Affix ordering in Optimal Construction Morphology

Sharon Inkelas
University of California, Berkeley

This paper sketches an integrated approach to affix ordering within Optimal Construction Morphology, a bottom-up, competition based model of word production in which each step of affixation is the optimal choice among competing possibilities (Caballero & Inkelas 2013). Optimality-theoretic models are natural fits for affix ordering, a complex phenomenon governed by a mix of conflicting universal and language-specific factors which interact differently in every language. This study covers familiar, global cross-linguistic principles such as semantic relevance (e.g., Bybee 1985) and scope (e.g., Baker 1988; Rice 2000), integrating them with local lexical selectional restrictions (e.g., Fabb 1988); it also incorporates usage-based factors such as Complexity-Based Ordering (e.g., Hay & Plag 2004). The study innovates in adding affix informativity to the mix.

1. Introduction

This paper approaches the topic of affix ordering from the theoretical perspective of Optimal Construction Morphology (OCM), a bottom-up, competition-based model of word production in which each step of affixation is the optimal choice among competing possibilities (Caballero & Inkelas 2013).[1]

Descriptions of affix ordering in individual languages have appealed variously to language-specific affix ordering templates, usage-based statistics about affix co-occurrence, and broad, cross-linguistic, functionally motivated principles of affix ordering.

All of these components are likely part of the ultimate, complex story. Dealing with their interaction is a particularly suitable task for a lexicalist, competition-based approach with defeasible constraints, such as OCM, since virtually every in-depth probe of affix ordering generalizations within a language with complex morphology results in competing generalizations with lexical exceptions that require expert adjudication.

1. Sincere thanks to two anonymous reviewers of this volume, and to the editors, for helpful feedback on an earlier version. This work is a direct offshoot of research conducted with Gabriela Caballero. None of these people are responsible for any errors herein.

doi 10.1075/la.229.16ink

Optimal Construction Morphology (Caballero & Inkelas 2013) is a highly lexicalist, constraint-based approach to morphology in which affix ordering emerges from the lexicon but may also be subject to defeasible grammatical constraints. Like other lexicalist theories such as Lexical Morphology and Phonology (Kiparsky 1982a) or Construction Morphology (Booij 2010), OCM has not previously addressed the topic of affix ordering head-on; this paper provides that opportunity.

This paper will attempt an integration, within the formal framework of OCM, of arbitrary language-specific constraints, overarching cross-linguistic grammatical principles, and usage-based tendencies that are all too often discussed separately, such that their conflicts with one another are overlooked rather than probed. The paper also proposes to introduce an information-theoretic component into the affix ordering discussion, based on a pilot study showing an inverse correlation between informativity and distance from the root in Turkish suffixes.

2. Production vs. licensing as the goal of a morphological model

One reason for choosing OCM for an integrated discussion of factors contributing to relative affix order is that OCM is explicitly a model of word production. It therefore forces the analyst to ask not 'what words are possible in language X, and what meaning does each such word convey', but the even more challenging question of 'what is the best word in language X to express meaning Y (even if imperfectly)?'.

The first of these questions is the one typically addressed in familiar item-and-arrangement theories, e.g. Lexical Morphology and Phonology (Kiparsky 1982b). But as Caballero & Inkelas observe in their paper introducing OCM, morphological grammaticality is often relative, not absolute. For example, morphological blocking, in which portmanteau forms like *mice* block compositional forms like **mouse-s*, is about relative goodness. According to Caballero & Inkelas (2013), there is nothing wrong with **mouse-s* except that *mice* is better, in a way to be made clear in Section 5. In the competition between *mice* and **mouse-s*, *mice* wins.

Because word-production is ultimately an exercise in optimization rather than perfection, it is a useful perspective from which to discuss constraints on affix order which can rarely all be completely satisfied.

3. Overview: Determinants of affix order

Rice (2011) succinctly lays out three very general types of factors influencing affix ordering. (Other excellent recent overviews can be found in Manova & Aronoff 2010 and Saarinen and Hay 2014).

(1) – Grammatical principles
  a. Syntactic and semantic (relative affix order is determined by function, either paradigmatic or syntagmatic)
  b. Phonological considerations (P » M)
– Arbitrary, stipulated via language-specific position class templates
– Extra-grammatical factors such as frequency, productivity, parsability

Of these, grammatical principles of affix ordering have been the focus of theoretical study (and are the focus of Rice's (2011) overview). The most famous two such principles are the syntactically based Mirror Principle of Baker (1988), built into the architecture of Distributed Morphology (e.g., Harley & Noyer 1999), and the semantically based Relevance Principle of Bybee (1985). The Mirror Principle is most directly applicable to valence-changing derivational morphology; it states that morphological derivation (and hence the linear order of morphemes) reflects the order of syntactic operations. The Relevance principle holds that affixes whose meanings are semantically more relevant to the root occur closer to the root than affixes whose meanings are less relevant. Derivation, which converts one lexeme to another, is predicted on both accounts to occur inside of (closer to the root than) inflection. Rice (2000) draws upon both of these proposals in developing Scope-based ordering, according to which the order of affix attachment correlates with the relative scope affixes have over one another. Where scopal relations are unclear or variable, so is affix ordering; where semantic scope is fixed, affix order is expected to be rigid language-internally and cross-linguistically.

Phonology can also provide an overarching principle for ordering affixes. McCarthy and Prince (1995) propose that phonological well-formedness constraints may determine the relative order of two affixes that occur in the same word. The degree to which this occurs is rather limited (Paster 2005), outside of a few cases of mobile affixes (Fulmer 1991; Kim 2008), morphological metathesis (Hargus 1993) and the analyses of infixation in McCarthy & Prince 1993 and related works (though see Yu (2007) for alternatives to this general approach to infixation).

On the lexical side, as is generally acknowledged if not highlighted in theoretical treatments, there are many kinds of language-specific and highly lexical considerations on the relative order of co-occurring affixes, as well as co-occurrence restrictions that arbitrarily ban two affixes from occurring even when their syntactic and semantic functions appear compatible. Fabb's (1988) treatment of English affix ordering is a well-known example of this kind of approach. On Fabb's analysis, each derivational English suffix is lexically equipped with a list of the morphemes which it may immediately follow. These lists are affix-specific and lexical; they do not follow from form or function. Position class templates (see, e.g., Simpson & Withgott 1986) can be classified in this same category, since they are also language specific and do not follow from form or function.

Although not classified as 'lexicalist' because of their realizational approach to morphology, Anderson's A-Morphous Morphology (1992) and Stump's Paradigm Function Morphology (2001) share with lexicalist and templatic approaches the ability to stipulate, for each affix(ational rule) in the language, its order of application relative to other affixation rules in the language; this is comparable to Fabb's listing approach, and provides complete, language-specific, templatic coverage of affix ordering.

Theoretical work on grammatical principles of affix ordering tends to transcend the details of arbitrary, language-specific templates. And theoretical work on arbitrary, language-specific templates tends not to focus on broader, obviously defeasible semantic and syntactic principles of affix ordering.[2] One notable exception is Rice (2000), who argues for a rational division of labor between these two sources of affix ordering: grammatical principles for affixes whose relative scopal relations are clear, and language-specific templatic principles for affixes which do not clearly have scope over one another. Similarly, Hyman (2003), Aronoff and Xu (2010) and Spencer (2013), among others, discuss the tension between scope-based and templatic affix ordering within a single theoretical model.

A very recent body of work has started to show that usage statistics, which reflect or contribute to psycholinguistic properties, also correlate with affix order and may need to join cross-linguistic grammatical principles and language-specific details in the pantheon of factors contributing to affix order. Hay (2002), Plag (2002), Hay and Plag (2004), Hay and Baayen (2005), Plag and Baayen (2009) have all shown in various ways that less 'parsable' affixes tend to occur closer to the root that more 'parsable' affixes, where parsability is a function of productivity, phonological transparency, and other psycholinguistically relevant factors. Parsability is not directly a function of meaning, and form plays only a minor role. However, Ryan (2010) has shown that by observing the most commonly occurring order within local pairs of affixes (bigrams) in Tagalog, it is possible to derive statistical predictions about the likelihood of different ordering possibilities for words with more than two affixes. Parsability and bigram frequency have not yet been incorporated into a broader formal theoretical model of word formation, despite the evidence for their influence.

In sum, there is a a huge amount of knowledge about the many quite different factors that cause, or at least correlate with, affix ordering, and these factors tend to be theoretically modeled in isolation from one another. No single existing theory of morphology incorporates all of them into a model of affix ordering. This paper

2. This is not to say that such approaches are not principled; for example, Lexical Morphology and Phonology (Kiparsky 1982a) and A-Morphous Morphology (Anderson 1992) both make heavy use of the Paninian principle of disjunctive 'elsewhere' ordering. The point is simply that both frameworks are focusing at a level of language-specificity under the radar of broader, cross-linguistic syntactic and semantic affix ordering generalizations.

attempts, in necessarily brief and cursory fashion, to incorporate a relatively full range of factors into the theoretical model of OCM. The mismatch between the vast scope of the problem and the short length of this paper makes this a programmatic exercise rather than a definitive theory. It is hoped that future research can further this ambitious venture.

4. Optimal Construction Morphology

OCM is a suitable vehicle for attempting an integration of the various factors affecting affix order, for two reasons:

1. OCM is competition-based, making it a natural model for handling competing contradictory pressures
2. OCM combines highly articulated lexical entries with general grammatical constraints, making it a natural model for handling lexical interference with higher-order principles

5. The structure of OCM

OCM was developed, by Caballero and Inkelas (2013), as a theory that could generate blocking effects and multiple exponence by building words from the bottom (root) up, choosing among competing additional layers of morphology based on considerations of target faithfulness and well-formedness.

The foundation of OCM is the target meaning, M (e.g., (2)), and the Constructicon (e.g., (3)), a set of roots and bilevel constructions which include affixation, compounding, reduplication, and all manner of morphological constructions which take as input a morphological constituent (or constituents) and produce as output a constituent with properties conferred by the construction. OCM is a target-driven, competition-based version of Construction Morphology (Booij 2010).

