Allomorphy and Vocabulary Insertion*

Maria Gouskova (New York University) and Jonathan David Bobaljik (Harvard)

1 Introduction

1.1 Separation and Vocabulary Insertion

Any theory of morphology must answer two questions about allomorphy: first, what considerations govern the form of morphemes, and second, under what conditions can one morpheme influence the shape of another. Distributed Morphology (Halle and Marantz 1993) allows allomorphy to be conditioned in many components of the grammar: syntax, phonology, and the lexicon. In no small part, the specific pronunciation of morphemes depends on rules of Vocabulary Insertion—the subject of this review.

Distributed Morphology (DM) inherits certain assumptions about the lexicon from *the Sound Pattern of English* (Chomsky and Halle 1968). The lexicon stores the pronunciations of morphemes, sometimes as abstract phonological forms that do not correspond to any surface allomorph. Surface allomorphs can then be derived by phonological rules. Furthermore, central to the theory is the assumption that morphemes are abstract entities that relate syntactic features (e.g., PLURAL or PAST) to pronunciations (*vocabulary items*), and that there is a feeding relationship between syntactic derivations and morphological and phonological rules. The syntactic component operates on abstract morphemes, not on pronunciations. The specific details of any given morpheme's pronunciation cannot affect its behavior in the narrow syntax or semantics, but the pronunciation may play a role in morphological patterns.

In a precursor to DM, Halle (1990) articulates the argument for separating the abstract syntactic features from their phonological realizations as follows. In English, the plural takes many forms: it is null in *sheep* (SG)∼*sheep* (PL), it effects a vowel change in *goose*∼*geese*, it entails the removal of *-us* and addition of *-i* in *alumn-us*∼*alumn-i*, and it is realized as the suffix */-z/* for the vast majority of nouns such as *dog*∼*dog-s*. Regardless of how the plural is pronounced for any given item, all of the plurals share in some syntactic behaviors: for example, they condition the determiner *these (these {sheep, dogs},* etc.) and agreement on the verb (*{sheep, dogs, alumni} walk*, not **walks*). Indeed, in languages with more elaborate morphological systems, there are certain morphological features that systematically affect other words in the sentence (e.g., gender, tense, or case), whereas others affect only the shape of morphemes inside individual words (declension and conjugation classes). The explanation Halle offers is that gender and number are properly syntactic features, and the rules that concern them are active in the syntax before the pronunciations of

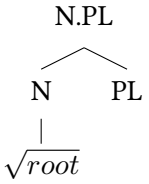
*We would like to thank the participants of the NYU Morphology Research Group for valuable feedback, which resulted in numerous improvements. Special thanks to Hagen Blix, Christos Christopoulos, Maddie Gilbert, Richie Kayne, Yining Nie, Mary Robinson, Juliet Stanton, and Guy Tabachnick.

specific morphemes are decided upon. The shape of *these* is determined by the abstract PLURAL. PLURAL is then converted to specific pronunciations, depending on the stem in question, by individual Vocabulary Insertion rules.

1.2 Vocabulary Insertion and Allomorphy, a preview

The architectural assumption of Separation makes available at least two sources of allomorphy, which we can continue to illustrate with the English plural. We may assume that the syntax of plural nouns in English is uniform, as in (1)—all plural nouns include at a minimum an abstract PLURAL morpheme that combines with a nominal root.[1] As Halle emphasized, syntax does not care about the differing pronunciations of the PL node (including zero).

(1)



Vocabulary Insertion, the focus of this overview, is the primary mechanism for relating abstract morphosyntactic representations like (1) to more concrete phonological representations. The *Vocabulary* of a language is a list of rules (corresponding to rules of exponence in other frameworks, such as Matthews 1972), where each individual rule is a *vocabulary item*. In the basic case, rules of Vocabulary Insertion apply at each terminal node in the tree, proceeding cyclically from the most deeply embedded node upwards/outwards (Bobaljik 2000, Wolf 2008, Myler 2017). (Some versions of DM treat the root node as special in this regard, an issue we return to below). An important potential source of allomorphy, then, is competition among multiple rules that share a structural description. English, like many languages, has a variety of exponents of the feature PLURAL, perhaps including those in (2):

