Consistent affix linearization in subanalysis learning

Sebastian Bank
sebastian.bank@uni-leipzig.de

Department of Linguistics
University of Leipzig

American International Morphology Meeting 2
University of California, San Diego
November 8–10, 2013

Background: Linearization in morphological grammar

Consistent general order

template, slots, position classes

"…the linear descriptive sequence of affixes is reflected in the rule system as the relative **ordering of the corresponding rules**. That is, the fact that a word has the form *stem* + $affix_1$ + $affix_2$ reflects the fact that the rule attaching $affix_1$ applied 'before' the rule that attaches $affix_2$." (Anderson 1992)

Contextually restricted reordering

deviations

Ordered **rule blocks**, reversible by Rules of Referral (Stump 2001)

Serialized **tree nodes**, overridden by Fission and Readjustment (Halle / Marantz 1993)

Consistency requirements

given **AB**, **BA** if **A** < **B** then ***B** < **A** *no symmetry*
hence $\mathbf{A_1B}$, $\mathbf{BA_2}$

given **AB**, **BC**, **CA** if **A** < **B** and **B** < **C** then ***C** < **A** *no cycles*
hence $\mathbf{A_1B}$, **BC**, $\mathbf{CA_2}$

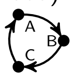
… *alternatives to homophony*: non-subanalysis, reordering

Background: Templatic linearization + reordering

Expectations

- morphologies with reordering or cycles are structurally more complex
- typologically overriding the general order should be dispreferred
- sequences like **AB**, **BA** or **AB**, **BC**, **CA** should be harder to subanalyze
- diachronic pressure to realign order-inconsistent morphemes

Alternative hypotheses

1. affix linearization reflects pairwise predecessor-successor relations (precedence constraints, e.g. Paster 2005) or bigram constraints (Ryan 2010):
   they can contain **cycles** 

   where the parts never co-occur *en bloc*
   allow cycles, nontransitivity
2. learners do not generalize over the affix linearization (rote memorization):
   different occurrences of the same affix can be **inconsistent** 

   allow symmetry

Mission: Subanalysis in morphological learning

Algorithm-based **learning** of morphological grammar(s)

Components	meaning assignment, segmentation, linearization, blocking, …
Challenges	search-space (analytical options), local vs. global optimization
	interaction of components, look-ahead, syncretism vs. homonymy
Goals	more falsifiable predictions for analysis/framework comparison
	determine the distinct logical properties of different formalisms
	explore logical space of morphological theories

In this talk

combined meaning assignment and **subanalysis learner** (Bank & Trommer 2013)
using iterative local optimization (Harmonic Serialism, McCarthy 2010)
that can satisfy three different **linearization requirements**

A templatic ordering (position classes)

B binary predecessor-successor relations

C no restrictions

$\mathbf{A} \subset \mathbf{B} \subset \mathbf{C}$

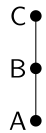
Morpheme order as asymmetric transitive relation

Each morpheme string in a paradigm cell gives a **chain** (strict total ordering), i.e. a relation that is *asymmetric*, *transitive* and *total*:

ABC $\Leftrightarrow \{\langle A,B\rangle,\langle A,C\rangle,\langle B,C\rangle\} \Leftrightarrow$

<	A	B	C
A		×	×
B			×
C			

$\Leftrightarrow \begin{bmatrix} 0&1&1\\ 0&0&1\\ 0&0&0 \end{bmatrix} \Leftrightarrow$



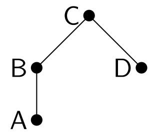
asymmetry	$a < b \Rightarrow b \nless a$	*antisymmetic, irreflexive*
transitivity	$a < b \wedge b < c \Rightarrow a < c$	
totality	$a < b \vee b < a$	*trichotomous*

Morpheme chains are **consistent** if their union is a strict **partial ordering**, i.e. a relation that is *asymmetric* and *transitive*: *directed acyclic graph*

ABC, **DC** $\Leftrightarrow$ $\{\langle A,B\rangle,\langle A,C\rangle, \langle B,C\rangle,\langle D,C\rangle\}$ $\Leftrightarrow$

<	A	B	C	D
A		×	×	
B			×	
C				
D			×	

$\Leftrightarrow \begin{bmatrix} 0&1&1&0\\ 0&0&1&0\\ 0&0&0&0\\ 0&0&1&0 \end{bmatrix} \Leftrightarrow$

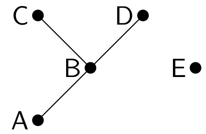


From partial ordering to total ordering (cf. Kahn 1962)

Partial orderings can always be *linearly extended* into a compatible **total** ordering.