(2) Target meanings (M) of 'mice', 'dogs', 'carried', 'carrying'
   a. M=[Mouse, cat=noun, +animate, +plural]
   b. M=[Mouse, cat=noun, +animate, +singular]
   c. M=[carry, cat=verb, +past]
   d. M=[carry, cat=verb, +prog]

(3) Constructicon fragment (4 roots, 3 suffixes)
   a. $[\text{Mouse}]_N$
      $[\text{Mice}]_{N,Pl}$
      $[\text{Dog}]_N$
      $[\text{Carry}]_V$

b. $[[\ \ ]_N\ s\ ]_{Pl}$
$[[\ \ ]_V\ d\ ]_{Past}$
$[[\ \ ]_V\ ing]_{Prog}$

In OCM, competition between competing forms is modeled via ranked or weighted constraints, as in classic Optimality Theory (Prince & Smolensky 1993; Prince & Smolensky 2004; McCarthy 2002) or Harmonic Grammar (Legendre, Miyata & Smolensky 1990; Hayes & Wilson 2008). OCM is a bottom-up theory of word production with a serial character. In OCM, each tableau operates on two inputs: the target, which never changes, and the stem under construction, which is augmented in form and/or meaning in every cycle of evaluation. The word in progress is the pairing between meaning and form constituted by the stem under construction and the target. On each cycle of evaluation, the candidate outputs are those new stems which are produced by combining the input stem with every possible individual entry in the Construction whose selectional restrictions are compatible with that specific input. In the respect that each tableau evaluates the results of applying a single morphological step to the input, OCM resembles Harmonic Serialism (e.g. McCarthy 2010), applied to morphological derivations by Wolf (2008).

For instance, at a stage in the derivation where the input stem is *pluck*, the candidate outputs in OCM would include *plucky, plucking, plucker,* and any other forms that can be produced by combining the input *pluck* with a single morphological construction. The output of one tableau can serve as input to another, potentially producing words with multiple layers of morphology.

Within any given tableau, selection among the candidates operates on the optimization along two dimensions: (a) faithfulness to the target meaning and (b) well-formedness.

The target meaning (M), present throughout the derivation, is analogous to the familiar 'property sets' or featural inputs which drive many realizational approaches to inflectional morphology (Anderson 1992; Stump 2001), including Noyer's (1993) and Xu and Aronoff's (2011) OT approaches to inflection. (It is a more semantic representation than the syntactic configuration which serves as the input to realizational morphology in Distributed Morphology, in that OCM aspires to derive the internal branching structure of words from constraint interaction, rather than provide the syntactic structure of a word in the input.) M includes standard features (e.g., part of speech, person, polarity, number, tense, aspect, mood, diminutive, evidential), valence structures, (including number of arguments and the thematic roles to which they link), as well as the meanings commonly (but not exclusively) associated with 'root' morphemes.

Faithfulness to the target meaning is assessed through the FAITH constraints in (4). Each constraint compares the target meaning, M_T, to the meaning associated with each candidate under consideration (M_C):

(4) Faithfulness to the target meaning (M-FAITH)
  a. M-MAX: Violated once by each property $m_i \in M_T$ that is not present in M_C
  b. M-DEP: Violated once by each property $m_i \in M_C$ that is not present in M_T

Caballero and Inkelas (2013) show that from this basic lexicalist architecture, OCM can derive classic blocking effects with ease. Consider, for example, the target meaning [MOUSE]$_{N,PL}$. There are two logically possible ways to create a word matching this target meaning from the Constructicon fragment in (3): *mice* and *mouse-s*. 'Blocking' is the term applied by Aronoff to the fact that **mouse-s* is ungrammatical, even though its formation follows the general rule for plural suffixation in English. It is ungrammatical not because it is ill-formed in any absolute sense, but because *mice* is better.

In OCM, blocking follows from maximal satisfaction of Faithfulness constraints. As can be seen in the tableau below, whose target is [MOUSE]$_{N,PL}$ and for which there is not yet an input stem, the viable candidates are those morphologically well-formed stems in the lexicon whose meaning comes closest to the target. Here, as in subsequent tableaux, the meaning target M_T is represented simply as M:

(5)

Cycle 1	Input stem: N/A		M: MOUSE CAT=N NUM=PL	M-FAITH
☞ a.	[MICE] NUM=PL	*mice*	MOUSE CAT=N NUM=PL	
b.	[MOUSE]	*mouse*	MOUSE CAT=N	*! (NUM=PL)

On the first selection cycle, the singular (or unmarked) stem *mouse* loses to *mice*, since *mice* does a better job of matching the target. Because the target is perfectly matched, no further morphology is required. Hypothetical candidate *mouse-s* is not judged ungrammatical in this or any other tableau; rather, it never even has a shot at being evaluated, because the path that would lead to it, namely the selection of *mouse*, is prematurely ended.

In a case where one cycle of construction selection does not produce a candidate output that perfectly matches the target, a second cycle takes place. In OCM, cycling stops only when the optimal candidate output is identical to the input stem – when the 'ID candidate' is better than any of the candidates with a new layer of morphology. (The ID candidate is not defined on the first cycle, as in (5), as there is no input stem to compare the output to; but it is defined and relevant on all subsequent cycles.) The following example, again drawing on the Constructicon fragment in (3), demonstrates a case where two cycles are called for. On the first cycle, because our constructicon only has one verb, we introduce *mice* as a (silly) competitor to *carry*. Candidate *carry* wins uncontroversially on target faithfulness alone:

(6)

Cycle 1	Input stem: N/A		M: CARRY CAT=V TENSE=PAST	M-FAITH
☞ a.	[CARRY] CAT=V	*carry*	CARRY CAT=V	* (TENSE=PAST)
b.	[MICE] NUM=PL	*mice*	MOUSE CAT=N PL	**!**** (all)
Cycle 2	Input stem: [CARRY] CAT=V		M: CARRY CAT=V TENSE=PAST	M-FAITH
☞ a.	+ [ ]V d] TENSE=PAST	*carried*	CARRY CAT=V TENSE=PAST	
b.	ID	*carry*	CARRY CAT=V	*! (TENSE=PAST)

In the second cycle, the identity candidate 'ID' loses to the candidate with a past tense suffix, *carried*, which carries the day as the optimal stem for target meaning CARRY, TENSE=PAST.

6. Affix ordering in the original version of OCM

The original formulation of OCM, by Caballero and Inkelas (1993), broadly skirts the issue of affix ordering principles of the sort highlighted in Section 3. It does incorporate two means of ordering affixes: selectional restrictions on constructions (Section 6.1), and a constraint BE-WORD that favors Words over Stems and Stems over Roots (Section 6.2). We begin, in this section, with a discussion of these dimensions of OCM and their descriptive and explanatory limitations.

In subsequent sections we discuss how, or whether, more global, functionally motivated principles of affix ordering might be implemented in a bottom-up, competition based word production model, including scope-based ordering (Section 7.1) and Relevance (Section 7.2). Sections 8 and 9 address ways of bridging the gap between overly arbitrary templatic constraints on affix order and overly strict global universal principles by adding to the mix local ordering principles that prioritize affixes according to the language-specific, lexical criteria of parsability (Section 8) and informativity (Section 9). The latter proposal, in particular, is a novel one and is supported by a pilot study of Turkish suffix order.

6.1 Selectional restrictions

Local selectional restrictions (subcategorization frames) on affixes have been used in many analyses, e.g. Fabb (1988), to achieve local affix ordering. They are similar or equivalent to the affix bigram constraints used in Ryan (2010) to model affix ordering pattern in Tagalog. The idea is that an overall set of affixes, e.g. A, B, C, D, can be ordered by imposing local orderings among pairs of suffixes. For example, from the information that A < B, A < C, B < C, C < D, a total order is computable: A < B < C < D.

Fabb (1988) argues that local selectional frames are sufficient to determine ordering among derivational affixes in English. Fabb sorts English suffixes into those which attach freely to bases of the appropriate part of speech, those which attach outside of one specific affix, those which attach only to bare roots, and those which attach to bare roots or to one of a small set of suffixes (not depicted here). (See also Hay & Plag 2004 for amplification of Fabb's hypothesis.)

(7)

Condition	Examples
Freely attaching suffixes	– *-able*, e.g. *manage-able, magnet-iz-able, indemn-ifi-able* – deverbal *-er*, e.g. *manag-er, magnetiz-er, indemn-ifi-er* – *-ness*, e.g. *happi-ness, sorrow-ful-ness, penni-less-ness*
Suffixes which attach outside one specific suffix only	– Noun-forming *-ary* may only follow *-ion*, e.g. *revolutionary* – *-ic* may only follow *-ist*, e.g. *modernistic* – deadjectival *-y* may follow only *ent,* e.g. *residency*
Suffixes which never attach to an already-suffixed word	– *-ful,* e.g. *sorrowful* – *-hood*, e.g. *motherhood* – *-ify*, e.g. *solemnify* – *-ment,* e.g. *entrapment* – adjective-forming *-y,* e.g. *shadowy*

Ryan (2010) takes a related approach to Tagalog, proposing bigram constraints of the form XY which are violated by words containing affix X in which X does not immediately precede Y.

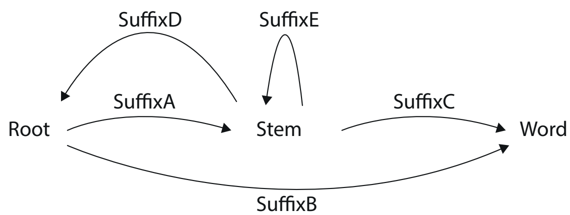
Whether affix-specific selectional constraints need to mention the identity of the preceding morpheme per se or whether they mention the type of stem they can combine with is another possibility within this same general space. Selkirk (1982), Lieber (1980), Kiparsky (1982b), and Inkelas (1989), among others, posit that affixes in English select for different types of base, such as Root, Stem, or Word. Consider the simple example below. The existence of affixes whose lexical frames take a Root as input and produce a Stem as output (but not vice versa), and which take a Stem as input and produce a Word as output (but not vice versa), predicts a fixed ordering among the two types of suffix:

(8) a. Root → Stem affixation: $[[\quad]_{\text{Root}}$ SuffixA $]_{\text{Stem}}$
b. Stem → Word affixation: $[[\quad]_{\text{Stem}}$ SuffixB $]_{\text{Word}}$
Predicted ordering: Root-SuffixA-SuffixB

A more complicated schema is depicted in (9). Here, one suffix (E) is self-recursive, producing stems of the same type that it attaches to. Another (B) jumps over an intermediate stem type, going directly from Root to Word. Suffix D reverses direction on the usual stem type hierarchy, taking Word back to Root.

