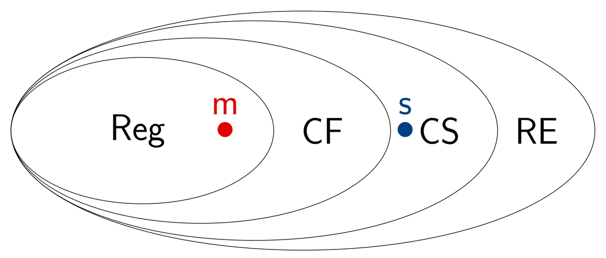
Distributed Morphology as a regular relation

THE UNIVERSITY OF CHICAGO

Marina Ermolaeva
mermolaeva@uchicago.edu
Daniel Edmiston
danedmiston@uchicago.edu

Introduction

- **Implicit claim of DM**: syntax all the way down; morphology over trees
- DM can in spirit be treated as a tree transducer, but how similar are morphology and syntax?
- Morphology appears **(at most) regular**; can we get this for free?
- Work in NLP treats morphology with finite-state methods (e.g. Karttunen et al. 1992); syntax cannot be done this way (Shieber 1985)



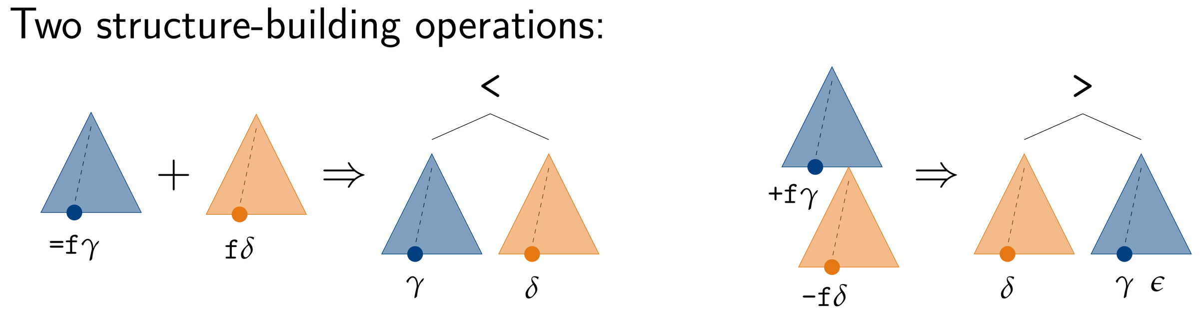
(morphology vs. syntax)

- We reshape Distributed Morphology (DM, Halle & Marantz 1993) to operate over strings rather than trees
- **Principle change**: Flatten structure to strings **before** morphology

Syntax-Morphology interface

Framework: Minimalist Grammars (MGs, Stabler 1997)

- A set of **syntactic features** *Syn*
- A **lexicon**: $Lex \subset \Sigma^* \, Syn^*$, where Σ is a set of pronounced segments
- Two structure-building operations:



(*merge*) (*move*)

Modification 1: syntax assembles morphological words (separated by #)

- Lowering and Head Movement as *merge* with concatenation of heads
- Mirror Theory (Brody 1997, Kobele 2002): **strong** and **weak** nodes
- Three subtypes of selector features:
  =f (normal *merge*)
  =>f (strong node; *merge* + Head Movement)
  <=f (weak node; *merge* + Lowering)

Modification 2: full separation of syntax and phonology

- **Feature structures**: $FS = \mathcal{P}(M) \times (\Sigma \cup \{\epsilon, None\})$, where *none* denotes the "placeholder" exponent and M is a finite set of morphological features; For $s = \langle x, y \rangle \in FS$, $feat(s) = x$ and $exp(s) = y$.
- **Redefining lexicon**: $Lex \subset \{s \mid s \in FS \;\&\; exp(s) = None\} \; Syn^*$