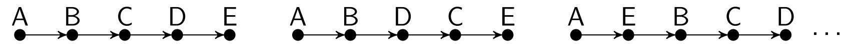
cycle-free

ABC, **BD**, **E** $\Leftrightarrow$



$\Leftrightarrow$ **A** < **B** < **C** **C** ∥ **D**
A < **B** < **D** **E** ∥ **A**, **B**, **C**, **D**

e.g.:



Such *topological sorting* is known from all kinds of **dependency** resolution tasks:

Enumerate a set A with a strict partial ordering '<' in a topologically sorted order

```
while A ≠ ∅ do                              ▷ items to assign
    M ← {a ∈ A | ∀b ∈ A : b ≮ a}            ▷ minimal elements
    for all m ∈ M do                        ▷ in any order
        yield m
    A ← A ∖ M                               ▷ remaining items
```


From partial ordering to linearization template

Templates assign each morpheme to a **slot** $i \in \mathbb{N}$ such that slots are totally ordered.

$\mathbf{A} \mapsto 1, \mathbf{B} \mapsto 2, \mathbf{C} \mapsto 3, \mathbf{D} \mapsto 3, \mathbf{E} \mapsto 3 \qquad 1 < 2 < 3$ *ordered partition*

The ordering between the slots gives a strict **weak ordering** on the morphemes, i.e. a partial ordering in which also the *incomparability* is transitive.

$\{\mathbf{A}\}_1 < \{\mathbf{B}\}_2 < \{\mathbf{C},\mathbf{D},\mathbf{E}\}_3 \qquad \mathbf{C} \parallel \mathbf{D} \wedge \mathbf{D} \parallel \mathbf{E} \Rightarrow \mathbf{C} \parallel \mathbf{E}$ *negative transitivity*

Deriving a **template** from a partial ordering is almost identical to topological sort:

Generate compatible $\langle item, slot\rangle$ pairs for a set A with a strict partial ordering '<'

```
i ← 0
while A ≠ ∅ do                              ▷ items to assign
    i ← i + 1                               ▷ next slot index
    M ← {a ∈ A | ∀b ∈ A : b ≮ a}            ▷ minimal elements
    for all m ∈ M do                        ▷ in any order
        yield ⟨m, i⟩                        ▷ m ↦ i
    A ← A ∖ M                               ▷ remaining items
```


Position class template generator (left-anchored version)

$C = \{\langle \mathrm{b}, \mathrm{in}\rangle, \langle \mathrm{b}, \mathrm{ist}\rangle, \langle \mathrm{ist}\rangle, \langle \mathrm{s}, \mathrm{in}, \mathrm{d}\rangle, \langle \mathrm{s}, \mathrm{ei}, \mathrm{d}\rangle\}$

$R = \{\langle \mathrm{b}, \mathrm{in}\rangle, \langle \mathrm{b}, \mathrm{ist}\rangle, \langle \mathrm{s}, \mathrm{in}\rangle, \langle \mathrm{s}, \mathrm{d}\rangle, \langle \mathrm{in}, \mathrm{d}\rangle, \langle \mathrm{ei}, \mathrm{d}\rangle\} = R^+$

$M = \{\langle \mathrm{b}, \varnothing\rangle, \langle \mathrm{in}, \{\mathrm{b}, \mathrm{s}\}\rangle, \langle \mathrm{ist}, \{\mathrm{b}\}\rangle, \langle \mathrm{s}, \varnothing\rangle, \langle \mathrm{d}, \{\mathrm{s}, \mathrm{in}, \mathrm{ei}\}\rangle, \langle \mathrm{ei}, \{\mathrm{s}\}\rangle\}$

$A_1 = \{\mathrm{b}, \mathrm{in}, \mathrm{ist}, \mathrm{s}, \mathrm{d}, \mathrm{ei}\}$ $\quad A_2 = \{\mathrm{in}, \mathrm{ist}, \mathrm{d}, \mathrm{ei}\}$ $\quad A_3 = \{\mathrm{d}\}$

$S_1 = \{\mathrm{b}, \mathrm{s}\}$ $\quad S_2 = \{\mathrm{in}, \mathrm{ist}, \mathrm{ei}\}$ $\quad S_3 = \{\mathrm{d}\}$

$Slots = \langle \{\mathrm{b}, \mathrm{s}\}_1, \{\mathrm{in}, \mathrm{ist}, \mathrm{ei}\}_2, \{\mathrm{d}\}_3\rangle$

German 'to be'

	sg	**pl**
1	b-in	s-in-d
2	b-ist	s-ei-d
3	ist	s-in-d

$R = \begin{bmatrix} 0&1&1&0&0&0\\ 0&0&0&0&0&0\\ 0&0&0&0&0&0\\ 0&1&0&0&1&1\\ 0&0&0&0&0&0\\ 0&0&0&0&1&0 \end{bmatrix}$