(9) a. Root → Stem affixation: $[[\quad]_{\text{Root}}$ SuffixA $]_{\text{Stem}}$
Root → Word affixation: $[[\quad]_{\text{Root}}$ SuffixB $]_{\text{Word}}$
b. Word → Stem affixation: $[[\quad]_{\text{Stem}}$ SuffixC $]_{\text{Word}}$
Stem → Root affixation: $[[\quad]_{\text{Stem}}$ SuffixD $]_{\text{Root}}$
c. Stem-preserving affixation: $[[\quad]_{\text{Stem}}$ SuffixE $]_{\text{Stem}}$

These lexical entries determine a partial order among the affixes; from them, a flowchart like the following emerges as a depiction of what the lexicon allows:[3]

(10)


If no other principles – scope-based, relevance-based, phonological – interfere, this system predicts the following possible suffix orders, irrespective of meaning and function:

(11) B
A((DA)*E*)*C *i.e.* AC, ADAC, ADAEC, AEC, *etc.*

More difficult to incorporate graphically into the flow chart, but possible in languages, are affixes that can attach to stems of any type and are either type-preserving or produce a stem of fixed type.

(12) $[[\quad]_i$ Suffix $]_i$ Type-neutral suffix
$[[\quad]_i$ Suffix $]_{\text{Word}}$ Suffix that attaches to stems of any type and produces Word

3. See e.g. Hankamer 1986 on a finite-state chart of this kind for Turkish, and Beesley & Karttunen 2003 for an overview of finite-state morphological models in general.

Inkelas (1989), expanding on Selkirk (1982), models Lexical Morphology and Phonology-style level ordering via stem-type selectional affix frames, obviating the need to divide the grammar into modules. Level 1 affixes attach to and produce stems of type 1; stem 2 affixes attach to and produce stems of type 2; etc. Taking the model further, Inkelas (1993) uses lexical selectional frames to derive the entire position class templatic system of Nimboran, in which no two affixes may occur in the same position. In Nimboran, every affix produces a stem of a type higher than the type it attaches to.

A reason to prefer selectional frames which mention an affix and the type of stem it may follow or precede is that selectional frames are also useful in modeling phonological restrictions on the type of stem an affix may combine with, as is done by Paster (2006) in modelling suppletive allomorphy and by Yu (2007) in characterizing infix placement. Stem selection is also an essential part of the analysis of systems built around morphomic stem distinctions of the sort discussed by Aronoff (1994), Blevins (2003), and Maiden (2005).

Wrapping up this discussion, templatic affix order can be captured in OCM by means of local templatic constraints which list information, for each affix, about the type of base it attaches to and the type of stem it produces. To the extent that global affix order is predictable from local affix order, it can be implemented in OCM or any other bottom-up theory of word formation with the equivalent of selectional frames.

6.2 Wordhood

As seen from examples (10) and (11), lexical frames can specify where, in the hierarchy of stem types, an affix attaches and what type of stem the affix produces. Nothing in this system forces stem type to progress monotonically from Root to Stem to Word in the formation of a complex word. But ultimately, Wordhood is required of a well-formed word in order to enter into syntactic constructions. To this end, OCM posits a constraint Be-Word which prefers candidates which are closer to Word on the Wordhood scale:

(13) Be-Word: violated once for each degree of separation from Word on the Wordhood scale
Root………………………..Stem(s) ………………………..Word

The Wordhood scale is language-specific in the number of points between Root and Word, and may only impose a partial order on stem types. See Inkelas and Orgun 1998 on level 'nonordering' in Turkish, and Mohanan 1986, 1995 on the level ordering loop in Malayalam, phenomena which correspond to a flexible hierarchy of stem types.

Caballero and Inkelas discuss the Wordhood scale as a motivator of some kinds of semantically empty morphology, including the Latin theme vowels discussed by Aronoff (1994):

> The [Latin] theme vowel is … a marker of the category *verb* only in the sense that it is determined by the category *verb*, just as final devoicing is a marker of the word boundary. In itself, it has no significance. It is empty. Nonetheless, it is not useless. It has a use in the language…it serves to determine the conjugation of the verb stem, or which inflectional affixes will realize the various morphosyntactic properties that the verb bears in a particular instance.
> [Aronoff 1994: 46]

The table in (14), reproduced from Caballero & Inkelas 2013, illustrates four different derivational verb suffixes and the theme vowels that must be added to each to prepare the stem for future inflection (Aronoff 1994: 46):

(14) Latin Derivational Verb Suffixes

Meaning	Derivational Suffix	Theme Vowel	Example	Gloss
desiderative	-ur-	-ī	ēsurīre	'be hungry'
iterative	-it-	-ā	vīsitāre	'see often'
inceptive	-sc-	-e	calescere	'get warm'
intensive	-ess-	-e	capessere	'seize'

Caballero and Inkelas (2013) model Aronoff's insight in OCM by listing, for each suffix, the stem type(s) it combines with and the type of stem it produces, as defined on the following stem type scale:

(15) Root….D(erivational)Stem….I(nflectional)Stem….Word

Derivational suffixes combine with Roots and produce constituents of type DStem, while theme vowels combine with Roots or DStems and produce IStems. Inflectional suffixes attach to IStems. Examples of lexical entries are given below:

(16) Derivational suffix (e.g. iterative): $[[\ \ ]_{\text{Root}}\ \text{it}\ ]_{\text{Dstem}}$
Theme vowel (e.g. -ā-) $[[\ \ ]_{\text{Root/Dstem}}\ \bar{\text{a}}\ ]_{\text{Istem}}$
Inflectional suffix (e.g. 1st conjugation 1st person) $[[\ \ ]_{\text{Istem}}\ \bar{\text{o}}\ ]_{\text{Word}}$

On Caballero and Inkelas's account, the Latin theme vowel is the only mechanism to promote a Dstem to Istem status, which is necessary in order for it to take inflection. Globally, it is clear that theme vowels are helpful. Locally, on the cycle at which theme vowels are added, the structural advantage of adding a theme vowel is simply advancement on the Wordhood scale:

(17) Addition of theme vowel advances Dstem towards wordhood:

	Input = [vīs-it]$_{\text{DSTEM}}$ SEE, ITERATIVE		M: SEE ITERATIVE	M-FAITH	BE-WORD
☞ a.	[[ ]$_{\text{ROOT/DSTEM}}$ ā]$_{\text{ISTEM}}$	[vīs-it-ā]$_{\text{ISTEM}}$	SEE ITERATIVE		*
b.	ID	[vīs-it]$_{\text{DSTEM}}$	SEE ITERATIVE		**!

Of course, as a reviewer aptly points out, theme vowels do not always advance stems on an ordered scale of stem types. Theme vowels in Latin can also create a stem which is the input to derivation, which could represent either non-movement or even regression on the Wordhood scale. Furthermore, not all stem types are necessarily ordered with respect to one another; this is especially true of morphomic stems. The total fixed ordering sketched for stem types in (15) is only a property of some systems. In a system in which stem types are not ordered with respect to one another on a Wordhood scale, the BE-WORD constraint would simply be irrelevant.

6.3 Interim summary

Between affixal selectional frames and the wordhood scale, OCM – or any theory employing such devices – can effectively model a variety of affix ordering patterns.

There are a variety of reasons why this approach to affix order is incomplete, however. First, it fails to capture any sort of functional motivation for why affixes are ordered in the way that they are. Second, local selectional restrictions limit the impossible but are not always deterministic, a flaw in a step-wise bottom-up theory of word formation. For example, many languages exhibit pairs of affixes that can occur in either order – with consequences for meaning. The association between meaning and order cannot be captured by means of selectional restrictions alone. Both of these objections refer to the connection between meaning and affix order, to which we now turn.

7. Scope and relevance

A great body of work in morphology has shown a cross-linguistic tendency for relative affix order to be related to syntactic and/or semantic scope: if the function associated with affix A has scope over the function associated with affix B, then affix B will be closer to the root, or 'inside', affix A (see e.g. Bybee 1985; Baker 1988; Alsina 1999; Rice 2000, among many others). Can overarching generalizations like scope- and relevance-based ordering be captured in a bottom-up theory in which the choice to add any given affix is based on a comparison between the meanings of competitor candidates to the meaning of the target word? In this section we offer a qualified yes, concluding that while some scopal ordering restrictions emerge automatically from the basic architecture of OCM, global scopal principles (such as Bybee's Relevance) are challenging to implement.

7.1 Scope

In lexicalist approaches to morphology, scopal ordering restrictions are often modeled by means of explicit grammatical constraints that filter out affix orders that do not conform to the scope-based ordering generalization. This is the approach taken, for example, by Hyman (2003) and Aronoff and Xu (2010). (The statement of Hyman's Mirror constraint, in (18), has been modified for greater generality; Hyman formulates separate Mirror constraints for each affix pair in the language). The statement of Scope, in (19), is adapted by Aronoff and Xu (2010) from Spencer 2003.

(18) Mirror (X, Y) The morphosyntactic input [[[….] X] Y] is realized *Verb-x-y*, where *x* is the exponent of X and *y* is the exponent of Y [Hyman 2003: 251]

(19) Scope: Given two scope-bearing features f_1 and f_2, if f_1 scopes over f_2, then I_2, an exponent of f_2, cannot be farther away from the same stem than I_1, an exponent of f_1. [Aronoff & Xu 2010: 389]

Both of these constraints are assessed globally on forms already containing both *x* and *y* (Mirror) or I_1 and I_2 (Scope).

