Distributed Morphology as a regular relation

Marina Ermolaeva and **Daniel Edmiston**
University of Chicago
{mermolaeva,danedmiston}@uchicago.edu

1 Introduction

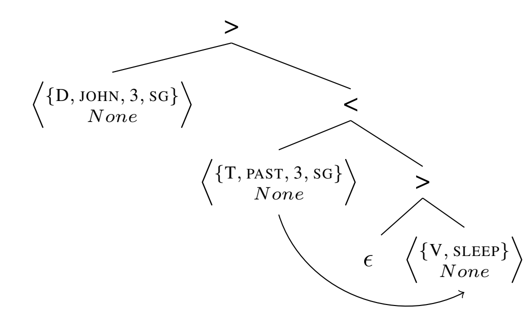
This research reorganizes the Distributed Morphology (DM, (Halle and Marantz, 1993)) framework to work over strings. That the morphological module should operate over strings is desirable, since it is assumed that most (arguably all) morphological processes can be modelled with regular languages (Karttunen et al. 1992). As is, DM typically operates on binary trees, with the syntax- morphology interface implicitly treated as a tree-transducer. We contend that using (binary) trees is overpowered, and predicts patterns which are unattested in natural language (e.g. iterable nested dependencies). If however, we restrict the morphological component to working on strings, we correctly predict that morphology can be modelled with regular string languages, and so we treat the morphological component as a finite-state string-transducer, i.e. as a regular relation.

Assuming the Y-model, standard DM operating on trees presumes that the *flattening* of the derivation for PF takes place post-morphology. We push this flattening of the structure to *above* the morphological module, between the syntax and morphology.

2 Syntax-morphology interface

We use a modified version of Minimalist Grammars (Stabler, 1997) as our syntactic framework. We enhance the formalism with the notion of *node strength* imported from Mirror Theory (Brody, 1997; Kobele, 2002) to implement both Head Movement and Lowering as mechanisms for constructing morphological words. Following (Trommer, 1999), we assume that the basic unit of syntactic computation is a *feature structure* (FS). An FS F is defined as a pair $\langle M, E\rangle$, where the feature bundle $M = feat(F)$ is a subset of some finite set of features (including syntactic category labels) and the phonological exponent $E = exp(F) \in (\Sigma \cup \{None, \epsilon\})$, where Σ is a finite set of phonemes. The default exponent of each syntactic unit is $None$, a place-holder to be replaced by vocabulary insertion (VI) in the morphological module; $None$ is distinct from the null exponent ϵ. An example derivation and its yield serving as input to the morphology are shown in (1); we assume that word boundaries (denoted by #) are part of the linearized representation.

(1)



$$\Downarrow$$

$$\left\langle \substack{\{\text{D, JOHN, 3, SG}\} \\ None} \right\rangle \# \left\langle \substack{\{\text{V, SLEEP}\} \\ None} \right\rangle \# \left\langle \substack{\{\text{T, PAST, 3, SG}\} \\ None} \right\rangle$$

3 Morphology as regular relations

As in standard DM, our operations are context-dependent. Context-dependent *rewriting rules* have the form $A \rightarrow B/C\_\_D$, where A, B, C, D are regular expressions over some alphabet. With a few caveats, such rules (and ordered finite sequences thereof) have been shown to define regular relations on strings (Kaplan and Kay, 1994). Assuming that morphology can be modelled by regular relations, it

Proceedings of the Society for Computation in Linguistics (SCiL) 2018, pages 178-181.
Salt Lake City, Utah, January 4-7, 2018

should be possible to state morphological rules in a way compatible with this format. We take A, B to be sequences of FSs and C, D to be regular expressions over FSs. As a proof of concept, we define VI and Readjustment as follows. VI rewrites a single node as a sequence of FSs which can be thought of as *morphophonemes* and can be further manipulated by Readjustment rules:

(2) A rule r of the form $A \rightarrow B/C\_\_D$ is a *VI* rule iff:
$|A| = 1, |B| \geq 1$;
$exp(A_i) = None$ for $1 \leq i \leq |A|$;
$exp(B_j) \neq None$ for $1 \leq j \leq |B|$;
$\bigcup_{1 \leq i \leq |A|} feat(A_i) = \bigcup_{1 \leq j \leq |B|} feat(B_j)$ (r is *feature-preserving*);
$feat(A_1) = ... = feat(A_{|A|}) = feat(B_1) = ... = feat(B_{|B|})$ (r is *set-preserving*).

(3) A rule r of the form $A \rightarrow B/C\_\_D$ is a *Readjustment* rule iff:
$exp(A_i) \neq None$ for $1 \leq i \leq |A|$;
$exp(B_j) \neq None$ for $1 \leq j \leq |B|$;
r is feature-preserving and set-preserving.