Input: a set $C := \{\langle m_1, \ldots, m_n\rangle, \ldots\}$ of morpheme sequences ms of length $n > 0$
Output: an initially empty list $Slots := \langle S_1, \ldots, S_n\rangle$ of morphemes in the same slot S

```
R ← {⟨ms_i, ms_{i+1}⟩ | ms ∈ C, i ∈ [1 .. |ms| − 1]}     ▷ predecessor-successor relation
if R⁺ ∩ (R⁺)⁻¹ = ∅ then                                 ▷ transitive closure is asymmetric
    M ← {⟨m, P⟩ | ms ∈ C, m ∈ ms, P = {l | ⟨l, m⟩ ∈ R}}  ▷ m ↦ predecessors
    A ← {m | ⟨m, P⟩ ∈ M}                                ▷ m to be assigned to a slot
    while A ≠ ∅ do
        S ← {m | m ∈ A, ⟨m, P⟩ ∈ M, P ∩ A = ∅}           ▷ no predecessors to be assigned
        append S to Slots
        A ← A ∖ S                                       ▷ remaining m to assign
```

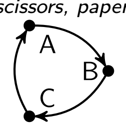
Morpheme order as predecessor-successor relation

rock, scissors, paper

AB, **BC**, **CA** ⇔ $\{\langle A,B\rangle,\langle B,C\rangle,\langle C,A\rangle\}$ ⇔

<	A	B	C
A		×	
B			×
C	×		

⇔ $\begin{bmatrix}0&1&0\\0&0&1\\1&0&0\end{bmatrix}$ ⇔



↻ **A < B < C** ↺ **A ≯ C > B**

asymmetry	$a < b \Rightarrow b \nless a$	$R \cap R^{-1} = \varnothing$
transitivity	$a < b \wedge b < c \Rightarrow a < c$	$R \circ R \subseteq R$

$\begin{bmatrix}0&1&0\\0&0&1\\1&0&0\end{bmatrix} \cup \begin{bmatrix}0&1&0\\0&0&1\\1&0&0\end{bmatrix} \circ \begin{bmatrix}0&1&0\\0&0&1\\1&0&0\end{bmatrix} = \begin{bmatrix}0&1&1\\1&0&1\\1&1&0\end{bmatrix}$ *composition* $\begin{bmatrix}0&1&0\\0&0&1\\1&0&0\end{bmatrix}^{+} = \begin{bmatrix}1&1&1\\1&1&1\\1&1&1\end{bmatrix}$ *transitive closure*

If the relation has a **cycle**, it is not possible to enumerate **all** elements in consistent order.
Topological sort ends up in an **infinite loop**.
Only the actually present **morpheme chains** are guaranteed to be serializable.

Still, morpheme chains cannot have inconsistent precedence with each other given **AB**, **BA** if **A** < **B** then ***B** < **A**

Maintaining precedence in incremental segmentation

Annotate $_{left}$*string*$_{right}$ with *left* = **predecessors** and *right* = **successors** sets.
Discharge (sub)strings only when $\bigcup$ *left* and $\bigcup$ *right* are **disjoint**.

⟨$_{\varnothing}$a**phɨk**ta$_{\varnothing}$⟩ ⟨$_{\varnothing}$a**phɨk**ti$_{\varnothing}$⟩ ⟨$_{\varnothing}$a**phɨk**tini$_{\varnothing}$⟩
⟨$_{\varnothing}$**phɨk**ta$_{\varnothing}$⟩ ⟨$_{\varnothing}$**phɨk**ti$_{\varnothing}$⟩ ⟨$_{\varnothing}$ham**phɨk**ta$_{\varnothing}$⟩

1: discharge phɨk with *left* $\subseteq \varnothing$ **and** *right* $\subseteq \varnothing$

⟨$_{\varnothing}$a$_{1}$, $_{1}$**t**a$_{\varnothing}$⟩ ⟨$_{\varnothing}$a$_{1}$, $_{1}$**t**i$_{\varnothing}$⟩ ⟨$_{\varnothing}$a$_{1}$, $_{1}$**t**ini$_{\varnothing}$⟩
⟨$_{1}$**t**a$_{\varnothing}$⟩ ⟨$_{1}$**t**i$_{\varnothing}$⟩ ⟨$_{\varnothing}$ham$_{1}$, $_{1}$**t**a$_{\varnothing}$⟩

2: discharge t with *left* $\subseteq \{1\}$ **and** *right* $\subseteq \varnothing$

⟨$_{\varnothing}$**a**$_{1,2}$, $_{1,2}$a$_{\varnothing}$⟩ ⟨$_{\varnothing}$**a**$_{1,2}$, $_{1,2}$i$_{\varnothing}$⟩ ⟨$_{\varnothing}$**a**$_{1,2}$, $_{1,2}$ini$_{\varnothing}$⟩
⟨$_{1,2}$a$_{\varnothing}$⟩ ⟨$_{1,2}$i$_{\varnothing}$⟩ ⟨$_{\varnothing}$ham$_{1,2}$, $_{1,2}$a$_{\varnothing}$⟩

3: discharge a with *left* $\subseteq \varnothing$ **and** *right* $\subseteq \{1,2\}$