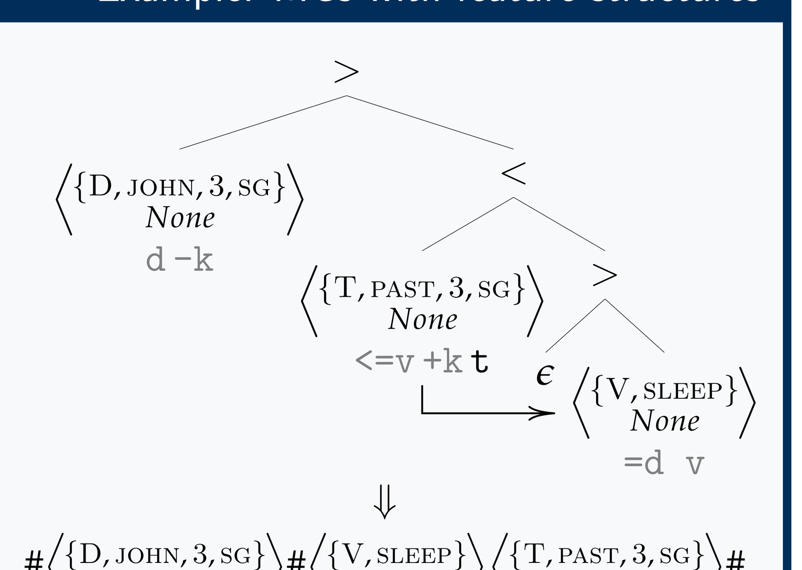
Example: MGs with feature structures

Lexicon:

$\left\langle {\{D, JOHN, 3, SG\} \atop None} \right\rangle$ :: d -k

$\left\langle {\{V, SLEEP\} \atop None} \right\rangle$ :: =d v

$\left\langle {\{T, PAST, 3, SG\} \atop None} \right\rangle$ :: <=v +k t



Morphological rules

- Morphological rules operate on **underspecified feature structures**: $FS_U = \mathcal{P}(M) \times (\Sigma \cup \{\epsilon, None, ?\})$, where ? stands for "any exponent"
- Kaplan & Kay 1994:
  - Rewriting rules that do not overwrite their own output define **regular relations** over strings;
  - Simultaneous application as **batch rules**, ordered rules as **composition** of regular relations.

Analogy: phonological rules

A rule in feature matrix notation is equivalent to a set of rules over atomic symbols. For instance, $[-syl, +voi] \to [-voi] \;/\; \_\_ \; [-voi]$ abbreviates multiple rules: $b \to p \;/\; \_\_ \, (p \mid t \mid k)$, $d \to t \;/\; \_\_ \, (p \mid t \mid k)$, ...

- **Rule format**:

structural description		structural change	/	left context		right context
A	$\to$	B	/	C	__	D
$\Vert$ $A_1,\ldots,A_m \in FS_U$		$\Vert$ $B_1,\ldots,B_n \in FS_U$		(regular expressions over $FS_U \cup \{\#\}$)		

- Vocabulary Insertion (VI): for $1 \le i \le |A|$, $1 \le j \le |B|$, $exp(A_i) = None$, $exp(B_j) \ne None$, $feat(A_i) = feat(B_j)$; $|A| = 1$, $|B| \ge 1$.
- Readjustment: for $1 \le i \le |A|$, $1 \le j \le |B|$, $exp(A_i) \ne None$, $exp(B_j) \ne None$, $feat(A_i) = feat(B_j)$.

Example: root alternations

Vocabulary Insertion:

$\left\langle {\{SLEEP\} \atop None} \right\rangle \to \left\langle {\{SLEEP\} \atop s} \right\rangle \left\langle {\{SLEEP\} \atop l} \right\rangle \left\langle {\{SLEEP\} \atop i:} \right\rangle \left\langle {\{SLEEP\} \atop p} \right\rangle$

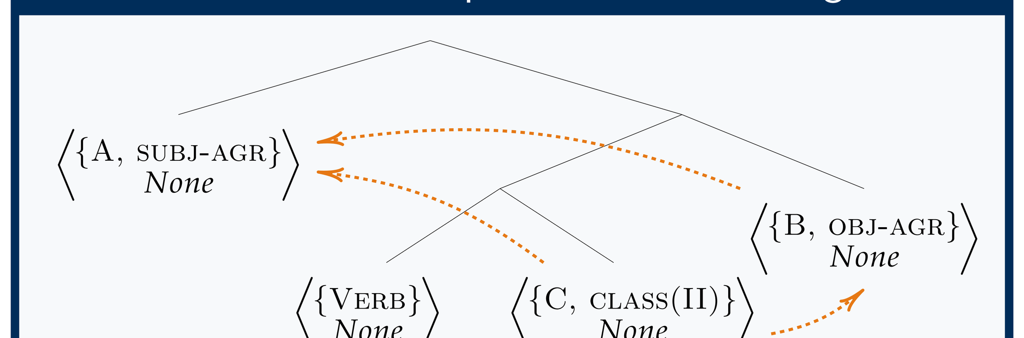
Readjustment:

$\left\langle {\{SLEEP\} \atop i:} \right\rangle \to \left\langle {\{SLEEP\} \atop \varepsilon} \right\rangle / \_\_ \left[\left\langle {\{SLEEP\} \atop ?} \right\rangle \right]^* \left\langle {\{PAST\} \atop ?} \right\rangle$

Cyclicity and Rewriting

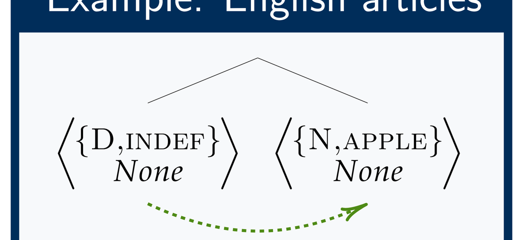
- Bobaljik 2000:
  - **Cyclicity**: VI starts at the root and proceeds outwards
  - **Rewriting**: VI deletes morphosyntactic features it expresses
  - Outward sensitivity to morphosyntactic features; inward sensitivity to morphophonological features

Example: Itelmen verbal agreement



(adapted from Bobaljik 2000)

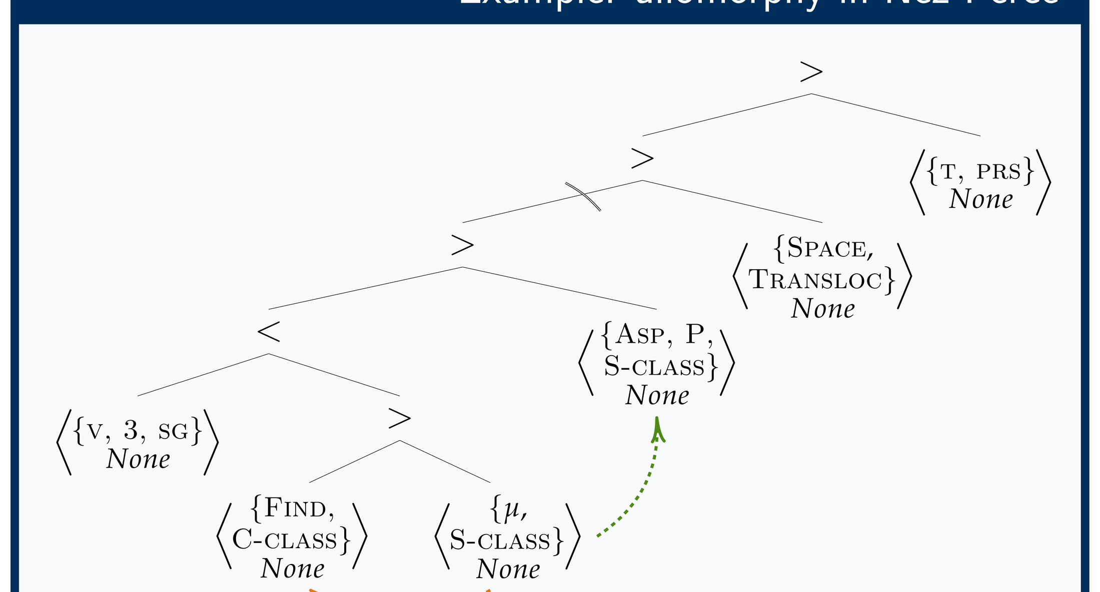
Example: English articles



- Simulating cyclicity with rule ordering: **Hierarchy of Projections** (HoP, Adger 2003)
  Clausal: V ⟨ v ⟨ (Pass) (Prog) ⟨ (Perf) ⟨ (Neg) ⟨ T ⟨ C
  Nominal: N ⟨ n ⟨ (Poss) ⟨ D

- Are these constraints **desirable**?
- Multiple counter-examples
  Svenonius 2012:
  – reconsider Cyclicity?
  Gribanova & Harizanov 2017:
  – eliminate Rewriting?
  Deal & Wolf 2017:
  – weaken Rewriting to Monotonicity: VI strictly adds information
  – inside-out insertion of cycles (≈phases)
  – VI inside cycles in any order
- Cyclicity as **ordered rules over strings**: follow HoP, allowing for mismatches

Example: allomorphy in Nez Perce



(adapted from Deal & Wolf 2017)