(2) a. [PL] ↔ -en / $]_N$ ___, where N ∈ {ox}

b. [PL] ↔ Ø/ $]_N$ ___, where N ∈ {sheep, fish, ...}

c. [PL] ↔ [−back, +tense] / $]_N$ ___, where N ∈ {foot, goose, ...}

d. [PL] ↔ -i / $]_N$ ___, where N ∈ {alumn-, syllab-, ...}

e. [PL] ↔ -z / $]_N$ ___

In the simplest of cases, there is no need to specify ordering among these rules, other than the intrinsic order imposed by the Elsewhere Condition (Kiparsky 1973). All but the last rule in (2) have mutually exclusive lexical conditions, which characterize the lexically-specific nature of this allomorphy. The Elsewhere Condition ensures that insertion of the context-free (default) vocabulary item /-z/ is blocked when there is a more specific rule that takes precedence.

Competition among vocabulary items constitutes *suppletive* allomorphy, and represents distinct memorized forms. Suppletive allomorphy contrasts with allomorphy governed by phonological rules, such as

[1] For illustration, we ignore further questions not immediately relevant, such as the question of whether there is a separate categorizing morpheme, little *n*.

the voicing and schwa-zero alternations in the regular plurals: *bag-s* [z], *back-s* [s], *batch-es* [əz]. This alternation is rule governed, phonologically predictable, and is therefore modeled in DM as arising from the insertion of a single vocabulary item /-z/, whose variation in surface form is effected by subsequent phonological rules.

In this way, basic architectural assumptions allow for a crisp theoretical distinction between *allomorphy by competition* and *allomorphy by rule-governed change*, but the architecture itself does not guarantee analytical determinism: much ink has been spilled over analytical debates about particular examples. Some such debates extend beyond the mere analytical uncertainty and bear, for example, on venerable questions about how abstract phonological representations may be, and what types of (morpho)-phonological rules should be countenanced. In the course of this overview, we will highlight some of these debates, without taking a stand on their resolution.

1.3 Roadmap

The remainder of the article is organized as follows. We start with a general review of what allomorphy is, and how it has been classified analytically both in DM and in other theories of morphology and phonology (§2). In §3, we show how a couple of examples (more complex than the English plural) would be analyzed in traditional DM; these present various issues that we consider in subsequent sections. One is whether, from a structural point of view, allomorphy is conditioned differently by inner vs. outer context (§4.1). Another is how far away the conditioning environment can be from the affected morpheme (§4.2). Section 5 considers various theory-internal debates within DM, including the role of linear adjacency, how competition for insertion is resolved, and whether this competition is a property of roots. In §6 we consider broader questions about allomorphy as a phonological and a morphosyntactic phenomenon. Section 7 is the conclusion.

2 Kinds of allomorphy

The term "allomorphy" is used in slightly different ways in different traditions. In one view, any alternation in the form of a morpheme is allomorphy. Many authors, on the other hand, tend to use the term when they mean suppletive allomorphy (defined later in this section). We will use "allomorphy" in the more general morphological sense, since the line between phonological alternations and suppletion is largely a matter of analysis.

Some alternations are general—every eligible morpheme undergoes them, and the alternation is part of a strong phonotactic generalization. Final devoicing in Russian is like this (e.g., [sapo<u>g</u>-om] 'boot INST.SG' vs. [sapo<u>k</u>] 'boot NOM.SG'; see (3)). A traditional phonological analysis would argue that the rule to derive voiceless stops word-finally is independently needed, since even non-alternating morphemes always obey it. If general alternations are analyzed this way, then morphology has nothing to say about them. A single vocabulary item is inserted to realize the morpheme, and all surface variants of that pronunciation are derived by phonological rules, not morphological ones (Chomsky and Halle 1968 et seq.).[2]

[2]Not all phonologists adopt the unique/abstract UR approach. Exemplar Theory posits that all variant pronunciations are