⟨$_{1,2,3}$**a**$_{\varnothing}$⟩ ⟨$_{1,2,3}$i$_{\varnothing}$⟩ ⟨$_{1,2,3}$ini$_{\varnothing}$⟩,
⟨$_{1,2}$**a**$_{\varnothing}$⟩ ⟨$_{1,2}$i$_{\varnothing}$⟩ ⟨$_{\varnothing}$h$_{1,2,3}$, $_{\varnothing}$m$_{1,2,3}$, $_{1,2,3}$**a**$_{\varnothing}$⟩

4: discharge a with *left* $\subseteq \{1,2\}$ **and** *right* $\subseteq \varnothing$

Dumi 'to get up'

2s	a-phɨk-t-a
2d	a-phɨk-t-i
2p	a-phɨk-t-ini
3s	phɨk-t-a
3d	phɨk-t-i
3p	ham-phɨk-t-a

(van Driem 1993)

/a-/ ↔ [+2]
/-a/ ↔ [−past]

Preventing to combine inconsistent form occurrences

After removal of phɨk and t (step **3**)

⟨$_{\varnothing}$a$_{1,2}$, $_{1,2}$a$_{\varnothing}$⟩$_{2SG}$ ⟨$_{\varnothing}$a$_{1,2}$, $_{1,2}$i$_{\varnothing}$⟩$_{2DU}$ ⟨$_{\varnothing}$a$_{1,2}$, $_{1,2}$ini$_{\varnothing}$⟩$_{2PL}$
⟨$_{1,2}$a$_{\varnothing}$⟩$_{3SG}$ ⟨$_{1,2}$i$_{\varnothing}$⟩$_{3DU}$ ⟨$_{\varnothing}$ham$_{1,2}$, $_{1,2}$a$_{\varnothing}$⟩$_{3PL}$

$\bigcup$ *left* and $\bigcup$ *right* of $_{\varnothing}$a$_{1,2}$, $_{\varnothing}$a$_{1,2}$, $_{\varnothing}$a$_{1,2}$ from 2SG, 2DU, 2PL:
∅ and {1, 2} ✓ *disjoint*

$\bigcup$ *left* and $\bigcup$ *right* of $_{1,2}$a$_{\varnothing}$, $_{\varnothing}$a$_{1,2}$, $_{\varnothing}$a$_{1,2}$ from 2SG, 2DU, 2PL:
{1, 2} and {1, 2}
⇒ **inconsistent**

$\bigcup$ *left* and $\bigcup$ *right* of $_{1,2}$a$_{\varnothing}$, $_{1,2}$a$_{\varnothing}$, $_{1,2}$a$_{\varnothing}$, from 2SG, 3SG, 3PL:
{1, 2} and ∅ ✓ *disjoint*

Consistency alternatives and extensions

	1	12	2	3
SG	phɨktə		aphɨkta	phɨkta
DU	phɨktɨ	phɨkti	aphɨkti	phɨkti
PL	phɨkktɨ	phɨkkti	aphɨktini	hamphɨkta

- remove markers either from left to right or from right to left?
- identify stem and then remove affixes either inside out or outside in?

- prefer **full strings** over edge-including substrings over internal substrings (Bank & Trommer to appear)
- all other things being equal, prefer **longer** strings
- exclude removals that would introduce a cycle in the transitive closure of the union of all the precedence pairs ⇒ **templatic order**

Incremental learning with cyclic form removal

- iterative local optimization akin to Harmonic Serialism *greedy algorithm*
- competing **segmentation options** are bled by form removal
- no exhaustive search through analytical options segmentations × meaning assignments
- abstracts away from marker hypothesis generation and evaluation

Input: a set $P := \{\langle f, c\rangle, \dots\}$ of cell form f and cell meaning c pairs
a set $\mathbb{U}$ of possible meanings with a partial order $\sqsubseteq$ on them
Output: an initially empty lexicon *Lex*

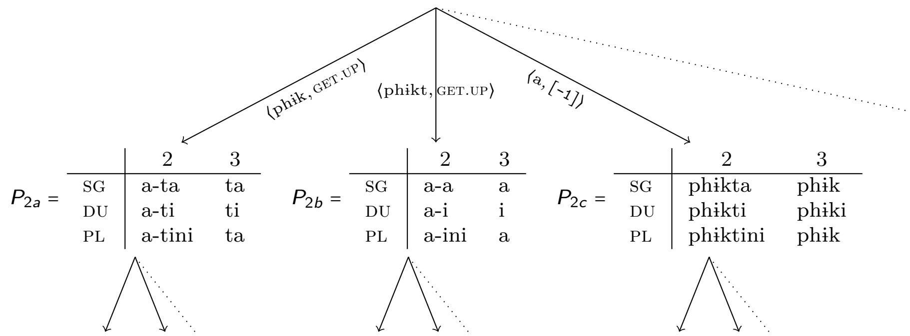
$P \leftarrow \{\langle f, c, i\rangle \mid \langle f, c\rangle \in P, i = 1\}$ ▷ add position index
while $\exists\langle f, c, i\rangle \in P : f \neq \varepsilon$ **do** ▷ there is a cell with a non-empty form
  $H \leftarrow \{\langle f, m\rangle \mid$ form f and meaning m of a marker hypothesis for $P\}$
  $\langle f, m\rangle \leftarrow$ an optimal hypothesis $\in H$
  $P \leftarrow \text{DISCHARGE}(P, f, m)$
  $Lex \leftarrow Lex \cup \{\langle f, m\rangle\}$