Two questions arise with such approaches. First, a procedural question: how can a constraint like Mirror or Scope be utilized in a model of grammar that builds words one step at a time to ensure that the would-be inner affix is selected first? Second, is it necessary to stipulate Scope or Mirror in a model that already has a target-matching constraint? The answer to both of these questions is that OCM can derive the effects of Mirror and Scope through M-Faith in its bottom-up word-forming process.

To see how this works, consider one of the Chichewa verbal examples discussed by Hyman (2003): reciprocalized causatives (Root-caus-rec-) vs. causativized reciprocals (Root-rec-caus-). In Chichewa, both orders are possible, with corresponding scope-based differences in meaning.

In Hyman's constraint-based approach to affix ordering, the meaning of a word is given, and the challenge to the grammar is to order the exponents of meaning appropriately. In OCM, however, the challenge is different: the challenge is building a word that means the same thing as the target. Though this difference may seem subtle, it is important.

One of Hyman's examples is the verb meaning *cause to tie each other*, represented as [[[tie] rec] caus]. In OCM, the first cycle of construction selection will choose the root *mang-*, meaning 'tie'. On the second round of construction selection, the competing candidates will include the causative and the reciprocal, as shown in (20). Each candidate matches the target meaning only imperfectly, such that a second round of affixation will need to be undertaken.

(20)

Cycle 2	Input stem: *mang* TIE		M: [[[TIE]RECIP]CAUS]	M-FAITH
☞ a.	[[ ] an] RECIP	*mang-an*	[[TIE]RECIP]	* (CAUS)
☞ b.	[[ ] its] CAUS	*mang-its*	[[TIE]CAUS]	* (RECIP)
c.	ID	*mang*	[TIE]	** (CAUS, RECIP)

Let us assume, for sake of argument, that on this first round of suffix selection, candidates (20a) (causative) and (20b) (reciprocal) tie. Both perform better on M-FAITH than the ID candidate, and both are missing one major component of M, so this is a plausible assumption.[4]

The second round of affixation improves further on M-FAITH. Since the first affixation cycle produced two outputs, the cycle depicted in (21) must consider two inputs, in what Itô, Mester and Padgett (1995) term a 'tableau des tableaux'.

(21)

Cycle 3	(two inputs)		M: [[[TIE]RECIP]CAUS]	M-FAITH
	Input stem: *mangan* [[TIE]RECIP]			
☞ a.	[[ ] its] CAUS	*mang-an-its*	[[[TIE]RECIP]CAUS]	
b.	ID	*mang-an*	[[TIE]RECIP]	* (CAUS)
	Input stem: *mangits* [[TIE]CAUS]			
c.	[[ ] an] RECIP	*mang-its-an*	[[[TIE]CAUS]RECIP]	* (*scope*)
d.	ID	*mang-its*	[[TIE]CAUS]	* (RECIP)

4. Quantifying M-FAITH violations is a bit of a black box in OCM, as it requires a more fine-grained representation of the contents of meaning than OCM is currently able to provide. The reader is encouraged to suspend disbelief in this area in the interests of exploring the potential of the framework.

On this cycle, both orders of affixation are generated and compete; *mang-an-its* (21a) and *mang-its-an* (21c) possess the meanings that result from the observed order of affixation. The candidate that wins (21a) is the one whose meaning matches the target. Note particularly that candidate (21c) loses because its scopal relations are incorrect. No special Scope or Mirror constraint has to be stipulated to doom this candidate. Its failure emerges simply from target meaning faithfulness, in the same way that blocking emerges in OCM without the need for a specific stipulation to that effect (Section 7 and Caballero and Inkelas 2013).

As Hyman shows, languages do not always obey Mirror (see also Spencer 2013). In the terms of OCM, this means that scopal relations in the winning candidate do not always perfectly match those in the meaning target. Sometimes lexical combinatorics make a perfect match impossible. In Chicheŵa, as Hyman argues, the causative and applicative suffixes are always ordered in the same way, regardless of semantic scope; words containing both are therefore technically ambiguous, though pragmatics often privileges one reading over another. In (22a), the order of attachment of the causative and applicative affixes gives a meaning that matches the meaning target scopally; in (22b), the order of attachment of the causative and applicative affixes should give the meaning of 'use a spoon to cause the woman to stir', which mismatches the intended, target meaning of 'cause the woman to use a spoon to stir'.

(22) a. Applicativized causative (M target = [[[cry]**caus**]**appl**]
alenjé a-ku-líl-**its**-**il**-a mwaná ndodo
hunters 3pl-prog-cry-**caus**-**appl**-fv child sticks
'the hunters are making the child cry with sticks'

b. Causativized applicative (M target = [[[cry]**appl**]**caus**]
alenjé a-ku-tákás-**its**-**il**-a mkází mthíko
hunters 3pl-prog-stir-**caus**-**appl**-fv woman spoon
'the hunters are making the woman stir with a spoon'

Hyman generates what he calls the 'anti-scope' ordering in (22b) by ranking a bigram constraint, Template(caus,appl) above Mirror. Template forces the observed linear order.

In OCM, the Template constraint per se is not needed, any more than Mirror is; the effects of effects emerge from lexical selectional frames of the sort discussed in Section 6.1. For example, if we assume the lexical frames in (23), we correctly predict the possibility of Root-caus-appl, Root-caus-rec, Root-rec-caus, and the impossibility of *Root-appl-caus.[5]

5. Hyman (2003) discusses a very interesting complication with Applicative and Reciprocal suffixation. The order Applicative-Reciprocal is a legitimate realization of the meaning [[[Root]appl]recip]. But when the intended meaning is [[[Root]recip]appl], then either the

(23) a. [[$]_{ROOT}$ Caus $]_{ROOT}$
b. [[$]_{ROOT}$ Appl $]_{STEM}$
c. [[$]_{i}$ Rec $]_{i}$ (type-preserving, either Root→Root or Stem→Stem)

No constraint has to rule out the sequence *Root-APPL-CAUS; it simply cannot be generated as a candidate, since the selectional requirements of Causative (verb, type Root) in (23a) are not satisfied by an input base derived by Stem-forming Applicative suffixation (23b).

It is an important part of this story that GEN is constrained by selectional frames. Instead of producing every imaginable string of affixes, GEN is limited to producing candidates in which the selectional frames of affixes are satisfied. OCM is thus a theory in which only legitimate lexical combinations are considered by the grammar. OCM chooses among these candidate combinations based on faithfulness to target meaning.

7.2 Relevance

Bybee's (1984) principle of Relevance holds that the more relevant an affix is to the meaning of a verb, the closer it will appear to the verb stem. Relevance predicts a cross-linguistic ordering among derivation (closer to the root) and inflection (farther). It also applies within inflection, the focus of Bybee's cross-linguistic study of affix order in verbs in which she reports a tendency for aspect to be ordered closer to the root than tense and mood, and for object agreement to be ordered closer to the root than subject agreement. According to Rice (2000), these relationships can also be described in terms of semantic scope. For example, tense has wider scope than aspect which pertains to internal event structure; within aspect; subjects have wider scope in the clause (are higher up in syntactic structure) than objects.

It is not obvious how to implement the principle of Relevance in a model like that of OCM, because the architecture of the model confers no apparent local benefit to attaching a 'more relevant' affix over a 'less relevant' affix at a stage where the two might be competing. From the OCM perspective, given a choice between adding one inflectional affix or another, both of which expone values in the M-target, both choices should be equally good, as they are locally harmony-increasing. And if the meanings of words containing two or more such affixes are compositional, no M-Faith-based postfilter can distinguish between candidates obeying and candidates violating Relevance-based ordering, either. In Turkish, for example, tense suffixes precede suffixes marking subject person and number agreement (e.g. *git-ti-m* 'go-PAST-1SG'); in nouns, possessor suffixes occurs inside case (e.g. *elma-sı-na* 'apple-3POSS-DAT'). But

'anti-scope' order of affixes is used, *or* the reciprocal suffix is doubled: Root-APPL-RECIP-APPL. Hyman offers an account in which TEMPLATE and MIRROR are conjoined.

from the point of view of interpretation, of compositionally constructing the meaning of the complex word from the meanings of its parts, the reverse affix orders do not seem logically impossible; if case is dative and possessor is 1st person, it should not matter which piece of information is encoded 'first', as they do not interact.[6] In an OCM framework, Relevance would have to be a stipulative post-hoc filter that favors certain orders. Its effects do not emerge from a chain of local optimization decisions based on M-Faith.

One obvious possibility within OCM would be to attempt to correlate relevance with 'amount of meaning expressed', and use *degrees* of M-Faith satisfaction to privilege the early attachment of affixes which express a lot of meaning, vs. those which express only a single inflectional feature. This might capture some of what falls under the Relevance generalization. As acknowledged earlier, meaning differences are hard to quantify. It stands to reason that the semantic differentation between EAT and CAUSE TO EAT, i.e. the syntactic and semantic chunk of meaning expressed by a causative suffix, is greater than that between EAT and EATS, i.e. than the chunk of meaning expressed by the 3rd person singular suffix. M-Faith thus already has the potential to capture *some* of the relevance bias in affix ordering. But the quantitative difference in meaning between, say, poss=sg and case=dat is not self-evident. The use of M-Faith to capture Relevance quantitatively must thus remain only a speculative possibility without a system to quantify the relative magnitude of the contribution a given property makes to the overall meaning of the target.

7.3 Why not let M-Faith govern affix order?

A natural question to ask is whether OCM could split M-Faith into individual, property-specific M-Faith-F_i constraints and use their relative ranking to order affixes.

