, t$$

......

'naive English' SVO MCFG (page 2)

$$\begin{array}{rcl}
 & \ldots\ldots & \\
\langle 0,\mathsf{T,\text{-}wh}\rangle(s_2 s_0, s_1) & \text{:-} & \langle 0,\mathsf{+ep\ T,\text{-}wh,\text{-}ep}\rangle(s_0, s_1, s_2) \\
\langle 0,\mathsf{+ep\ T,\text{-}wh,\text{-}ep}\rangle(s_0 t_0, t_1, t_2) & \text{:-} & \langle 1,\mathsf{=v\ +ep\ T}\rangle(s_0),\ \langle 0,\mathsf{v,\text{-}wh,\text{-}ep}\rangle(t_0, t_1, t_2) \\
\langle 0,\mathsf{v,\text{-}wh,\text{-}ep}\rangle(s_0, s_1, t_0) & \text{:-} & \langle 0,\mathsf{=D\ v,\text{-}wh}\rangle(s_0, s_1),\ \langle 0,\mathsf{D\ \text{-}ep}\rangle(t_0) \\
\langle 0,\mathsf{=D\ v,\text{-}wh}\rangle(s_0 t_0, t_1) & \text{:-} & \langle 1,\mathsf{=V\ =D\ v}\rangle(s_0),\ \langle 0,\mathsf{V,\text{-}wh}\rangle(t_0, t_1) \\
\langle 0,\mathsf{V,\text{-}wh}\rangle(s_0, t_0) & \text{:-} & \langle 1,\mathsf{=D\ V}\rangle(s_0),\ \langle 0,\mathsf{D\ \text{-}wh}\rangle(t_0) \\
\langle 1,\mathsf{=D\ V}\rangle(\text{praised}) & & \\
\langle 0,\mathsf{D\ \text{-}wh}\rangle(s_0 t_0) & \text{:-} & \langle 1,\mathsf{=N\ D\ \text{-}wh}\rangle(s_0),\ \langle 1,\mathsf{N}\rangle(t_0) \\
\langle 1,\mathsf{=N\ D\ \text{-}wh}\rangle(\text{which}) & & \\
\langle 1,\mathsf{N}\rangle(\text{teachers}) & & \\
\langle 1,\mathsf{N}\rangle(\text{idea}) & & \\
\langle 1,\mathsf{N}\rangle(\text{students}) & & \\
\langle 0,\mathsf{D\ \text{-}ep}\rangle(s_0 t_0) & \text{:-} & \langle 1,\mathsf{=N\ D\ \text{-}ep}\rangle(s_0),\ \langle 1,\mathsf{N}\rangle(t_0) \\
\langle 1,\mathsf{=N\ D\ \text{-}ep}\rangle(\text{the}) & & \\
\langle 0,\mathsf{T,\text{-}wh}\rangle(s_1 s_0, s_2) & \text{:-} & \langle 0,\mathsf{+ep\ T,\text{-}ep,\text{-}wh}\rangle(s_0, s_1, s_2) \\
\langle 0,\mathsf{+ep\ T,\text{-}ep,\text{-}wh}\rangle(s_0 t_0, t_1, t_2) & \text{:-} & \langle 1,\mathsf{=v\ +ep\ T}\rangle(s_0),\ \langle 0,\mathsf{v,\text{-}ep,\text{-}wh}\rangle(t_0, t_1, t_2) \\
 & \ldots\ldots &
\end{array}$$

‘naive English’ SVO MCFG (page 3)