For instance, the root alternation in *sleep/slept* can be captured by the VI and Readjustment rules in (4), represented as rewriting rules in (5), with ? as shorthand for "any exponent":

(4) [SLEEP] $\rightarrow$ **sli:p**
i: $\rightarrow$ **ε** / X__Y [PAST], where X__Y $\in$ {CREEP, SLEEP, WEEP, ...}

(5) $\left\langle \substack{\{\text{SLEEP}\} \\ None} \right\rangle \rightarrow \left\langle \substack{\{\text{SLEEP}\} \\ \text{s}} \right\rangle \left\langle \substack{\{\text{SLEEP}\} \\ \text{l}} \right\rangle \left\langle \substack{\{\text{SLEEP}\} \\ \text{i:}} \right\rangle \left\langle \substack{\{\text{SLEEP}\} \\ \text{p}} \right\rangle$
$\left\langle \substack{\{\text{SLEEP}\} \\ \text{i:}} \right\rangle \rightarrow \left\langle \substack{\{\text{SLEEP}\} \\ \varepsilon} \right\rangle / \_\_ \left[\left\langle \substack{\{\text{SLEEP}\} \\ ?} \right\rangle \right]^* \left\langle \substack{\{\text{PAST}\} \\ ?} \right\rangle$

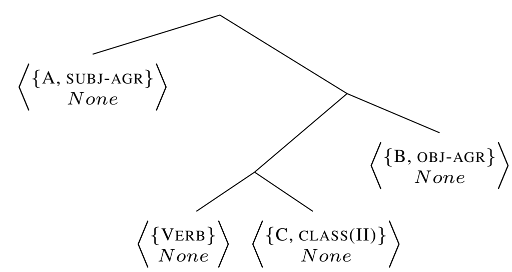
Importantly, morphological rules – both in traditional DM and in our formalism – are stated over *underspecified* feature structures: they apply to any item whose features match the structural description. Put in practice, each underspecified rule like the ones in (5) can be converted, via a well-defined procedure, into a set of *instances* – regular relations over the alphabet of fully specified FSs. This alphabet is guaranteed to be finite, as the set of features and Σ are finite by definition. Relations encoding rules are then composed into a single regular relation (transducer), following the procedure laid out in (Kaplan and Kay, 1994).

4 Cyclicity and Rewriting

As we have proposed a move away from tree-structures and toward strings, it is worth revisiting one of the primary motivations behind tree structures in the literature, namely Cyclicity/Rewriting (Bobaljik, 2000).

(6) **Cyclicity**: VI starts at the root and proceeds outwards
Rewriting: VI deletes morphosyntactic features it expresses

The two constraints in (6) make strong predictions concerning possible contexts for VI rules. In particular, Cyclicity predicts that phonologically conditioned allomorphy can only be inward-sensitive, and Rewriting only allows outward-sensitivity for morphosyntactic features. One motivating example behind Cyclicity comes from Itelmen, where it appears that dependencies are sensitive to embedding-depth rather than linear ordering. Consider the following schematic of an Itelmen example adapted from Bobaljik 2000.

(7)


In (7), the surface form B is dependent upon the morphosyntantic features of A, and the the form of C is dependent upon the features of B and A. Given the tree structure above, there is a clear pattern of more-embedded items being dependent on less-embedded items morphosyntactically; however, such dependencies cannot be reduced to linear order.

Such effects in morphology are handled by applying VI rules from the root outwards – which is impossible after linearization. Fortunately, using transducer composition to build our morphology out of individual rules imposes a strict rule ordering. For this ordering, we can follow the standard order of projections, working from bottom to top (e.g. V→T→C in the clausal domain). In this way, we reproduce Cyclicity effects over strings. For example, with cases of phonologically conditioned allomorphy like English DPs of the form $[_{DP}D[_{NP}N]]$, we can capture the *a*/*an* allomorphy by having rules that apply to FSs containing N precede rules that target FSs containing D. Likewise for examples like (7) above, arranging VI in the C→B→A order yields the right pattern with no need for tree structure.

Thus, trees are not necessary to produce Cyclicity effects. However, it is not even clear that Cyclicity/Rewriting are desirable constraints for our grammar, as there are numerous apparent counter-examples to Rewriting (Gribanova and Harizanov, 2017) and Cyclicity (Svenonius, 2012).