In the analytic gray zone of allomorphy are cases such as English prefixes that show nasal place assimilation (*i[n]-tolerant* vs. *i[m]-possible* vs. *i[ŋ]-coherent*; *e[n]-tomb* vs. *e[m]-bed; co[n]-temporary vs. co[m]-patriot*). It is not clear that English requires nasal place assimilation in general; there are certainly morphemes that do not not assimilate or do so only optionally (*u[n]-deserving, u[n]-balanced*). Morpheme-internally, most morphemes have place agreement in nasal-stop clusters, but there are some morpheme-internal exceptions (*Canberra*), and the rule does not apply to stem-final nasals (*hem*∼ *he[md]*, not **he[nd]*), or to phonological clitics such as the prepositions *in* and *on*, etc. There is no general agreement among phonologists or morphologists about the nature of allomorphy in cases such as *in-*/*im-*/*iŋ*; it is not even clear whether undergoing the rule is the norm or the exception (Chomsky and Halle 1968, Mohanan 1982, Borowsky 1986). Complicating the picture is the observation that non-automatic rules often strongly resemble rules that are fully general in other languages, which some take as an argument in favor of treating them as phonological rules with a limited scope (e.g., Jusczyk et al. 2002 on English nasal place assimilation). If this is right, then these alternations are a phonological problem, not a morphological one. In some phonological frameworks, however, any non-automatic alternations are treated as suppletive allomorphy or readjustment rules that apply in the morphology (Green 2007, Mascaró 2007, Bye and Svenonius 2012).

Relatedly, some allomorphy is conditioned by specific morpheme triggers. This can mean a number of things, from specific vocabulary items (as in (3)) to morphosyntactic contexts (see §3.1 and §3.2, as well as §4)). Example (3) is from Russian, where the velar consonants [k, x, g] become [ʧ, ʂ, ʐ] in the context of some suffixes (such as the diminutive suffix [-ok]), but not others that are phonologically similar ([-om] INST.SG). There is no rule in Russian that requires velars to turn into stridents, either across a morpheme boundary or morpheme-internally (see Padgett 2003 and others). But while the triggers of velar mutation appear to be arbitrary, the undergoers are not—any noun that combines with the diminutive will undergo the rule (Kapatsinski 2010, Gouskova et al. 2015). The treatment of this type of allomorphy has been controversial both in phonology and in morphology (see 6.1.1 and §6.1.2, which discuss some of the issues of morphological interest).

(3) Russian velar palatalization/mutation: triggered by the diminutive, but not the inst. sg. suffix

		NOM. SG.	INST. SG.	DIM. NOM. SG.	Gloss
a.	/sapog/	sapok	sapog-om	sapoʐ-ok	‘boot’
b.	/kulak/	kulak	kulak-om	kulaʧ-ok	‘fist’
c.	/grex/	grex	grex-om	greʂ-ok	‘sin’
d.	/zub/	zup	zub-om	zub-ok	‘tooth’
e.	/god/	got	god-om	god-ok	‘year’

At the other end of the continuum from the fully productive phonological alternations lies suppletion. In suppletive allomorphy, there is no phonological resemblance between the allomorphs:

(4) Examples of suppletive allomorphy

stored, including fine-grained phonetic detail (Pierrehumbert 2001, Bybee 2001, Green 2007 and others). Phonological generalizations are extracted through analogy in this framework—the more alternations support an analogical rule, the more likely it is to be productive.

a. The forms of the English verb "be" (*am, is, were*, etc.) are suppletive across person and tense.

b. Several English adverbs and adjectives show suppletion in the comparative and the superlative (*good* ∼ *better* ∼ *best*).

c. In Moroccan Arabic, the 3 p.sg. possessive is either *-h* or *-u*, depending on whether the noun ends in a consonant or a vowel ([ktab-u] 'his book', [htˤa-h] 'his error', Mascaró 1996, Wolf 2015).

d. The realizations of the plural in English as *-i*, *-en*, or *-s* are suppletive (see (2)).

Suppletion of roots (as in (4b)) is sometimes analytically distinguished from affixal suppletion (as in (4c–d))—not all authors consider this type of affixal allomorphy to be suppletion. Mel'čuk (1976, 1994) defines both as suppletion, but Corbett (2007) considers only root allomorphy in inflectional contexts to be suppletion. Root suppletion has been debated in DM, as well; see §5.4. To get a sense of why root suppletion is controversial, consider typical examples seen in the literature, such as the English *person* (sg) vs. *people* (pl) and Russian [reb^{j}onok-Ø] 'child.SG' [det^{j}-i] 'child.PL'. In the English case, it is unclear what to make of the existence of *persons* and *peoples*, as well as *a people* vs. *many people*. The Russian case is similar: [dit^{j}-a] 'child.SG' exists alongside [reb^{j}-onok-Ø], as does [reb^{j}-at-a] 'guys, kids' (see §3.2 for more on *-onok/-at*). The question for a theory such as DM is whether root suppletion and affix suppletion differ in any substantial way; the existing evidence suggests that they obey similar locality constraints (see §4.1).