$$
\begin{array}{rcl}
\ldots & \ldots & \\
\langle \text{0,v,-ep,-wh} \rangle(s_0, s_1, t_0) & \text{:-} & \langle \text{0,=D v,-ep} \rangle(s_0, s_1),\ \langle \text{0,D -wh} \rangle(t_0) \\
\langle \text{0,=D v,-ep} \rangle(s_0 t_0, t_1 & \text{:-} & \langle \text{1,=V =D v} \rangle(s_0),\ \langle \text{0,V,-ep} \rangle(t_0, t_1) \\
\langle \text{0,V,-ep} \rangle(s_0, t_0) & \text{:-} & \langle \text{1,=D V} \rangle(s_0),\ \langle \text{0,D -ep} \rangle(t_0) \\
\langle \text{0,T,-wh} \rangle(s_0, s_1) & \text{:-} & \langle \text{0,+ep T,-ep -wh} \rangle(s_0, s_1)] \\
\langle \text{0,+ep T,-ep -wh} \rangle(s_0 t_0, t_1) & \text{:-} & \langle \text{1,=v +ep T} \rangle(s_0),\ \langle \text{0,v,-ep -wh} \rangle(t_0, t_1) \\
\langle \text{0,v,-ep -wh} \rangle(s_0, t_0) & \text{:-} & \langle \text{0,=D v} \rangle(s_0),\ \langle \text{0,D -ep -wh} \rangle(t_0) \\
\langle \text{0,D -ep -wh} \rangle(s_0 t_0) & \text{:-} & \langle \text{1,=N D -ep -wh} \rangle(s_0),\ \langle \text{1,N} \rangle(t_0, t_1) \\
\langle \text{1,=N D -ep -wh} \rangle(\mathrm{which}) & & \\
\langle \text{0,v,-ep -wh} \rangle(t_0 s_0, s_1) & \text{:-} & \langle \text{0,=D v,-ep -wh} \rangle(s_0, s_1),\ \langle \text{0,D} \rangle(t_0) \\
\langle \text{0,=D v,-ep -wh} \rangle(s_0 t_0, t_1) & \text{:-} & \langle \text{1,=V =D v} \rangle(s_0),\ \langle \text{0,V,-ep -wh} \rangle(t_0, t_1) \\
\langle \text{0,V,-ep -wh} \rangle(s_0, t_0) & \text{:-} & \langle \text{1,=D V} \rangle(s_0)\ \langle \text{0,D -ep -wh} \rangle(t_0) \\
\langle \text{0,D} \rangle(s_0 t_0) & \text{:-} & \langle \text{1,=N D} \rangle(s_0),\ \langle \text{1,N} \rangle(t_0)] \\
\langle \text{1,=N D} \rangle(\mathrm{the}) & & \\
\langle \text{0,V} \rangle(s_0 t_0) & \text{:-} & \langle \text{1,=D V} \rangle(s_0),\ \langle \text{0,D} \rangle(t_0) \\
\langle \text{0,v,-ep} \rangle(t_0 s_0, s_1) & \text{:-} & \langle \text{0,=D v,-ep} \rangle(s_0, s_1),\ \langle \text{0,D} \rangle(t_0)
\end{array}
$$

Constituent order

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

  $$=D\ =D\ V \quad \text{is analagous to} \quad (D\backslash V)/D.$$

  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

order	C	MG	order	C	MG
1234	**4**	**4**	1324	**0**	**0**
1243	**3**	**3**	1342	**1**	**3**
1423	**1**	**3**	1432	**3**	**2**
4123	**2**	**3**	4132	**1**	**2**
2134	**0**	**0**	2314	**0**	**0**
2143	**0**	**0**	2341	**1**	**3**
2413	**0**	**0**	2431	**2**	**2**
4213	**0**	**0**	4231	**2**	**2**
3124	**0**	**0**	3214	**0**	**0**
3142	**0**	**2**	3241	**0**	**0**
3412	**1**	**3**	3421	**1**	**2**
4312	**2**	**2**	4321	**4**	**2**

	Cinque	MG
0=	unattested	underivable
1=	very few	4 licensors
2=	few	3 licensors
3=	many	2 licensors
4=	very many	0-1 licensors

Constituent order

order	C	MG	order	C	MG
1234	**4**	**4**	1324	**0**	**0**
1243	**3**	**3**	1342	**1**	**3**
1423	**1**	**3**	1432	**3**	**2**
4123	**2**	**3**	4132	**1**	**2**
2134	**0**	**0**	2314	**0**	**0**
2143	**0**	**0**	2341	**1**	**3**
2413	**0**	**0**	2431	**2**	**2**
4213	**0**	**0**	4231	**2**	**2**
3124	**0**	**0**	3214	**0**	**0**
3142	**0**	**2**	3241	**0**	**0**
3412	**1**	**3**	3421	**1**	**2**
4312	**2**	**2**	4321	**4**	**2**

	Cinque	MG
0=	unattested	underivable
1=	very few	4 licensors
2=	few	3 licensors
3=	many	2 licensors
4=	very many	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:

  $$\text{C} < \text{T} < \text{v} < \text{V} < \text{D} < \text{N}$$

- Rizzi'04 'left periphery'