An interesting case is Nez Perce (Deal and Wolf, 2017). The verb suffix denoted as μ exhibits allomorphy conditioned by the *class feature* of the root, as exemplified by the so-called 'S-class' verb *eat* in (8) and the 'C-class' verb *find* in (9). This is an instance of inward sensitivity and, since class membership is a morphosyntactic feature, a counter-example to Rewriting:

(8) 'aw- 'ap- **an'y-** o'- qa
3OBJ- eat- μ- PROSP.ASP- PST

(9) 'aw- 'yáx̂- **nan'y-** o'- qa
3OBJ- find- μ- PROSP.ASP- PST

Moreover, μ appears in the *long form* (9) if the following suffix is smaller than a sequence of one consonant and one vowel (CV). The *short form* (10) is realized when μ is followed by a CV (or longer) morpheme. Suffixes that appear to the right of μ, such as Aspect, attach outside μ, making this alternation a case of outward-looking phonologically-conditioned allomorphy – contrary to Cyclicity.

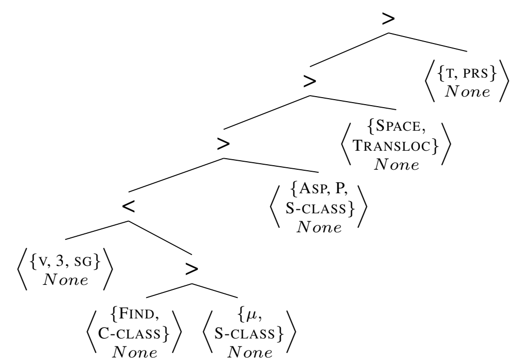
(10) 'aw- 'yáx̂- **nay'-** sa- qa
3OBJ- find- μ- IPFV.ASP- PST

Importantly, this alternation is conditioned only by the suffix immediately following μ; if that suffix is realized by a null exponent (11), μ will appear in the long form, even if the material further to the right is CV in form:

(11) 'aw- 'yáx̂- **nan'i-** ∅- ki- ∅
3OBJ- find- μ- P.ASP- TRANSLOC- PRS

Using our feature structure notation, the structure of (11) can be represented as follows:

(12)



The solution proposed by Deal and Wolf (2017) involves weakening of both Rewriting and Cyclicity. The former is replaced with Monotonicity: instead of deleting morphosyntactic features, VI *strictly adds* information. The latter is restated in terms of cycles, roughly corresponding to syntactic phases. Within a cycle, lexical items can be inserted in any order, constrained only by their phonological context requirements. Cycles themselves are spelled out in an inside-out fashion, one at a time. As there is a phase boundary above Aspect, the realization of μ can depend on the phonological exponent of Aspect, but not of a morpheme belonging to the higher cycle.

This can be straightforwardly translated into our formalism. The contexts for *-nay'* (10) and *-nan'i* (11) are shown in (13a) and (13b) respectively:

(13) a. $\left\langle \begin{smallmatrix}\{\text{C-CLASS}\}\\ ?\end{smallmatrix} \right\rangle \text{—} \left\langle \begin{smallmatrix}\{\,\}\\ \text{C}\end{smallmatrix} \right\rangle \left\langle \begin{smallmatrix}\{\,\}\\ \text{V}\end{smallmatrix} \right\rangle$

b. $\left\langle \begin{smallmatrix}\{\text{C-CLASS}\}\\ ?\end{smallmatrix} \right\rangle \text{—}$

Our definition of VI in (2) already conforms to Monotonicity rather than Rewriting. The weakening of Cyclicity pushes the responsibility for the correct order of VI steps inside a cycle to the ordering of VI rules; the correct choice of μ allomorphs can be ensured by placing all VI rules realizing Aspect before those realizing μ. Finally, the locality restriction that prompted the adoption of weakened Cyclicity can be stated in terms of linear adjacency. One way to achieve this on strings is to incorporate phase boundaries into the linearized input to morphology. This would prevent the right context in (13a) from crossing the phase boundary.

To summarize, the string-based formalism can be used both to capture patterns that require VI to (apparently) make reference to word-internal tree structure, and to handle counter-examples to Cyclicity in a simple and natural way.

5 Inching toward Word & Paradigm

What we find when we simplify DM as we have, is that we reveal how DM and certain models of W&P end up looking a lot alike. Per Karttunen (2003), Stump's (2003) Paradigm Function Morphology (PFM) can also be recast as regular relations. That both PFM and DM can be recast as regular relations means that whatever dress these frameworks may be adorned with, they are in some fundamental and real ways the same. For a direct comparison, consider the PFM-style realization (15) of a Swahili example (14):

(14) ni-ta-taka
1SG-FUT-want

(15) PF($\langle$TAKA, σ:$\{$1sg fut$\}\rangle$) =
[II: [I: ***Stem***($\langle$TAKA, $\sigma\rangle$)]]

a. [I: $\langle taka, \sigma$: $\{$1sg fut$\}\rangle$] = $\langle tataka, \sigma\rangle$

b. [II: $\langle tataka$, σ: $\{$1sg fut$\}\rangle$] = $\langle nitataka, \sigma\rangle$

This style of analysis is clearly just a series of rewrite rules, where the ordering of the blocks is handled as composition of VI-rules in our framework. Compare with our formulation.