Resolving subanalysis options by iterative optimization

P_1 =

	2	3
SG	aphikta	phikta
DU	aphikti	phikti
PL	aphiktini	phikta

$H_1 = \{\langle \text{phik}, \text{GET.UP}\rangle, \langle \text{phikt}, \text{GET.UP}\rangle, \langle \text{t}, [\text{-past}]\rangle, \dots\}$



P_{2a} =

	2	3
SG	a-ta	ta
DU	a-ti	ti
PL	a-tini	ta

P_{2b} =

	2	3
SG	a-a	a
DU	a-i	i
PL	a-ini	a

P_{2c} =

	2	3
SG	phikta	phik
DU	phikti	phiki
PL	phiktini	phik

Evaluating the accuracy of form-meaning pairs

Meaning assignment seeks optimal **patterns** of paradigmatic distributions (Pertsova 2007)

Informal and formal paradigm representation

	[+y]	[−y]
[+x]	*a*	***b***
[−x]	*a*	***ab***

$\{ \langle a, [+x\ +y]\rangle, \langle \boldsymbol{b}, [+x\ -y]\rangle, \langle a, [-x\ +y]\rangle, \langle \boldsymbol{ab}, [-x\ -y]\rangle \}$

Affix hypotheses with accuracy evaluation

	⟨form, meaning⟩	false positives	false negatives	implica. relation	precision	recall
a.	⟨*b*, [−y]⟩	–	–	↔	1	1
b.	⟨*a*, [+y]⟩	–	**yes**	←	1	$\frac{2}{3}$
c.	⟨*a*, [−x]⟩	–	yes	←	1	$\frac{2}{3}$
d.	⟨*a*, []⟩	**yes**	–	→	$\frac{3}{4}$	1
e.	⟨*b*, [+x]⟩	yes	yes	none	$\frac{1}{2}$	$\frac{1}{2}$
f.	⟨*a*, [−y]⟩	yes	yes	none	$\frac{1}{2}$	$\frac{1}{3}$

Avoid false positives ~ optimize precision ~ underinsertion

*OVERINSERTION: Assign * to every paradigm cell where the hypothesis predicts the marker but the marker does not occur

Ainu verbal agreement (Tamura 2000)

	1sg	1pl	2sg	2pl	3sg	3pl	–
1sg			eci-	eci-	ku-	ku-	ku-
1pl			eci-	eci-	ci-	ci-	-as
2sg	en-	un-			e-	e-	e-
2pl	eci-en-	eci-un-			eci-	eci-	eci-
3sg	en-	un-	e-	eci-	Ø-	Ø-	Ø-
3pl	en-	un-	e-	eci-	Ø-	Ø-	Ø-

Meaning assignment for *eci-* with precision-favoring ranking

	*OVERINS	*UNDERINS
☞ a. ***eci-*:[+2 +pl]**		$*_2$ **(1:2sg)**
b. *eci-*:[+2]	$*_7$! (2sg,2sg:X,3:2sg)	
c. *eci-*:[Acc +2]	$*_2$! (3:2sg)	$*_5$ (2pl,2pl:X)

Avoid false negatives ~ optimize recall ~ overinsertion

*UNDERINSERTION: Assign $*$ to every paradigm cell where the marker occurs but the hypothesis does not predict it

Ainu verbal agreement

	1sg	1pl	2sg	2pl	3sg	3pl	–
1sg			eci-	eci-	ku-	ku-	ku-
1pl			eci-	eci-	ci-	ci-	-as
2sg	en-	un-			e-	e-	e-
2pl	eci-en-	eci-un-			eci-	eci-	eci-
3sg	en-	un-	e-	eci-	Ø-	Ø-	Ø-
3pl	en-	un-	e-	eci-	Ø-	Ø-	Ø-

Meaning assignment for *eci-* with recall-favoring ranking

	*UNDERINS	*OVERINS
a. *eci*-:[+2 +pl]	$*_2$ (1:2sg)	
☞ b. ***eci*-:[+2]**		**$*_7$! (2sg,2sg:X,3:2sg)**
c. *eci*-:[Acc +2]	$*_5$ (2pl,2pl:X)	$*_2$! (3:2sg)

Consistent linear order subanalysis learner (main loop)

MARKERS() provides (sub)strings with their set of occurrences, one cell per occurrence
HYPOTHESES() provides consistent marker hypotheses with accuracy-determining info
DISCHARGE() removes material from the paradigm keeping track of precedence relations