  Force < Top1 < Foc < Top2 < Fin < Infl

- Cinque'99 adverbials

Mood <	Evidential <	Epistemic <	Habitual <	Inceptive
frankly	allegedly	probably	usually	suddenly

- Manzini&Savoia'04 clitic positions

  Def(uninfl) < Quant(3) < N(3pl) < P(1,2) < Origin < Loc < Measure

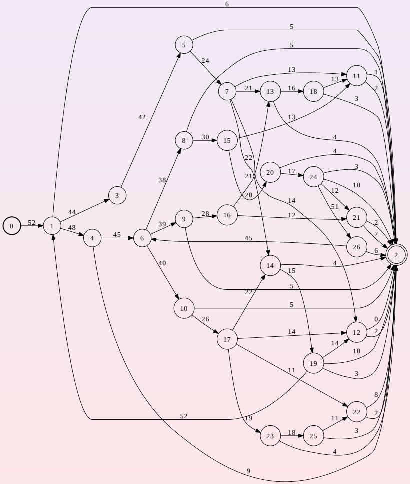
Constituent order and paths

- **(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)

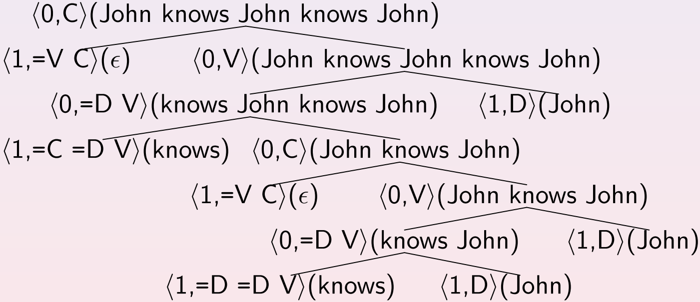
Paths in naive English



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

Tiny English

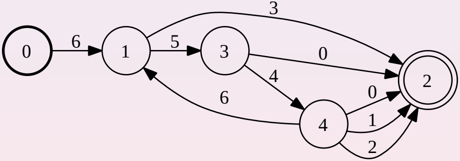
John	D
knows	=D =D V
knows	=C =D V
ϵ	=V C



Tiny English

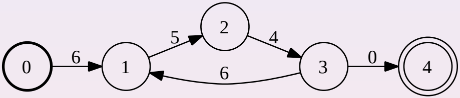
$0 = \langle 1, \mathrm{D} \rangle \quad 1 = \langle 1, \texttt{=}\mathrm{D}\texttt{=}\mathrm{DV} \rangle \quad 2 = \langle 1, \texttt{=}\mathrm{C}\texttt{=}\mathrm{DV} \rangle \quad 3 = \langle 1, \texttt{=}\mathrm{VC} \rangle$

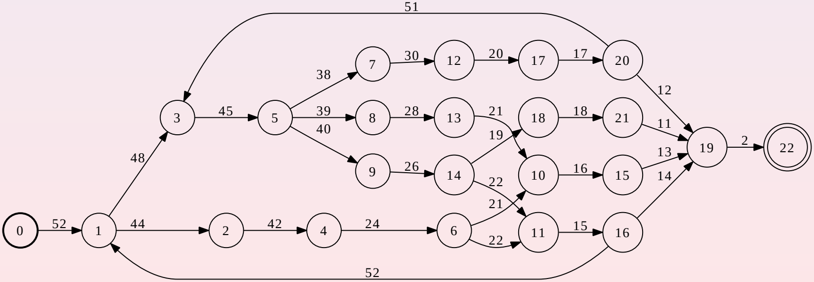
$4 = \langle 0, \texttt{=}\mathrm{DV} \rangle \quad 5 = \langle 0, \mathrm{V} \rangle \quad 6 = \langle 0, \mathrm{C} \rangle$



- 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.