3 Basic features of vocabulary insertion in DM

3.1 Sources of allomorphy

We start by working through a more elaborate case of allomorphy, the traditional DM analysis of which involves some suppletion, some readjustment rules, and some phonological rules. German "strong" roots alternate in verbal paradigms, appearing with different vowels depending on person and tense (e.g., the root of *geben* appears with [e], [a] or [ɪ]—see (6)). The undergoers here are lexically specific—vowels in phonologically similar weak verbs do not alternate (see (5)). Alongside root alternations, the past tense suffixes show allomorphy between strong and weak verbal paradigms, as well: a [-t] suffix consistently appears in the past tense of weak verbs, but not in strong ones. One could say that the vowel change [e] to [a] in [gab-ən] is conditioned morphosyntactically, by the tense feature. Alternatively, one could say that the vowel itself (or [+back]) expresses the past tense feature in strong verbs, whereas [-t] expresses it in weak ones. Moreover, as shown in (7), verbs such as *denken* show suppletion in the root between present and past tense, while the affixes are the same as in weak verbs.

(5) German regular (weak) verbs lack ablaut: *leben* "to live"

present	sg	pl
1	leb-ə	leb-ən
2	lep-st	lep-t
3	lep-t	leb-ən

past	sg	pl
1	lep-t-ə	lep-t-ən
2	lep-t-əst	lep-t-ət
3	lep-t-ə	lep-t-ən

(6) German vowel alternations in strong verbs: *geben* "to give"

present	sg	pl
1	geb-ə	geb-ən
2	gɪp-st	gep-t
3	gɪp-t	geb-ən

past	sg	pl
1	gap	gab-ən
2	gap-st	gap-t
3	gap	gab-ən

(7) German suppletive alternations: *denken* "to think"

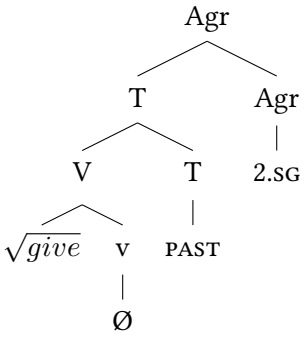
present	sg	pl
1	dɛŋk-ə	dɛŋk-ən
2	dɛŋk-st	dɛŋk-t
3	dɛŋk-t	dɛŋk-ən

past	sg	pl
1	dax-t-ə	dax-t-ən
2	dax-t-əst	dax-t-ət
3	dax-t-ə	dax-t-ən

Vocabulary Insertion rules take morphosyntactic trees such as (8) as their input (Halle and Marantz 1993, Bobaljik 2000), and apply rules to each node from the most embedded node outward[3]: the root is inserted first, then the categorizing head *v* (null in this case), then tense, and finally, agreement. As shown in (9), some of the morphosyntactic entries are associated with multiple vocabulary items. For example, the root for "think" is [dax] when [+past] is in the context, but is [dɛŋk] otherwise. For the root "give", [geb] is inserted, its [+STRONG] diacritic conditions the more specific Ø allomorph of [+past]. A further readjustment rule will change the height of [e] in present tense 2nd and 3rd person singular parts of the paradigm, and the vowel will change to [a] in the past. Once /gib-t/ has been thus assembled in the morphological component, it is submitted to the phonology, where the voicing of [b] is changed to [p]. The relationship between the allomorphs is thus mediated in three different parts of the system: lexical storage/VI rules, readjustment rules, and phonological rules.

[3]VI from the most embedded node outward has been formalized in several ways—see, for example, Wolf (2015), Myler (2017). Myler formulates the Temporal order of Vocabulary Insertion as follows: "For a pair of terminal nodes x and y: If x is the head of a maximal projection M such that M is categorially distinct from y and M dominates y, then y > x (">" means *is more deeply embedded than*). If y > x, then y undergoes Vocabulary Insertion prior to x." (Myler 2017, p. 102, modified slightly).