Input: a set $P := \{\langle f, c\rangle, \dots\}$ of cell form f and cell meaning c pairs
a set $\mathbb{U}$ of possible meanings with a partial order $\sqsubseteq$ on them
a morpheme hypothesis evaluation function EVALUATE
Output: an initially empty lexicon *Lex*
an initially empty linear precedence relation *Ord*

```
P ← {⟨f, c, i⟩ | ⟨f, c⟩ ∈ P, i = 1}                          ▷ form, meaning, position index
LR ← {⟨c, i, L, R⟩ | ⟨f, c, i⟩ ∈ P, L = ∅, R = ∅}             ▷ cell, index, predecessors, successors
while ∃⟨f, c, i⟩ ∈ P : f ≠ ε do                              ▷ there is a cell with a non-empty form
    M ← MARKERS(P)                                           ▷ s ↦ cell-consistent contexts
    H ← HYPOTHESES(M, 𝕌, LR, Ord)
    ⟨(f, C, L, R), m, (TP, FP, FN, TN)⟩ ← EVALUATE(H)         ▷ optimal hypothesis
    D ← {⟨c, i⟩ | ⟨c, i, b⟩ ∈ C, c ∈ TP}                      ▷ true positive occurrence coordinates
    ⟨P, LR⟩ ← DISCHARGE(P, LR, f, m, D)
    Ord ← Ord ∪ {⟨l, (f, m)⟩ | l ∈ L} ∪ {⟨(f, m), r⟩ | r ∈ R}   ▷ prede-/successors
    Lex ← Lex ∪ {⟨f, m⟩}
```


(Not) learning linearization cycles: algorithm run

1 $P_1 = \{\langle \textbf{rischi}, \text{INCH.ASST}, 1, \varnothing, \varnothing\rangle, \langle \text{schina}, \text{ASST.RECP}, 1, \varnothing, \varnothing\rangle, \langle \text{nari}, \text{RECP.INCH}, 1, \varnothing, \varnothing\rangle\}$
$Lex_1 = \varnothing$
$Ord_1 = \varnothing$

2 $P_2 = \{\langle \textbf{schi}, \text{INCH.ASST}, 2, \{\text{ri:INCH}\}, \varnothing\rangle, \langle \textbf{schi}\text{na}, \text{ASST.RECP}, 1, \varnothing, \varnothing\rangle, \langle \text{na}, \text{RECP.INCH}, 1, \varnothing, \{\text{ri:INCH}\}\rangle\}$
$Lex_2 = \{\ \textbf{ri:inch}\}$
$Ord_2 = \varnothing$

3 $P_3 = \{\langle \textbf{na}, \text{ASST.RECP}, 1, \{\text{schi:ASST}\}, \varnothing\rangle, \langle \text{na}, \text{RECP.INCH}, 1, \varnothing, \{\text{ri:INCH}\}\rangle\}$
$Lex_3 = \{\ \text{ri:INCH}, \text{schi:ASST}\}$
$Ord_3 = \text{ri:INCH} < \text{schi:ASST}$

4 $P_4 = \{\langle \textbf{na}, \text{RECP.INCH}, 1, \varnothing, \{\text{ri:INCH}\}\rangle\}$
$Lex_4 = \{\ \text{ri:INCH}, \text{schi:ASST}, \text{na:ASST.RECP}\}$
$Ord_4 = \text{ri:INCH} < \text{schi:ASST}$, $\text{schi:ASST} < \text{na:ASST.RECP}$

5 $P_5 = \varnothing$
$Lex_5 = \{\ \text{ri:INCH}, \text{schi:ASST}, \text{na:ASST.RECP}, \text{na:RECP.INCH}\}$
$Ord_5 = \text{ri:INCH} < \text{schi:ASST}$, $\text{schi:ASST} < \text{na:ASST.RECP}$, $\text{na:RECP.INCH} < \text{ri:INCH}$

Nontransitivity in Chumbivilcas Quechua derivation affixes

INCH.ASST	ri-schi
ASST.RECP	schi-na
RECP.INCH	na-ri

(Muysken 1988)
see also Ryan (2010)

Preventing to combine cycle-introducing form occurrences

After removal of ri **and** schi **(loop 3)**

$P_3 = \{\langle \text{na}, \text{ASST.RECP}, 1, \{\text{schi:ASST}\}, \varnothing\rangle, \langle \text{na}, \text{RECP.INCH}, 1, \varnothing, \{\text{ri:INCH}\}\rangle\}$
$Lex_3 = \{\text{ri:INCH}, \text{schi:ASST}\}$
$Ord_3 = \text{ri:INCH} < \text{schi:ASST}$

Consider marker hypothesis na:RECP

Linear relations from $\bigcup L \times \bigcup R$ of na:RECP
schi:ASST < ri:INCH
contradicts $Ord_3^{\,+}$ — *asymmetry violation* $\begin{bmatrix} 0&1&0\\ 1&0&0\\ 0&0&0 \end{bmatrix}$
⇒ **introduces cycle**