(16) $\left\langle \begin{matrix}\{\mathrm{D}, 1, \textsc{sg}\}\\ None\end{matrix} \right\rangle \left\langle \begin{matrix}\{\mathrm{T}, \textsc{fut}\}\\ None\end{matrix} \right\rangle \left\langle \begin{matrix}\{\mathrm{V}, \textsc{taka}\}\\ None\end{matrix} \right\rangle$

(17) $\left\langle \begin{matrix}\{\textsc{taka}\}\\ None\end{matrix} \right\rangle \rightarrow \left\langle \begin{matrix}\{\textsc{taka}\}\\ \mathrm{t}\end{matrix} \right\rangle \left\langle \begin{matrix}\{\textsc{taka}\}\\ \mathrm{a}\end{matrix} \right\rangle \left\langle \begin{matrix}\{\textsc{taka}\}\\ \mathrm{k}\end{matrix} \right\rangle \left\langle \begin{matrix}\{\textsc{taka}\}\\ \mathrm{a}\end{matrix} \right\rangle$

$\left\langle \begin{matrix}\{\textsc{fut}\}\\ None\end{matrix} \right\rangle \rightarrow \left\langle \begin{matrix}\{\textsc{fut}\}\\ \mathrm{t}\end{matrix} \right\rangle \left\langle \begin{matrix}\{\textsc{fut}\}\\ \mathrm{a}\end{matrix} \right\rangle$

$\left\langle \begin{matrix}\{1\textsc{sg}\}\\ None\end{matrix} \right\rangle \rightarrow \left\langle \begin{matrix}\{1\textsc{sg}\}\\ \mathrm{n}\end{matrix} \right\rangle \left\langle \begin{matrix}\{1\textsc{sg}\}\\ \mathrm{i}\end{matrix} \right\rangle$

6 Discussion

Above, we defined a portion of DM over strings and gave examples as a proof of concept. Given recent work in DM which seeks to limit the size of context available to morphological rules (Merchant, 2015), we believe that limiting DM to strings is a step in the right direction. Rather than limit the size of the relevant windows over trees, restricting ourselves to regular relations over strings gives us a solid formal grounding to limit contexts. Restrictions on context can be naturally stated by regular expressions, and the formal properties of these systems are well understood. As a final plus, in eliminating tree structure from DM and relying on rule-ordering, we bring DM closer to certain instantiations of W&P morphology, particularly Stump's PFM. We believe this bringing together of frameworks is enlightening, as it shows the cores of each are more similar than they at first appear to be.

Assuming that running DM over strings does not cost us empirical coverage, we contend that DM should be run on strings, providing a direct explanation for the generalization that morphology appears to be regular, and allows for efficient parsing and generation of surface forms (Kaplan and Kay, 1994).

References

Jonathan David Bobaljik. 2000. The ins and outs of contextual allomorphy. In *University of Maryland Working Papers in Linguistics*, pages 35–71.

Michael Brody. 1997. Mirror theory.

Amy Rose Deal and Matthew Wolf. 2017. Outwards-sensitive phonologically-conditioned allomorphy in Nez Perce. In *The morphosyntax-phonology connection: locality and directionality at the interface*. Oxford University Press.

Vera Gribanova and Boris Harizanov. 2017. Locality and directionality in inward-sensitive allomorphy: Russian and Bulgarian. In *The morphosyntax-phonology connection: locality and directionality at the interface*. Oxford University Press.

Morris Halle and Alec Marantz, 1993. *Distributed Morphology and the pieces of inflection*, pages 111–176. The MIT Press.

Ronald M. Kaplan and Martin Kay. 1994. Regular models of phonological rule systems. *Computational Linguistics*, 20(3):331–378.

Lauri Karttunen, 2003. *Computing with Realizational Morphology*, pages 203–214. Springer Berlin Heidelberg, Berlin, Heidelberg.

Gregory M. Kobele. 2002. Formalizing mirror theory. *Grammars*, 5(3):177–221.

Jason Merchant. 2015. How much context is enough? Two cases of span-conditioned stem allomorphy. *Linguistic Inquiry*, 46(2):273–303.

Edward P. Stabler. 1997. Derivational Minimalism. In Christian Retoré, editor, *LACL '96 Selected Papers*, pages 68–95. Springer Berlin Heidelberg.

Peter Svenonius. 2012. Look both ways: Outward-looking allomorphy in icelandic participles.

Jochen Trommer. 1999. Morphology consuming syntax' resources: generation and parsing in a minimalist version of Distributed Morphology. In *Proceedings of the ESSLLI Workshop on Resource Logics and Minimalist Grammars*.