For example, on a cycle of construction selection in which the target M contains information about number *and* case of a noun and the input stem as yet encodes neither property, the choice between the candidate adding a number marking affix and the candidate adding a case-marking affix could be made by the ranked M-Faith-num and M-Faith-case constraints (all other considerations being equal or less important). If M-Faith-num is ranked higher than M-Faith-case, then number marking

6. Indeed, counterexamples to relevance-based ordering exist. A helpful reviewer points out, for example, that possessive-case affix ordering varies in some Uralic languages, including Mari (Luutonen 1997) (e.g. *joltas-em-βlak-lan* 'friend-1sg.poss-pl-dat' ~ *joltas-βlak-em-lan* 'friend-pl-1sg.poss-dat' ~ *joltas-βlak-lan-em* 'friend-pl-dat-1sg.poss' (Luutonen 1997, cited in Marle 1998: 308; see also McFadden 2004 for discussion). In some Northeast Caucasian languages, subject markers can occur inside of tense (e.g., in Udi, *ak'-es-**ne**-d-**e*** 'see-inf-**3sg**-caus-**aorii** = (she) showed' (Harris 2000: 596)).

will be selected, and hence will occur closer to the root than case marking, which will be deferred to the next cycle of construction selection. This hypothetical scenario is illustrated below by the partial derivation of a Turkish word *gül-ler-e* 'rose-PL-DAT', starting on the cycle after root selection:

(24) Construction selection: *gül* → *gül-ler* 'rose-PL'

Cycle 2	Input stem: *gül* ROSE		M: ROSE PL DAT	M-FAITH- NUM	M-FAITH- CASE
☞ a.	[[ ] lEr] PL	*gül-ler*	ROSE PL		* (DAT)
b.	[[ ] E] DAT	*gül-e*	ROSE DAT	*! (PL)	
c.	ID	*gül*	ROSE	*! (PL)	* (DAT)

Because M-FAITH-NUM is ranked high, candidate (24a) beats out candidate (24b) on this cycle, leaving dative affixation for the subsequent cycle (25) and generating the order Root-PL-DAT:

(25) Construction selection: *gül-ler* → *gül-ler-e* 'rose-PL-DAT'

Cycle 3	Input stem: *gül-ler* ROSE, PL		M: ROSE PL DAT	M-FAITH- NUM	M-FAITH- CASE
☞ a.	[[ ] E] DAT	*gül-ler-e*	ROSE PL DAT		
b.	ID	*gül-ler*	ROSE PL		*! (DAT)

Property-specific M-faithfulness is reminiscent of the proposal by Aronoff and Xu 2010) and Xu and Aronoff 2011), who treat affixes as realizational constraints, e.g. the constraint "{past}: -ir: The past tense is realized by the suffix -ir". For Aronoff and Xu, however, all affixal exponence is evaluated in a single tableau; their approach is noncyclic, rather than cyclic. As a result, constraint ranking is not a direct means of enforcing affix order. Instead, Aronoff and Xu appeal to such means as an over-arching SCOPE constraint (Aronoff & Xu 2010: 390) as well as, when needed, feature-specific precedence constraints such as "Neg > Tense". These more global affix ordering constraints are suited to a noncyclic approach, while affix constraint ranking is not; the reverse is true in the bottom-up, cyclic, OCM approach, in which local ordering *could* be accomplished purely by ranking feature-specific M-Faithfulness constraints.

Such an approach is undesirable in several ways, however. First, it potentially duplicate the descriptive effects of lexical selection frames (e.g., Fabb 1988). Insofar as each affix corresponds to an individual M-Faith-F_i constraint, then ranking such constraints imposes a total or partial ordering on individual affixes. For example, if the negative suffix always precedes the past tense suffix, ranking the exponence constraint for negative above that of past tense would guarantee the desired order.

Ranked M-Faith-F_i constraints could not, however, replace selectional frames, which are needed independently to capture arbitrary phonological and root-based selectional restrictions on affixation. In so far as selectional frames are needed independently, Occam's razor suggest that individually ranked M-Faith-F_i constraints may not be necessary.

A second reason not to delegate affix ordering to the ranking of individual M-Faith-F_i constraints is Anderson's (1992) key point that perfect pairings between form and meaning – the canonical morpheme – are elusive. Specifically, in cases where a given feature F_x can be exponed by more than one affix (say, in cases of suppletive allomorphy) and where those affixes have different ordering properties, ranking a M-Faith-F constraint for feature [i] with respect to M-Faith-F constraints for features [j, k, etc.] will not be descriptively adequate.

A third argument against taking this approach in OCM is a perverse prediction that it potentially makes. The intuition behind faithfulness constraints is that the properties which are most important to express correspond to constraints which are ranked highest in the grammar. This logic is certainly familiar in phonology, where ranking Max-C above Max-V means that it is more important to express consonants in the output (if present in input) than it is to express vowels.

Ranking individual M-Faith-F_i constraints according to the obligatoriness of the feature or category exponed generates the prediction that morphological properties which are required to be exponed in all words (of a given part of speech) will be expressed closer to the root than categories that are not required. But this seems to be the wrong prediction. Cross-linguistically, the type of morphology most likely to be obligatory in any paradigm is inflectional, and inflection is invariably *farther* from the root than derivation (which is usually not required).[7]

Finally, affix ordering is occasionally context-dependent, with its variation conditioned by factors such as phonological well-formedness (see e.g. Hargus 1993; Aronoff & Xu 2010; cf. Paster 2005) or the presence of other affixes (see e.g. Ryan 2010 on Tagalog). No fixed ranking of M-Faith-F_i can handle this situation, so the rank-

7. This may of course by merely a superficial typological generalization; in any given word, all of its properties may be equally important. If so, then exponence importance cannot be quantified language-internally, and therefore it cannot be used to order affixes.

ing would have to be defeasible.[8] In sum, it may be the correct decision to stick with a single M-Faith constraint, rather than splitting M-Faith into rankable, feature-specific subconstraints.

8. Text corpus statistics: Parsability, frequency

The above components of affix ordering are all based on individual, lexical properties of each affix: its meaning, its function, or restrictions on the type of base it can attach to. In this section we turn to a recent proposal that affix order is also affected by a definitionally relative property: 'parsability', a complex equation whose main component is relative base frequency, measured over a corpus. To the extent that this property is a predictor of affix order, it reaffirms the concept embodied in OCM that affix order is governed by relative, not absolute, well-formedness.

Based largely on studies of derivational affix ordering in English, a body of recent work represented by Hay 2002, Plag 2002, Hay & Plag 2004, Hay & Baayen 2002, Hay & Baayen 2005, Plag & Baayen 2009 has advanced the compelling Complexity-Based Ordering hypothesis (Hay & Plag 2004), according to which relative morphological distance from the root should correlated with parsability: "an affix that can be easily parsed out should not occur inside an affix that cannot" (Hay 2002: 527–528). This hypothesis has also been supported for Russian suffixes by Parker and Sims (2012).

For each of 80 affixes in English, Hay and Baayen 2002 compare the relative frequency, for each stem containing such an affix, of the affixed stem and its unaffixed base, drawing for their corpus on the OED, BNC and CELEX. The degree to which an affix is considered parsable corresponds, primarily, to the degree that the frequency of the unaffixed base exceeds that of the unaffixed base. A parsability index can be constructed for each affix by averaging its parsability over all the types, or over all the tokens, of words containing it in a corpus.

Hay and Plag (2004) test the Complexity-Based Ordering hypothesis on 15 of the English derivational affixes studied by Fabb (1988) (see Section 6.1) *-en, -er, -ling, -ee, -ess, -ly, -dom, -ship, -hood, -ish, -less,* $\textit{-ful}_{Adj}$, *-ness,* $\textit{-ful}_{Noun}$, which occur in the order listed when they do (in smaller subsets) co-occur. Parsing ratios show a statistically significant correlation of parsability with observed relative order in the expected direction.

8. Ryan (2010) analyzes variable Tagalog ordering with affix bigram constraints. This, like OCM, is also a form-based approach, not one which can be implemented in a framework using (only) individualized M-Faith-F_i constraints to order affixes.

The Complexity-Based Ordering Hypothesis is alluring because it suggests that overall affix ordering is an emergent effect; though obtained from a corpus, parsability is a property of individual affixes, computed without reference to other affixes, and it can be represented lexically for each individual affix. Parsability indices permit the learner – or the stepwise word generator in a production model – to achieve a correct total ordering of all the affixes in a given word without recourse to global principles of affix ordering that could only act as post hoc filters, ruling out words with incorrect affix orders.

In OCM, implementing Complexity Based Ordering would amount to invoking a Parsability constraint which favors candidates formed by less parsable affixes over candidates formed by more parsable affixes. For example, in forming a word meaning 'in a manner that is fairly quick', the first cycle of word formation selects *quick* and the second has to choose between *-ly* and *-ish*. According to Hay & Plag's corpus results, *-ly-ish* is an attested order but *ish-ly* is not; correspondingly, *-ly* has a lower parsability score. In theory, either *quick-ish-ly* or *quick-ly-ish* could be interpreted with the desired meaning. By Parsability, candidate *quick-ly* would beat candidate *quick-ish* on the first suffix cycle, leaving it to the second suffix cycle to achieve the full desired meaning by adding *-ish.*[9]

However, it is not clear why such a synchronic constraint should exist. We understand historically that older affixes, grammaticalized earlier in a language's history, are closer to the root than new affixes, grammaticalized more recently; age correlates inversely with parsability, and so from a diachronic point of view the CBO hypothesis is well-grounded. As a synchronic constraint it lacks motivation.

The CBO hypothesis is also arguably more useful for semi-productive morphology of the kind exhibited in English rather than for more fully productive morphology in a highly affixing language such as Turkish. Although parsability scores have not, to my knowledge, been computed for Turkish suffixes, the obvious transparency and productivity of so much of the morphology suggests that the parsability metric is likely to result in ties, despite clear linear ordering tendencies among the suffixes.

9. A google search on 2/27/15 shows 3,850 hits for *quicklyish*, as in 'how to lose love handles quicklyish', and 4,750 hits for *quickishly*, as in 'how can I lose weight quickishly?'. This does not, of course, mean that both would be judged equally grammatical or used with equal likelihood in natural speech. However, finding both lends support to the possibility that these are legitimate candidates to consider in a word production model.