(8) Tree for past tense of "give" in German (Bobaljik 2000), after insertion of the root and agreement operations



(9) Vocabulary insertion rules for German

Roots	*Tense*	*Agreement*
$\sqrt{\text{GIVE}} \leftrightarrow \text{geb}_{[+\text{STRONG}]}$	[+PAST] ↔Ø/[+STRONG] ___	[2SG] ↔st
$\sqrt{\text{THINK}} \leftrightarrow$ dax / ___[+PAST]	[+PAST] ↔te	
$\sqrt{\text{THINK}} \leftrightarrow$ dɛŋk		
$\sqrt{\text{LIVE}} \leftrightarrow$ leb		

(10) Readjustment rules for "give" in German

a. $\begin{bmatrix} -\text{high} \\ +\text{STRONG} \end{bmatrix} \rightarrow$ [+high] / ___ . . . $\begin{bmatrix} -\text{PAST} \\ -1\text{PERS} \\ +\text{SG} \end{bmatrix}$

b. [−high] →[+low] / ___ [+PAST]

In the first readjustment rule in (10), the undergoer is the last full (non-schwa) vowel of the stem if the stem bears the [+STRONG] diacritic—the assumption that diacritic markers of the morpheme spread to all of its segments dates back to the SPE (Chomsky and Halle 1968; see Zonneveld 1978 and Gouskova 2012).

The rules in (9) refer to specific roots and morphosyntactic features as inputs, and their contexts are either morphosyntactic features themselves (e.g., [+PAST]) or diacritic features such as [+STRONG], which identify a particular conjugation. Conjugation/declension class features are halfway between syntactic and phonological: they can be determined by syntactic features such as gender (though see X-REF TO SYNCRETISM CHAPTER), but they also belong to specific vocabulary items and can condition phonological rules, as we just saw. Insertion rules can also refer to phonological features in their context: for example, the insertion rules for the indefinite determiner in English reference [+syllabic] for the "an" allomorph, and [-syllabic] (or unspecified/elsewhere) for the "a" allomorph. Finally, the readjustment rule in (10) takes as its input a phonological feature (here, [−high]) of any vocabulary item with the diacritic feature [+STRONG], and refers to morphosyntactic features in its context.

3.2 Conditioning environments for VI rules

DM does not in principle restrict what sorts of features can appear in the conditioning environment of a VI rule: they can be phonological ([sonorant], [syllabic], etc.), morpho-lexical (referring to declension/conjugation classes such as [+STRONG] in the German example in (9)), or morphosyntactic. An example of morphosyntactically conditioned allomorphy is supplied by Russian in (11). The Russian baby diminutive suffix[4] can be translated roughly as "-ling" (as in "duckling"), and it alternates between *-onok* in the singular and *-at* in the plural. The *-at* allomorph appears in any plural—thus, "kitten" has *-on(o)k* in all the singular case forms (*kot^{j}-onk-a* GEN.SG, *kot^{j}-onk-om* INST.SG, *kot^{j}-onk-e* OBL.SG, etc.), and *-at* in all the plural ones (*kot^{j}-at-Ø* GEN.PL, *kot^{j}-at-ami* INST.PL, *kot^{j}-at-ax* OBL.PL).

(11) Outward sensitivity to plural vs. singular in Russian baby diminutives

NOM.SG	NOM.PL	Gloss	BABY-DIM NOM.SG	BABY-DIM NOM.PL	Gloss
kot	kot-ɨ	'cat'	kot^{j}-onok	kot^{j}-at-a	'kitten'
slon	slon-ɨ	'elephant'	slon^{j}-onok	slon^{j}-at-a	'baby elephant'
ʂakal	ʂakal-ɨ	'jackal'	ʂakal^{j}-onok	ʂakal^{j}-at-a	'baby jackal'
pauk	pauk-i	'spider'	pauʧ-onok	pauʧ-at-a	'baby spider'

A simple Distributed Morphology analysis would posit a rule that uses plural as the insertion context, as in (12). The singular *-onok* would be inserted by an analogous rule in the context of the singular, or alternatively one of the affixes could be inserted by a context-free default rule (see §5.3).

(12) DIM ↔ -at / ___ PLURAL

Various parts of this system could be questioned, and they raise questions that any theory of allomorphy must address. Why are some allomorphs stored in the lexicon, while others are created grammatically? If a morpheme has multiple stored allomorphs, are the insertion rules for them ordered in an arbitrary fashion, or is there another principle for selecting between them? Are there diagnostics that unambiguously distinguish morphosyntactic from phonological conditioning? Finally, what are the limits on the types of possible interactions between morphemes, and the contexts that the various rules can reference and change? We entertain these questions throughout the following discussion.