Inching toward Word & Paradigm

- Karttunen 2003:
  - Paradigm Function Morphology (PFM, Stump 2001) can be restated as **regular relations**
  - PFM can be viewed as series of ordered rewrite rules ... and transformed into FST via rule composition

Example: Swahili verbs in PFM

ni-ta-taka
1SG-FUT-WANT

REALIZE(⟨TAKA, σ:{1sg fut}⟩) =
[RULE II: [RULE I: STEM(⟨TAKA, σ⟩)]]
a. [I: ⟨*taka*, σ: {1sg fut}⟩] = ⟨*tataka*, σ⟩
b. [II: ⟨*tataka*, σ: {1sg fut}⟩] = ⟨*nitataka*, σ⟩
⇓
RULE II ∘ RULE I ∘ STEM(⟨TAKA, σ⟩)

(adapted from Bonami & Stump 2016, Karttunen 2003)

- PFM rules encode prefix/suffix distinction
- In string-based DM, ordering is decided by the string fed to morphology

Example: Swahili verbs in string-based DM

Input:
$\left\langle {\{D, 1, SG\} \atop None} \right\rangle \left\langle {\{T, FUT\} \atop None} \right\rangle \left\langle {\{V, TAKA\} \atop None} \right\rangle$

Vocabulary Insertion:
$\left\langle {\{TAKA\} \atop None} \right\rangle \to \left\langle {\{TAKA\} \atop t} \right\rangle \left\langle {\{TAKA\} \atop a} \right\rangle \left\langle {\{TAKA\} \atop k} \right\rangle \left\langle {\{TAKA\} \atop a} \right\rangle$
$\left\langle {\{FUT\} \atop None} \right\rangle \to \left\langle {\{FUT\} \atop t} \right\rangle \left\langle {\{FUT\} \atop a} \right\rangle$
$\left\langle {\{1SG\} \atop None} \right\rangle \to \left\langle {\{1SG\} \atop n} \right\rangle \left\langle {\{1SG\} \atop i} \right\rangle$

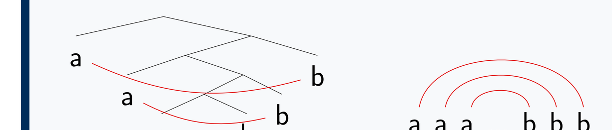
- However, external dress hides essence of each formalization: **both are faithfully reducible to FSTs**

Discussion

- Explanation for regular-ness of morphology
  – structures flatten between syntax and morphology
  – super-regular syntax, (sub)regular morphology

Abstract example: CF morphology

Unbounded nested dependencies: **incompatible** with regular DM



Any natural language examples?

- Restriction to regular relations instead of limiting size of windows over trees (e.g. spans, Merchant 2015)
  – formal grounding for limiting context
  – properties well understood, including efficient parsing and generation
- Apparent cyclicity effects treated as rule ordering
  – enough flexibility is retained to handle direct counter-examples
- Elimination of trees + reliance of rule orderings moves DM closer to W&P
  – formalization shows frameworks more alike than different

References: Adger, D. 2003. Core syntax: a minimalist approach. ▪ Bobaljik, J. D. 2000. The ins and outs of contextual allomorphy. ▪ Bonami, O., and G. Stump. 2016. Paradigm function morphology. ▪ Brody. M. 1997. Mirror theory. ▪ Deal, A. R., and M. Wolf. 2017. Outward-sensitive phonologically-conditioned allomorphy in Nez Perce. ▪ Gribanova, V., and B. Harizanov. 2017. Locality and directionality in inward-sensitive allomorphy: Russian and Bulgarian. ▪ Halle, M., and A. Marantz. 1993. Distributed Morphology and the pieces of inflection. ▪ Kaplan, R. M., and M. Kay. 1994. Regular models of phonological rule systems. ▪ Karttunen, L., R. M. Kaplan, and A. Zaenen. 1992. Two-level morphology with composition. ▪ Kobele, G. M. 2002. Formalizing Mirror Theory. ▪ Karttunen, L. 2003. Computing with Realizational Morphology. ▪ Merchant, J. 2015. How much context is enough? ▪ Shieber, S. M. 1985. Evidence against the context-freeness of natural language. ▪ Stabler, E. P. 1997. Derivational minimalism. ▪ Stump, G. T. 2001. Inflectional Morphology: A theory of paradigm structure.