9. Taking stock

The larger goal of any generative model of morphology is to generate affixes in the correct order. The more specific goal of this particular paper is to explore the extent to which this can be done in OCM, using local optimization in a bottom-up, step-wise, target-driven manner.

At this juncture, we have seen that OCM can model templatic, lexical restrictions on affix ordering and co-occurrence through the use of selectional requirements associated with individual affixes. Through its general M-Faith constraint, OCM can model blocking and affix ordering based on syntactic scope without needing explicit principles for either phenomenon. OCM does not have a secure means of capturing the Relevance principle, and, until one is devised, is limited to stipulating the order of inflectional affixes via lexical selectional frames (or ranking individual M-Faith constraints, if desired). OCM can implement the Complexity-Based Ordering hypothesis if desired, but it would be arbitrary, emerging from no grounded synchronic principle that would make one word form better than its competitor.

10. Proposal: Entropy reduction and informativity

This section introduces a new factor into the mix, namely the possibility of incorporating information theory into a formal theoretical model of word formation. The proposal invokes Informativity Bias (IB), a constraint that favors affixes which do the most to reduce the hypothesis space that the grammar must entertain in the further construction of the word in question. The claim is that, all else being equal, a candidate formed by a more informative affix is better than a candidate formed by a less informative affix.

IP is based on the concept of 'informativity', a negative log measure of probability in context (Shannon 1948) shown in recent research to correlate with a variety of linguistic effects. Informativity correlates with syntactic processing (Levy 2008) and lexical decision latencies (Moscoso del Prado Martín, Kostic & Baayan 2004; Balling & Baayen 2012). Hirschberg and Pan (2000) find for English that the greater the informativity of a word in context, the likelier it is to be accented. In English noun-noun compounds, Bell and Plag (2012a,b) find a correlation between informativeness of the second member of a compound and the probability that it bears compound stress (CREDIT card vs. silk SHIRT). A growing collection of studies, including Kuperman et al. 2007; Hirschberg & Pan 2000; Gahl & Garnsey 2004, and Cohen Priva 2012, have documented a robust connection between predictability and pronunciation reduction. In a highly relevant study, Cohen Priva (2012) measures informativity for phonological segment types (e.g. the vowel /i/ vs. the vowel /a/), computing it based on the identity of the preceding segment in a large corpus, and finds that overall phone informativity across the corpus is a strong predictor of the tendency of that phone to reduce in any given lexical item.

And most significantly, a recent body of work has shown that predictability, or minimal entropy, is relevant in morphological learning (Ackerman & Malouf 2013; see also Moscoso del Prado Martín, Kostic & Baayen 2004; Blevins 2013; Seyfarth, Ackerman & Malouf 2014). In a discussion of paradigm entropy, Ackerman and Malouf write that

> One important requirement of any morphological system is that it must be possible for speakers to make accurate guesses about unknown forms of words based on exposure to known forms… speakers must generalize beyond their direct and limited experience of particular words to make likely inferences about unknown forms of that word within the system of paradigmatic relations characteristic of a specific language. [Ackerman & Malouf 2013: 436]

The novel proposal of this paper is to apply this same technique to affixes, testing not whether affix informativity predicts phonological reduction (although that would be a worthy project as well) but rather whether it correlates with average affix order.

In the formula below, from Cohen Priva's (2012) study of phone informativity, "c" is the identity of the preceding segment string; "σ" is the segment whose informativity is being measured. The formula sums the predictability of σ over all of the contexts in which occurs in the corpus ("Pr(c|σ)"), weighted in each case by the log probability of that context (the term "$\log_2$ Pr(s|c)").

(26) Informativity of σ in context "c" (Cohen Priva 2012)

$$-\sum_{c} \Pr(c|\sigma)\log_2 \Pr(\sigma|c)$$

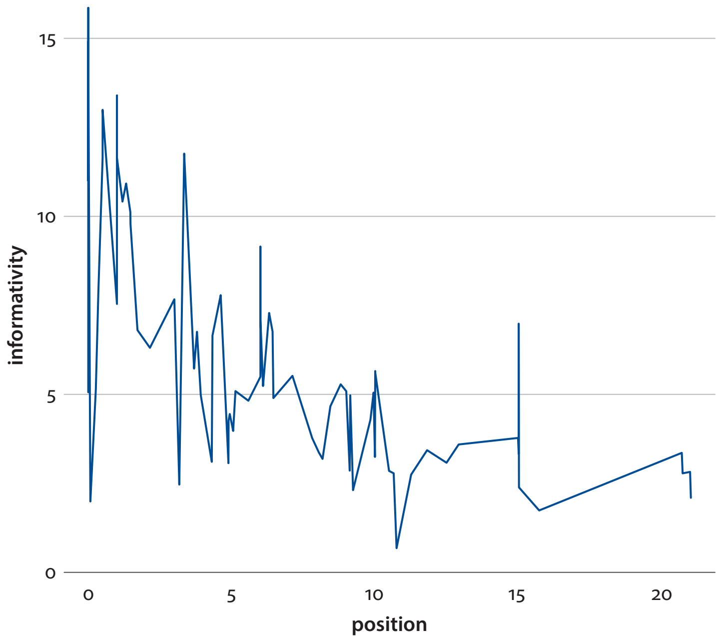
In a study of affix ordering, "σ" would represent the affix of interest; "c" would represent local morphological context. The hypothesis is that relative affix informativity, computed over a corpus of affixed words, correlates with relative affix order, or more precisely, that informativity will gradually increase (or gradually decrease) with average distance from the root.

As an illustration and preliminary test of this hypothesis, we summarize here the results of a preliminary study by Inkelas and J. Orgun (2015) of 54 of the 78 unique suffixes that occur in a corpus of over 611,000 unique words in Turkish, drawn from a text corpus of many million words compiled and morphologically analyzed by Kemal Oflazer.[10] Inkelas and J. Orgun computed, for each of 54 selected suffixes, its

10. Grateful acknowledgments to Kemal Oflazer for making this dataset available to Inkelas & J. Orgun. For documentation and discussion of the morphological anlayzer used to parse the forms in the word list, see e.g. Oflazer 1994 and Sak, Güngör & Saraçlar 2008.

informativity, based on the immediately preceding morpheme. An average position index was also calculated for each suffix, based on its relative distance from the root in words with 1–8 suffixes. Affix informativity was found to correlate inversely with position index. The more informative affixes are the inner ones, and informativity tends to decrease the farther away from the root an affix, on average, occurs. This negative correlation is graphically depicted below:

(27)



As an illustration, consider the Diminutive, Plural and Case suffixes. When these co-occur, they do so in a fixed relative order: Stem-DIM-PL-CASE.[11] Several examples from the corpus are given below:

(28)	badem-**cik**-**ler**-**e**	'almond-**DIM**-**PL**-**DAT**'
	tane-**cik**-**ler**-in-**den**	'piece-**DIM**-**PL**-2SG.POSS-**ABL**'
	gelin-**cik**-**ler**-**in**	'daughter-in-law-**DIM**-**PL**-**GEN**'
	tepe-**cik**-**ler**-in-**i**	'hill-**DIM**-**PL**-2SG.POSS-**ACC**'

Theoretically, as discussed in Section 7.2, the affixes could occur in a different order and still be interpretable. In particular, diminutive and plural affixes are known to vary in their order cross-linguistically (see e.g. Steriopolo 2013), though the more dominant

11. Turkish has five overt case endings: Accusative, Dative, Locative, Ablative, Genitive. Nominative case is not marked with an overt affix. The question of how or whether to calculate informativity for a phonologically null morphological category is not addressed here.

pattern is for diminutive affixation to occur closer to the root. Some construct in the theory determines this. Could it be informativity?

While there is some variation in informativity over the five overt case endings in Turkish, on average, informativity correlates with order for the suffixes illustrated in (27). The average informativity for case suffixes is 2.77, lower than for Diminutive and Plural:

(29)

Affix Id	Mean Position Index	Informativity	
DIM	0.50	11.64	11.64
PL	9.99	3.24	3.24
LOC	10.64	2.78	
ABL	20.68	3.36	
GEN	20.71	2.78	
DAT	20.97	2.82	
ACC	21.00	2.10	2.77

To take a second example, in Turkish verbs, aspect is expressed closer to the root than conditional mood, which in turn is inside of (most) person-number inflectional suffixes. Examples are given below:

(30) bil-iyor-sa-n 'know-PROG-COND-2SG'
ver-iyor-lar-sa 'give-PROG-PL-COND'
tüket-iyor-sa-nız 'prog-COND-2PL

There are numerous scattered counterexamples to Informativity-based ordering. For example, the Future tense suffix has a lower informativity score (2.86) than Conditional (3.77), yet Future precedes Conditional in verbs containing both.

The Informativity Privilege is clearly a violable principle. Yet the statistical correlation tests reported above indicate that it is capable of playing a predictive role.

Based on this preliminary result from one language, we posit a (relatively low-weighted) constraint, INFORMATIVITY BIAS, which helps to choose among affixes all of which would be compatible with a given input stem and all of which would improve target-matching. IB favors the affix with highest informativity.

(31) INFORMATIVITY BIAS: Given an input base B and a set of affixed candidates {B+Aff$_1$, B+Aff$_2$, … B+Aff$_n$}, penalize an affixed candidate B+Aff$_i$ by an amount proportional to the degree to which the informativity of Aff$_i$ is lower than that of the most informative affix in the set of competing candidates.

The candidate with the most informative affix will satisfy this comparative constraint; all others will exhibit some degree of violation.