Results after loop 5

$Lex_5 = \{\text{ri:INCH}, \text{schi:ASST}, \text{na:ASST.RECP}, \text{na:RECP.INCH}\}$ $\quad Ord_5 = \begin{bmatrix} 0&1&0&0\\ 0&0&1&0\\ 0&0&0&0\\ 1&0&0&0 \end{bmatrix}$

Costs and effects of transitive consistency

Homophony and extended exponence of the cycle-breaking na-variants
more markers than possible without linearization, less general markers, reduced accuracy

Locality of cycle detection effects generalization asymmetry in learning
derives perfect generalizations as early as possible, less accurate residue at later stages

Ensuring consistency: HYPOTHESES$(M, \mathbb{U}, LR, Ord)$

A	$L \cap R = \varnothing$ and $R \times L \cap Ord^{+} = \varnothing$	asymmetric relation with asymmetric transitive closure	templatic linearization
B	$L \cap R = \varnothing$	asymmetric relation	precedence constraints
C	*no restriction*	no generalization on morpheme order	memorize cell-chains

Output: an initially empty set $H := \{\langle (f, C, L, R), m, (TP, FP, FN, TN)\rangle, \dots\}$ of form, contexts, predecessors, successors, meaning, true positives, false positives, false negatives, true negatives

for all $\langle (f, C), m\rangle \in (M \times \mathbb{U})$ **do** ▷ occurrences meaning pairings
$L \leftarrow \bigcup\{L \mid \langle c, i\rangle \in C, m \sqsubseteq c, \langle c, i, L, R\rangle \in LR\}$ ▷ predecessors of subsumed
$R \leftarrow \bigcup\{R \mid \langle c, i\rangle \in C, m \sqsubseteq c, \langle c, i, L, R\rangle \in LR\}$ ▷ successors of subsumed
if $L \cap R = \varnothing$ **and** $R \times L \cap Ord^{+} = \varnothing$ **then** ▷ consistent and cycle-free
$C' \leftarrow \{\langle c, i\rangle \in C \mid \langle c, i'\rangle \in P, \langle c, i', L_c, R_c\rangle \in LR, L_c \cap R = \varnothing, R_c \cap L = \varnothing,$
$(R_c \times L \cup R \times L_c) \cap Ord^{+} = \varnothing\}$ ▷ consistent and cycle-free
$T \leftarrow \{c \mid \langle c, i\rangle \in C'\}$ ▷ target cells
$P \leftarrow \{c \in \mathbb{U}_c \mid m \sqsubseteq c\}$ ▷ positives
$N \leftarrow \{c \in \mathbb{U}_c \mid m \not\sqsubseteq c\}$ ▷ negatives
$H \leftarrow H \cup \{\langle (f, C', L, R), m, (P \cap T, P \setminus T, N \cap T, N \setminus T)\rangle\}$

Incremental templatic analysis of Dumi phɨkni 'to get up'

Marker evaluation *UNDERINSERTION ≫ *OVERINSERTION ≫ MAXCOVERAGE

1	1	12	2	3
SG	phɨktə		aphɨkta	phɨkta
DU	phɨktɨ	phɨkti	aphɨkti	phɨkti
PL	phɨkktɨ	phɨkkti	aphɨktini	hamphɨkta

phɨk:[]

2	1	12	2	3
SG	tə		a-ta	ta
DU	tɨ	ti	a-ti	ti
PL	ktɨ	kti	a-tini	ham-ta

t:[]

3	1	12	2	3
SG	ə		a-a	a
DU	ɨ	i	a-i	i
PL	k-ɨ	k-i	a-ini	ham-a

ɨ:[+1 −2 −sg]

4	1	12	2	3
SG	ə		a-a	a
DU		i	a-i	i
PL	k	k-i	a-ini	ham-a

k:[+1 +pl]

5	1	12	2	3
SG	ə		a-a	a
DU		i	a-i	i
PL		i	a-ini	ham-a

ham:[+3 +pl]

6	1	12	2	3
SG	ə		a-a	a
DU		i	a-i	i
PL		i	a-ini	a

a:[−1 +2]

7	1	12	2	3
SG	ə		a	a
DU		i	i	i
PL		i	ini	a

ini:[−1 +2 +pl]

8	1	12	2	3
SG	ə		a	a
DU		i	i	i
PL		i		a

ə:[+1 +sg]

9	1	12	2	3
SG			a	a
DU		i	i	i
PL		i		a

a:[−1 −du]