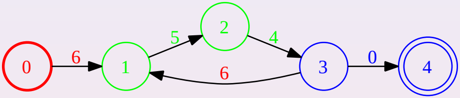
Spinal paths



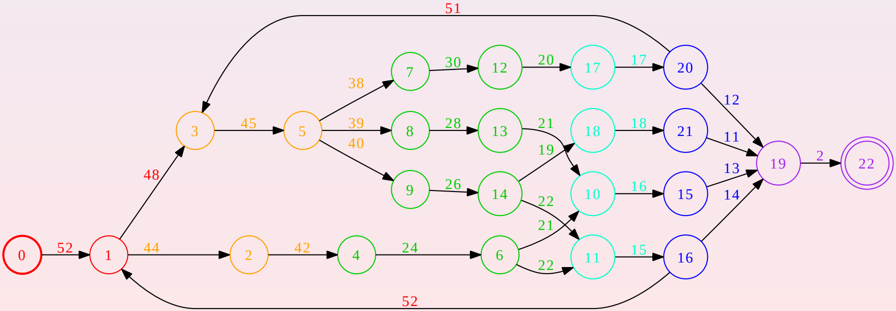


Spinal paths

tiny English: C, V, D:



Naive English: C, T, v, V, D, N:



'Cartography'

- **(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 a non-spinal path initiated at a spec position, and $f \neq g$, then either f covers$_{\leq}$ g or g minimal$_{\leq}$.

Movement order universals

- Naive English *which idea* moves twice, with −ep −wh

  the teachers knew which idea_i the students t_i praised t_i

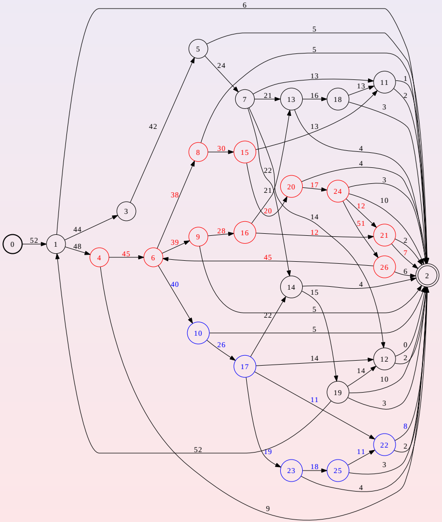
(BOIM) Other orders are 'improper', yielding ungrammaticality

* Who_i seems t_i will t_i leave?

* Who_i seems it is likely t_i to t_i leave?

- ep $\lhd$ scrambling $\lhd$ wh $\lhd$ 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 $\mathrm{Ep}_{\trianglelefteq}$ are ordered so that if a component moves for g and then for f, $f \trianglelefteq g$

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

Ordered MGs

$\text{OMG} = \langle\, \Sigma, \text{Cat}_{\leq}, \text{Ep}_{\trianglelefteq}, \text{Df}, \mu, \text{Lex}, \text{M}, \text{S}\,\rangle$

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

- *a* selects *b* iff *a* covers$_{\leq}$ *b* or *b* minimal$_{\trianglelefteq}$
- Semantic type may restrict selection and deficiency
- Df maps each Cat to at most one value – at most one specifier..

Many linguists assume $\text{UG} = \langle\, \text{Cat}_{\leq}, \text{Ep}_{\trianglelefteq}, \text{Df}, \text{M}, \text{S}\,\rangle$; all variation in Lex.

Ordered MGs

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	=C V

$C<T<v<V<D<N$

$\mathrm{wh} \lhd \mathrm{ep}$

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

the	D
the	D {ep}
students	N
idea	N
praised	V
ϵ	v
ϵ	T
ϵ	C
which	D {wh}
which	D {ep,wh}
teachers	N
ϵ	C
knew	V

Ordered MGs

Expressive power

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

Learnability

- $\forall$UG, $\mathrm{OML}_{\mathrm{UG}}$ 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)

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

Non-permuting OMGs

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

$$\text{im1: } \langle\, 0, \alpha, \delta_1, \ldots, \delta_{i-1}, \delta_{i+1}, \ldots, \delta_j \,\rangle(s_i s_0, s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_j) \,\text{:-}\\ \langle\, 0, \text{+}f\,\alpha, \delta_0, \ldots, \delta_j \,\rangle(s_0, \ldots, s_j)$$

? Can we change this to something like:

$$\text{im1??: } \langle\, 0, \delta_1, \ldots, \delta_{i-1}, \alpha \,\rangle(s_1, \ldots, s_{i-1}, s_i s_0) \,\text{:-}\\ \langle\, 0, \delta_1, \ldots, \delta_i, \text{+}f\,\alpha \,\rangle(s_1, \ldots, s_i, s0)$$

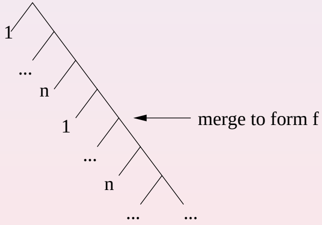
Yes, for <u>certain</u> OMGs we can formulate (strongly equivalent) non-permuting rules. E.g. suppose $\leq, \trianglelefteq$ total, Df total and 1-1. . .