The tableau in (32) illustrates IB at work, imposing a relative prioritization on the addition of Diminutive, Plural and Case endings in the bottom-up construction of the Turkish word *bademciklere*, from (28). Informativity scores for Diminutive (11.64), Plural (3.24) and Dative (2.77) suffixes are provided in the lookup table in (29). The most informative is the Diminutive; subtracting the informativity score of Plural and Dative from this yields the violations of IB incurred by candidates (32b) and (32c). In this tableau, the three candidates tie on M-Faith because all encode comparably-sized chunks of the meaning target M. IB breaks the tie in favor of the informative Diminutive (32a):

(32)

Cycle 2	Input stem: *badem* Almond		M: Almond dim pl dat	M-Faith- num	IB
☞ a.	[[ ] CIK] dat	*badem-cik*	almond dim	** (pl, dat)	0
b.	[[ ] lEr] pl	*badem-ler*	almond pl	** (dim, dat)	8.4
c.	[[ ] E] dat	*badem-e*	almond dat	** (dim, pl)	8.82
d.	ID	*badem*	almond	***! (dim, pl,dat)	N/A

IB continues to exert its effects on the next cycle of affixation, choosing (among the competitors which improve over the base in M-Faithfulness) the competitor with the relatively greater Informativity score, namely Plural:

(33)

Cycle 3	Input stem: *bademcik* Almond, dim		M: Almond dim pl dat	M-Faith- num	IP
a.	[[ ] CIK] dim	*bademcik-cik*	almond dim	**! (pl, dat)	0
☞ b.	[[ ] lEr] pl	*bademcik-ler*	almond dim pl	* (dat)	8.4
c.	[[ ] E] dat	*bademcik-e*	almond dim dat	* (pl)	8.82
d.	ID	*bademcik*	almond dim	**! (pl,dat)	N/A

The third cycle of affixation completes the derivation. Full M-faithfulness is achieved by adding the Dative suffix:

(34)

Cycle 4	Input stem: *bademcikler* ALMOND, DIM,PL		M: ALMOND DIM PL DAT	M-FAITH-NUM	IP
a.	[[ ] CIK] DIM	*bademcikler-cik*	ALMOND DIM PL	*! (DAT)	0
b.	[[ ] lEr] PL	*bademcikler-ler*	ALMOND DIM PL	*! (DAT)	8.4
☞ c.	[[ ] E] DAT	*bademcikler-e*	ALMOND DIM PL DAT		8.82
d.	ID	*bademcikler*	ALMOND DIM PL	*! (DAT)	N/A

Note that the candidates with recursive suffixation, e.g. (34a, b), are shown here just in order for the candidates to be comparable across cycles. The assumption in the above tableaus is that adding the same inflectional suffix twice does not affect meaning. Thus, a candidate which undergoes redundant affixation will always lose on M-FAITH to a candidate which adds a suffix that adds meaning and brings the candidate closer to the meaning target. These tableaux should not be taken to imply that in general, recursive affixation is semantically or syntactically vacuous; that is not at all true cross-linguistically, or even in Turkish (which exhibits recursive causative suffixation, for example).

11. Conclusion

This paper, while necessarily speculative and programmatic, offers a means within one single theoretical framework of integrating the many, diverse, and often competing factors which can contribute to the overall scheme of affix ordering in a given language. The particularly novel introduction is Informativity Bias; it is proposed that such a constraint may be responsible for some general ordering tendencies.

It is natural to ask, of Turkish or any other language with relatively fixed affix order, why any functionally grounded principles at all are necessary to the formal modeling of the system. Why is it not sufficient to observe attested orderings of affixes and, from there, compile a flow chart that illustrates all of the possible orderings? The answer to this question is that it could work (and for a suffix flow chart for Turkish, see Hankamer 1986). But at a conceptual level, we are always seeking principles that can motivate surface patterns. And in the acquisition process, something like Relevance or the Informativity Bias may help fill in gaps due to partial exposure to data. Eyigöz, Gildea and Oflazer (2013) report, of a 50,000-word corpus of Turkish texts, an average word length of 2.69 morphemes (vs. 1.57 in English words taken from CELEX). If, on average, the learner of Turkish is encountering less than two suffixes per word, any and all available bootstrapping methods and biases will be helpful in predicting affix order in a word containing a much larger number of suffixes.

Only future work will reveal if the correlation between informativity and relative affix order holds up on closer inspection in Turkish or in other languages. If it does, one natural question to ask is about the direction of causation over time. Does informativity determine order, or does order, over time, determine informativity? It is well known that affixes closest to the root tend to lexicalize with the root, change their meanings, and become less predictable in their distribution, just as they also become less parsable. Certainly in a synchronic model whose users are not aware of the time depth of their affixes, some synchronic principle is needed to capture the correlation. Whether this principle should be something like Bybee's Relevance (a difficult concept to quantify, but intuitive), parsability (easy to quantify, given a very large text corpus), or informativity (easy to quantify, given a modest corpus), or whether all are needed, is a subject for future research.

References

Ackerman, Farrell & Malouf, Rob. 2013. Morphological organization: The low conditional entropy conjecture. *Language* 89(3): 429–464. doi:10.1353/lan.2013.0054

Alsina, Alex. 1999. Where's the Mirror Principle? *The Linguistic Review* 16(1): 1–42. doi:10.1515/tlir.1999.16.1.1

Anderson, Stephen R. 1992. *A-morphous Morphology*. Cambridge: CUP. doi:10.1017/CBO9780511586262

Aronoff, Mark. 1994. *Morphology by Itself: Stems and Inflectional Classes*. Cambridge MA: The MIT Press.

Aronoff, Mark & Xu, Zheng. 2010. A realization Optimality-Theoretic approach to affix order. *Morphology* 20(2): 381–411. doi:10.1007/s11525-010-9181-2

Baker, Mark. 1988. *Incorporation: A Theory of Grammatical Function Changing*. Chicago ILI: Chicago University Press.

Balling, Laura Winther. & Baayen, R. Harald. 2012. Probability and surprisal in auditory comprehension of morphologically complex words. *Cognition* 125(1). 80–106. doi:10.1016/j.cognition.2012.06.003

Beesley, Kenneth & Karttunen, Lauri. 2003. *Finite State Morphology*. Stanford CA: CSLI.
Bell, Melanie & Plag, Ingo. 2012. Informativeness is a determinant of compound stress in English. *Journal of Linguistics* 48: 485–520. doi:10.1017/S0022226712000199
Blevins, James. 2003. Stems and paradigms. *Language* 79(4): 737–767. doi:10.1353/lan.2003.0206
Blevins, James. 2013. The information-theoretic turn. *Psihologija* 46(3): 355–375. doi:10.2298/PSI1304355B
Booij, Geert. 2010. *Construction Morphology*. Oxford: OUP.
Bybee, Joan. 1985. *Morphology: A Study of the Relation between Meaning and Form* [Typological Studies in Language 9]. Amsterdam: John Benjamins. doi:10.1075/tsl.9
Caballero, Gabriela & Inkelas, Sharon. 2013. Word construction: Tracing an optimal path through the lexicon. Special issue *New theoretical tools in the modeling of morphological exponence. Morphology* 23: 103–143. doi:10.1007/s11525-013-9220-x
Cohen Priva, Uriel. 2012. Sign and Signal: Deriving Linguistic Generalizations from Information Utility. Ph.D. dissertation, Stanford University.
Eyigöz, Elif, Gildea, Daniel & Oflazer, Kemal. 2013. Simultaneous word-morpheme alignment for statistical machine translation. *Proceedings of the 2013 meeting of the North American chapter of the Association for Computational Linguistics (NAACL-13)*.
Fabb, Nigel. 1988. English suffixation is constrained only by selectional restrictions. *Natural Language and Linguistic Theory* 6(4): 527–539. doi:10.1007/BF00134491
Fulmer, S. Lee. 1991. Dual-position affixes in Afar: an argument for phonologically-driven morphology. In *Proceedings of the Ninth Annual Meeting of the West Coast Conference on Formal Linguistics*, Aaron Halpern (ed.), 189–203. Stanford CA: CSLI.
Gahl, Susanne & Garnsey, Susan. 2004. Knowledge of grammar, knowledge of usage: syntactic probabilities affect pronunciation variation. *Language* 80(4): 748–775. doi:10.1353/lan.2004.0185
Hankamer, Jorge. 1986. Finite state morphology and left to right phonology. In *Proceedings of the Fifth West Coast Conference on Formal Linguistics,* Mary Dalrymple, Jeffrey Goldberg, Kristin Hanson, Michael Inman, Chris Piñon & Stephen Wechsler (eds), Stanford CA: Stanford Linguistics Association.
Hargus, Sharon. 1993. *Metathesis and Domain Membership in Sekani*. Berkeley CA: Department of Linguistics, UC Berkeley.
Harley, Heidi & Rolf Noyer. 1999. Distributed Morphology. *GLOT International* 4(4): 3–9.
Harris, Alice. 2000. Where in the word is the Udi clitic? *Language* 76(3): 593–616. doi:10.2307/417136
Hayes, Bruce & Wilson, Colin. 2008. A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* 39: 379–440. doi:10.1162/ling.2008.39.3.379
Hay, Jennifer. 2002. From speech perception to morphology: Affix ordering revisited. *Language* 78(3): 527–555. doi:10.1353/lan.2002.0159
Hay, Jennifer & Baayen, R. Harald. 2002. Parsing and productivity. In *Yearbook of Morphology 2001,* Geert Booij & Jaap van Marle (eds), 203–235. Dordrecht: Kluwer. doi:10.1007/978-94-017-3726-5_8
Hay, Jennifer & Harald Baayen. 2005. Shifting paradigms: Gradient structure in morphology. *Trends in Cognitive Sciences* 9(7): 342–348. doi:10.1016/j.tics.2005.04.002
Hay, Jennifer & Plag, Ingo. 2004. What constrains possible suffix combinations? On the interaction of grammatical and processing restrictions in derivational morphology. *Natural Language and Linguistic Theory* 22(3): 565–596. doi:10.1023/B:NALA.0000027679.63308.89

Hirschberg, Julia & Pan, Shimei. 2000. Modeling local context for speech accent prediction. *Proceedings of the 38th Annual Meeting of the Association for Computational Linguistics*.

Hyman, Larry M. 2003. Suffix ordering in Bantu: A morphocentric approach. In *Yearbook of Morphology 2002,* Geert Booij & Jaap van Marle (eds), 245–281. Dordrecht: Kluwer. doi:10.1007/0-306-48223-1_8

Inkelas, Sharon. 1989. Prosodic Constituency in the Lexicon. Ph.D. dissertation, Stanford University.