10	1	12	2	3
SG				
DU		i	i	i
PL		i		

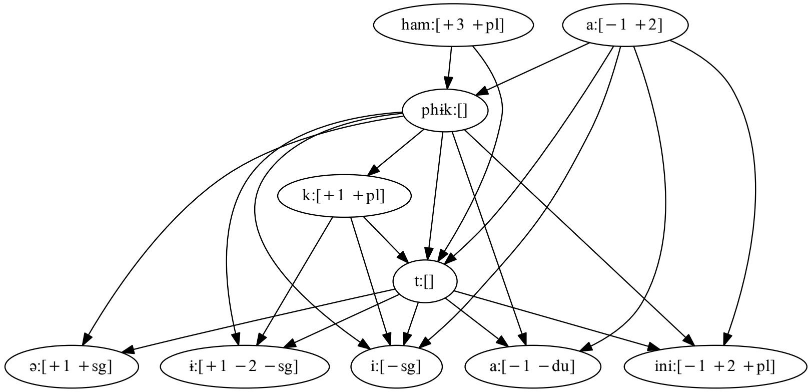
i:[−sg]

11	1	12	2	3
SG				
DU				
PL				

a:[−1 −du] has 1 *false positive*
i:[−sg] has 4 *false positives*

Resulting paradigm and morpheme precedence relation

	1	12	2	3
SG	phɨk-t-ə		a_1-phɨk-t-a_2	phɨk-t-a_2
DU	phɨk-t-ɨ	phɨk-t-i	a_1-phɨk-t-i	phɨk-t-i
PL	phɨk-k-t-ɨ	phɨk-k-t-i	a_1-phɨk-t-ini	ham-phɨk-t-a_2

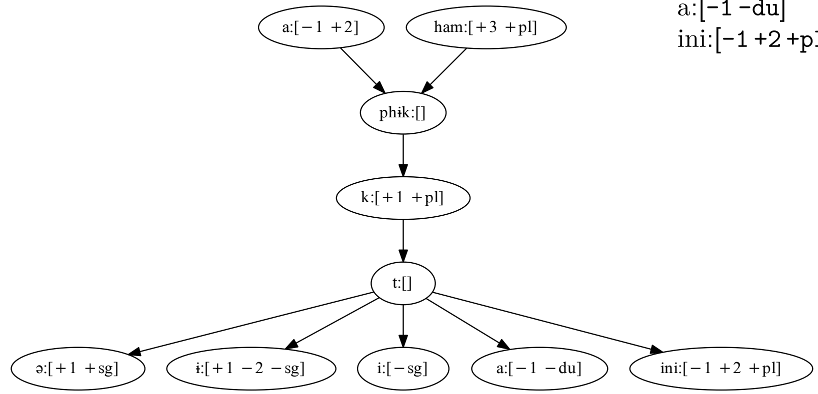


Convert precedence relation to morpheme template

extrinsic ordering

1	**2**	**3**	**4**	**5**
a:[−1 +2] ham:[+3 +pl]	phɨk:[]	k:[+1 +pl]	t:[]	ə:[+1 +sg] ɨ:[+1 −2 −sg] i:[−sg] a:[−1 −du] ini:[−1 +2 +pl]

transitive reduction



Conclusion

- subanalysis for inflectional grammars with extrinsic order like Anderson (1992) can be learned with *local* optimization and order consistency checking
- learner biases can be used to predict the complexity/markedness of certain linearization patterns like nontransitivity and classify them
- modelling learning of morphological grammar(s) uncovers their logical structure and helps to make their comparison more empirically informed
- how can an interaction between subanalysis and reordering or correspondences between linearization and blocking be learned in a local way?

Anderson, S. R. (1992). *A-Morphous Morphology*. Cambridge: Cambridge University Press.

Bank, S. & Trommer, J. (2013). *Inflectional learning as local optimization*. Manuscript, University of Leipzig.

Bank, S. & Trommer, J. (to appear). Paradigm learning and subanalysis complexity. *Proceedings of the 48th Annual Meeting of the Chicago Linguistic Society*.

van Driem, George (1993). *A Grammar of Dumi*. Berlin and New York: Mouton de Gruyter.

Halle, M. & Marantz, A. (1993). Distributed Morphology and the pieces of inflection. In Hale, K. & Keyser, S. J., editors, *The View from Building 20*, pages 111–176.Cambridge MA: MIT Press.

Kahn, A. B. (1962). Topological sorting of large networks. *Communications of the ACM* 5, 5(11): 558–562.

McCarthy, J. (2010). An introduction to harmonic serialism. *Language and Linguistics Compass*, 4(10): 1001–1018.

Muysken, P. (1988). Affix order and interpretation: Quechua. Martin Everaert, Arnold Evers, Riny Huybregts, and Mieke Trommelen, editors, *Morphology and modularity*. Dordrecht: Foris, 259–279.

Paster, M. (2005). Pulaar verbal extensions and phonologically driven affix order. In G. Booij & J. van Marle (eds.), *Yearbook of morphology* 2005. Dordrecht: Springer, 155–199.

Pertsova, K. (2007). *Learning Form-Meaning Mappings in Presence of Homonymy*. PhD thesis, UCLA.

Ryan, K. M. (2010). Variable affix order: Grammar and learning. *Language*, 86: 758–791.

Stump, G. T. (2001). *Inflectional Morphology*. Cambridge: Cambridge University Press.

Tamura, Suzuko (2000), *The Ainu Language*. Tokyo: Sanseido.