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 $\forall \mathrm{c} \in \mathrm{cat}, \mathrm{c} \mapsto +\mathrm{c}$ 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

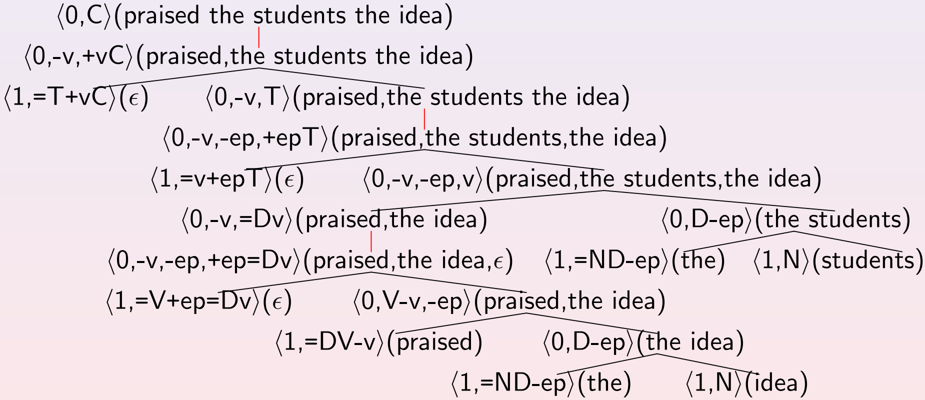
$$\begin{array}{l}
\text{em1: } \langle\, 0, \delta_0, \ldots, \delta_{i-1}, \alpha, \delta_{i+1}, \ldots, \delta_j \,\rangle(t_0, \ldots, t_{i-1}, s_0 t_i, t_{i+1}, \ldots, t_j) \text{ :-} \\
\qquad \langle\, 1, \text{=}f\,\alpha \,\rangle(s_0), \\
\qquad \langle\, x, \delta_0, \ldots, \delta_{i-1}, f, \delta_{i+1}, \ldots, \delta_j \,\rangle(t_0, \ldots, t_j) \\
\text{em2: } \langle\, 0, \eta_0, \ldots, \eta_{j+\ell+1} \,\rangle(u_0, \ldots, u_{j+\ell+1}) \text{ :-} \\
\qquad \langle\, 0, \delta_0, \ldots, \delta_{i-1}, \text{=}f\,\alpha, \delta_{i+1}, \ldots, \delta_j, \,\rangle(s_0, \ldots, s_j), \\
\qquad \langle\, x, \gamma_0, \ldots, \gamma_{j-1}, f, \gamma_{j+1}, \ldots, \gamma_\ell \,\rangle(t_0, \ldots, t_\ell)
\end{array}$$

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 $s_0, \ldots, s_{i-1}, t_k s_i, s_{i+1}, \ldots, s_j, t_0, \ldots, t_{k-1}, t_{k+1}, \ldots, t_\ell$.

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

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'



- 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

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

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 $\mathrm{Cat}_{\leq}$, $\mathrm{Ep}_{\trianglelefteq}$
  - Selection can be restricted with $\leq$ for 'Order universals'
  - Movement can be restricted with $\trianglelefteq$ for 'BOIM'
  - More succinct than strongly equivalent MGs, MCFGs.
  - Finite class $\mathrm{OML}_{\mathrm{UG}} \subsetneq (|\mathrm{Ep}|+1)$-MCFL(2) $\subsetneq$ 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, $\subseteq \mathrm{OMG}_{ug}$ non-permuting?

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

(EQ4) 'Semantically appropriate' HG $\subseteq \mathrm{MCFG}_{wn}$?

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

Abels, K.
Some implications of improper movement for cartography.
In *Alternatives to Cartography*, J. van Craenenbroeck, Ed. Mouton de Gruyter, New York, 2009.

Abner, N.
Selection in grammar: Evidence and consequences.
UCLA Working Papers in Syntax and Semantics, 2010.

Chomsky, N.
Conditions on transformations.
In *A Festschrift for Morris Halle*, S. R. Anderson and P. Kiparsky, Eds. Holt, Rinehard & Winston, NY, 1973.
Reprinted in Chomsky's (1977) *Essays on Form and Interpretation.* Amsterdam: North-Holland.