Inkelas, Sharon. 1993. Nimboran position class morphology. *Natural Language and Linguistic Theory* 11: 559–624. doi:10.1007/BF00993014

Inkelas, Sharon & Orgun, Cemil Orhan. 1998. Level (non)ordering in recursive morphology: Evidence from Turkish. In *Morphology and its Relation to Phonology and Syntax,* Steven Lapointe, Diane Brentari & Patrick Farrell (eds). Stanford CA: CSLI.

Inkelas, Sharon & Orgun, Jem. 2015. Informativity and affix ordering: A pilot study of Turkish. Portland OR. ⟨http://linguistics.berkeley.edu/~inkelas/Papers/InkelasOrgunLSAPoster2015.pdf⟩

Itô, Junko, Mester, Armin & Padgett, Jaye. 1995. Licensing and underspecification in Optimality Theory. *Linguistic Inquiry* 26(4): 571–614.

Kim, Yuni. 2008. Topics in the Phonology and Morphology of San Francisco del Mar Huave. Ph.D. dissertation, University of California, Berkeley.

Kiparsky, Paul. 1982a. Lexical morphology and phonology. In *Linguistics in the Morning Calm*, Linguistic Society of Korea (eds), 3–91. Seoul: Hanshin.

Kiparsky, Paul. 1982b. Word-formation and the lexicon. In *1982 Mid-America Linguistics Conference Papers*, Frances Ingemann (ed.), 3–32. Lawrence KS: University of Kansas.

Kuperman, Victor, Pluymaekers, Mark, Ernestus, Mirjam & Baayen, R. Harald. 2007. Morphological predictability and acoustic duration of interfixes in Dutch compounds. *Journal of the Acoustical Society of America* 12(4): 2261–2271. doi:10.1121/1.2537393

Legendre, Geraldine, Miyata, Yoshiro & Smolensky, Paul. 1990. *Harmonic Grammar: A Formal Multi-level Connectionist Theory of Linguistic Well-formedness. Theoretical Foundations.* Boulder CO: Department of Computer Science, University of Colorado.

Levy, Roger. 2008. Expectation-based syntactic comprehension. *Cognition* 106: 1126–1177. doi:10.1016/j.cognition.2007.05.006

Lieber, Rochelle. 1980. On the Organization of the Lexicon. Ph.D. dissertation, MIT.

Luutonen, Jorma. 1997. *The Variation of Morpheme Order in Mari Declension*. Helsinki: Suomalais-Ugrilainen Seura.

Maiden, Martin. 2005. Morphological autonomy and diachrony. In *Yearbook of Morphology 2004,* Geert Booij & Jaap van Marle (eds), 137–75. Dordrecht: Kluwer. doi:10.1007/1-4020-2900-4_6

Manova, Stella & Aronoff, Mark. 2010. Modeling affix order. *Morphology* 20: 109–131. doi:10.1007/s11525-010-9153-6

van Marle, Jaap. 1998. Review of Jorma Luutonen, The Variation of Morpheme Order in Mari Declension. In *Yearbook of Morphology 1997*, Geert Booij & Jaap van Marle (eds), 307–308. Dordrecht: Kluwer.

McCarthy, John. 2002. *A Thematic Guide to Optimality Theory*. Cambridge: CUP.

McCarthy, John. 2010. An introduction to harmonic serialism. *Language and Linguistics Compass* 4(10): 1001–1018. doi:10.1111/j.1749-818X.2010.00240.x

McCarthy, John & Prince, Alan. 1993. Generalized alignment. In *Yearbook of Morphology 1993*, Geert Booij & Jaap van Marle (eds), 79–153. Dordrecht: Kluwer. doi:10.1007/978-94-017-3712-8_4

McCarthy, John & Prince, Alan. 1995. Faithfulness and reduplicative identity. In *University of Massachusetts Occasional Papers in Linguistics 18: Papers in Optimality Theory*, Jill Beckman, Laura Dickey & Suzanne Urbanczyk (eds), 249–384. Amherst MA: GLSA.

McFadden, Thomas. 2004. The Position of Morphological Case in the Derivation: A Study on the Syntax-morphology Interface. Ph.D. dissertation, University of Pennsylvania.

Mohanan, Karuvannur P. 1986. *Lexical Phonology*. Dordrecht: Kluwer.

Mohanan, Karuvannur P. 1995. The organization of the grammar. In *The Handbook of Phonological Theory*, John Goldsmith (ed.). Malden MA: Blackwell.

Moscoso del Prado Martín, Fermín, Kostic, Aleksandar & Baayen, R. Harald. 2004. Putting the bits together: An information theoretical perspective on morphological processing. *Cognition* 94: 1–18. doi:10.1016/j.cognition.2003.10.015

Noyer, Rolf. 1993. Optimal Words: Towards a declarative theory of word formation. Rutgers University. ⟨ftp://babel.ling.upenn.edu/facpapers/rolf_noyer/optimal_words.ps⟩

Noyer, Rolf. 1994. Mobile affixes in Huave: Optimality and morphological wellformedness. In *Proceedings of the Twelfth West Coast Conference on Formal Linguistics*, Erin Duncan, Donka Farkas & Philip Spaelti (eds), 67–82. Stanford CA: Stanford Linguistics Association.

Oflazer, Kemal. 1994. Two-level description of Turkish morphology. *Literary and Linguistic Computing* 9(2): 137–148. doi:10.1093/llc/9.2.137

Parker, Jeff & Sims, Andrea. 2012. Affix ordering constraints and processing in Russian: A look at Complexity-Based ordering. Ms, University of Massachusetts, Amherst.

Paster, Mary. 2005. Pulaar verbal extensions and phonologically driven affix order. In *Yearbook of Morphology,* Geert Booij & Jaap van Marle (eds), 155–199. Dordrecht: Springer. doi:10.1007/1-4020-4066-0_6

Paster, Mary. 2006. Phonological Conditions on Affixation. Ph.D. dissertation, University of California, Berkeley.

Plag, Ingo. 2002. The role of selectional restrictions, phonotactics and parsing in constraining suffix ordering in English. In *Yearbook of Morphology 2001*, Geert Booij & Jaap van Marle (eds), 285–314. Dordrecht: Kluwer. doi:10.1007/978-94-017-3726-5_11

Plag, Ingo & Baayen, R. Harald. 2009. Suffix ordering and morphological processing. *Language* 85(1): 109–152. doi:10.1353/lan.0.0087

Prince, Alan & Smolensky, Paul. 1993. *Optimality Theory: Constraint Interaction in Generative Grammar*. New Brunswick NJ: Rutgers University.

Prince, Alan & Smolensky, Paul. 2004. *Optimality Theory: Constraint Interaction in Generative Grammar*. Malden MA: Wiley-Blackwell. doi:10.1002/9780470759400

Rice, Keren. 2000. *Morpheme Order and Semantic Scope: Word Formation in the Athapaskan Verb*. Cambridge: CUP. doi:10.1017/CBO9780511663659

Rice, Keren. 2011. Principles of affix ordering: An overview. *Word Structure* 4: 169–200. doi:10.3366/word.2011.0009

Ryan, Kevin. 2010. Variable affix order: Grammar and learning. *Language* 86: 758–791.

Saarinen, Pauliina & Hay, Jennifer. 2014. Affix ordering in derivation. In *Oxford Handbook of Derivational Morphology*, Rochelle Lieber & Pavol Stekauer (eds), 370–383. Oxford: OUP.

Sak, Haşim, Güngör, Tunga & Saraçlar, Murat. 2008. Turkish language resources: Morphological parser, morphological disambiguator and web corpus. *Proceedings of GoTAL 2008*, 417–427. doi:10.1007/978-3-540-85287-2_40

Selkirk, Elisabeth. 1982. *The Syntax of Words*. Cambridge MA: The MIT Press.
Seyfarth, Scott, Ackerman, Farrell & Malouf, Robert. 2014. Implicative organization facilitates morphological learning. In *Proceedings of the 40th annual meeting of the Berkeley Linguistics Society*, Herman Leung, Zachary O'Hagan, Sarah Bakst, Auburn Lutzross, Jonathan Manker, Nicholas Rolle & Katie Sardinha (eds), 480–494. Berkeley CA: BLS. doi: 10.3765/bls.v40i0.3154
Shannon, Claude. 1948. A mathematical theory of communication. *Bell System Technical Journal* 27: 379–423, 623–656. doi: 10.1002/j.1538-7305.1948.tb01338.x
Simpson, Jane & Withgott, Margaret. 1986. Pronominal clitic clusters and templates. In *Syntax and Semantics,* 19: *The Syntax of Pronominal Clitics,* Hagit Borer (ed.). New York NY: Academic Press.
Spencer, Andrew. 2003. Putting some order into morphology; Reflections on Rice (2000) and Stump (2001). *Journal of Linguistics* 39: 621–646. doi: 10.1017/S0022226703002123
Spencer, Andrew. 2013. *Lexical Relatedness: A Paradigm-Based Model*. Oxford: OUP. doi: 10.1093/acprof:oso/9780199679928.001.0001
Steriopolo, Olga. 2013. Diminutive affixes in the number domain: A syntactic variation. *Questions and Answers in Linguistics* 1(2): 33–56. doi: 10.1515/qal-2015-0003
Stump, Gregory. 2001. *Inflectional Morphology: A Theory of Paradigm Structure*. Cambridge: CUP. doi: 10.1017/CBO9780511486333
Wolf, Matthew. 2008. Optimal Interleaving: Serial Phonology-Morphology Interaction in a Constraint-based Model. Ph.D. dissertation, University of Massachusetts, Amherst.
Xu, Zheng & Aronoff, Mark. 2011. A realization Optimality-Theoretic approach to extended morphological exponence. *Journal of Linguistics* 47(3): 673–707. doi: 10.1017/S002222671100003X
Yu, Alan C.-L. 2007. *A Natural History of Infixation*. Oxford: OUP. doi: 10.1093/acprof:oso/9780199279388.001.0001