Chomsky, N.
Knowledge of Language.
Praeger, NY, 1986.

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

Cinque, G.
Adverbs and Functional Heads: A Cross-Linguistic Perspective.
Oxford University Press, Oxford, 1999.

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

Fukui, N.
A note on improper movement.

Linguistic Review 10 (1993), 111–126.

Greenberg, J.
Some universals of grammar with particular reference to the order of meaningful elements.
In *Universals of Language: report of a conference held at Dobbs Ferry, New York, April 13-15, 1961*, J. Greenberg, Ed. MIT Press, Cambridge, Massachusetts, 1963.

Grimshaw, J.
Complement selection and the lexicon.
Linguistic Inquiry 10 (1979), 279–326.

Harkema, H.
A characterization of minimalist languages.
In *Proceedings, Logical Aspects of Computational Linguistics, LACL'01* (Port-aux-Rocs, Le Croisic, France, 2001).

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.
IEICE Transactions on Information and Systems E75-D, 1 (1992), 78–88.

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.
Manuscript, UCLA, 2001.

Kuno, S., and Takami, K.
Functional Constraints in Grammar: On the Unergative-Unaccusative Distinction.
John Benjamins, Amsterdam, 2004.

Levin, B., and Rappaport Hovav, M.
Unaccusativity: At the Syntax-Lexical Semantics Interface.
MIT Press, Cambridge, Massachusetts, 1995.

Manzini, M. R., and Savoia, L. M.

Clitics: Cooccurrence and mutual exclusion patterns.
In *The Structure of CP and IP*, L. Rizzi, Ed. Oxford University Press, NY, 2004, pp. 211–250.

Manzini, M. R., and Savoia, L. M.
Syncretism and suppletion in clitic systems: Underspecification, silent clitics or neither?
In *Syntactic Variation: The Dialects of Italy* (Cambridge, 2010), R. D'Alessandro, A. Ledgeway, and I. Roberts, Eds., Cambridge University Press, pp. 86–101.

May, R.
Must comp-to-comp movement be stipulated.
Linguistic Inquiry 10 (1979), 719–715.

McNaughton, R., and Papert, S.
Counter-Free Automata.
MIT Press, Cambridge, Massachusetts, 1971.

Michaelis, J.
Derivational minimalism is mildly context-sensitive.
In *Proceedings, Logical Aspects of Computational Linguistics, LACL'98* (NY, 1998), Springer, pp. 179–198.

Michaelis, J.
Transforming linear context free rewriting systems into minimalist grammars.
In *Logical Aspects of Computational Linguistics* (NY, 2001), P. de Groote, G. Morrill, and C. Retoré, Eds., Lecture Notes in Artificial Intelligence, No. 2099, Springer, pp. 228–244.

Nakajima, H.
Adverbial cognate objects.
Linguistic Inquiry 37, 4 (2006), 674–684.

Obata, M., and Epstein, S. D.
Deducing improper movement from phase based C-to-T Phi transfer: Feature-splitting internal merge.
In *27th West Coast Conference on Formal Linguistics, WCCFL XXVII* (2008).

Ohara, M.
The interface between syntax and lexical-semantics: An analysis of benefactive constructions in Japanese.
Essex Graduate Student Papers in Language and Linguistics 1 (1997).

Pesetsky, D.
Paths and Categories.
PhD thesis, Massachusetts Institute of Technology, 1982.

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.
In *Structures and Beyond: Cartography of Syntactic Structures, Volume 3*, A. Belletti, Ed. Oxford, NY, 2004, pp. 104–131.

Seki, H., Matsumura, T., Fujii, M., and Kasami, T.
On multiple context-free grammars.
Theoretical Computer Science 88 (1991), 191–229.

Stabler, E. P.
Computational perspectives on minimalism.
In *Oxford Handbook of Minimalism*, C. Boeckx, Ed. Oxford University Press, Oxford, 2010, pp. 616–641.

Thatcher, J. W.
Characterizing derivation trees of context-free grammars through a generalization of finite automata theory.
Journal of Computer and System Sciences 1, 4 (1967), 317 – 322.

Yoshinaka, R., and Clark, A.
Polynomial time learning of some multiple context-free languages with a minimally adequate teacher.
In *Proceedings of the 15th Conference on Formal Grammar* (2010).