4 Vocabulary insertion, allomorphy, and structure

4.1 Inward and outward sensitivity

One of the most exciting facets of the study of allomorphy lies in its potential to shed light on a fundamental divide among approaches to morphology, namely the question of whether the morphosyntactic features that are are expressed in a complex word have an internal hierarchical arrangement independently of the forms that realize those features. The discussion of vocabulary insertion above instantiates this: the tree in

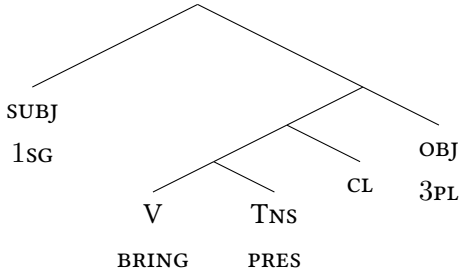
[4] This suffix is different from the diminutive *-ok* from (3), both in meaning and in morphosyntactic properties; for more, see Steriopolo (2008), Gouskova et al. (2015).

(8) represents a hierarchical arrangement of tense and agreement features within the German verb prior to (and thus independent of) the application of vocabulary insertion which provides phonological exponents (including Ø) to those abstract morphemes. Under the slogan 'Syntax all the way down', the existence of such structure is a central tenet of DM and related frameworks, but is denied, for example (at least for inflection) by Word-and-Paradigm and related approaches (Anderson 1992, Stump 2001, Blevins 2016).

With this difference in mind, one can ask whether there are conditions on allomorphy that are crucially stated in terms of such hierarchical structure. An affirmative answer would provide compelling support for the family of theories that incorporate such structure, with the potential that more fine-grained investigation of the types of constraint that are found will serve to distinguish among competing frameworks and theories.

One aspect of this investigation revolves around the distinction between *inwards*- and *outwards*-sensitivity (terms due to Carstairs 1987), namely, whether the triggering property is more or less peripheral to the word's root than the morpheme undergoing allomorphy. Within DM, Bobaljik (2000) argues that the hypothesis of cyclic vocabulary insertion from the root out (mentioned in §1.2), coupled with the proposal in Halle (1990) that vocabulary insertion rules rewrite the morphosyntactic features that they realize, yields a narrow set of constraints about the types of features that may trigger inwards or outwards sensitivity. Bobaljik's core argument comes from agreement morphology in the Chukotko-Kamchatkan languages, which show extensive affixal allomorphy. Bobaljik argues that the structure of the Chukotko-Kamchatkan verb can be represented as in (13).[5]

(13)



Relative to this structure, which has some independent motivation on syntactic and phonological grounds, all affixal allomorphy conditioned by morphosyntactic features (person, number, mood) is outwards-sensitive, while all affixal allomorphy conditioned by morphophonological features (in particular, root-specific conjugation class) is inwards-sensitive. This asymmetry in the direction of conditioning can be seen as an effect of the cyclic, bottom-up (root-outwards) application of VI, converting a morphosyntactic representation to a morphophonological one: When VI applies to any given node, more peripheral nodes are still abstract, consisting only of morphosyntactic information, thus only that information is available outwards.[6] Conversely, since lower nodes have already been converted to a phonological string, only information in the phonological representation, including inflection class diacritics, is visible for inwards-sensitive conditioning.

[5] We omit possible mood and aspect nodes, here both null. The object suffix is probably internally complex, with the glottal stop being a plural infix, not indicated here.

[6] This restriction was termed 'No Lookahead' in Simpson and Withgott (1986).

Some Itelmen verbs providing a partial illustration of these effects are given in (14). In (15), we sketch a stepwise derivation of one of these verbs, showing what information is available to condition allomorphy at each step. [7]

(14) Outwardly sensitive allomorphy in Itelmen

a. [ntxɬxkum]
/n-tɸɬ-xk-um/
IMPERS-bring-CL.II-1SG.OBJ
'(Someone) brought me.'

b. [ntxɬxkin]
/n-tɸɬ-xk-in/
IMPERS-bring-CL.II-2SG.OBJ
'(Someone) brought you.'

c. [tɸskiče?n]
/t-tɸɬ-z-ki-če?n/
1.SG-bring-PRES-CL.II-1>3.PL
'I am bringing them.'

d. [ttxɬaɬkičen]
/t-tɸɬ-aɬ-ki-čen/
1.SG-bring-FUT-CL.II-1>3.SG
'I will bring it.'

e. [tɸsčŋin]
/Ø-tɸɬ-z-čŋ-in/
2.SG-bring-PRES-CL.II-2>3.SG
'You are bringing it.'

f. [taβolaɬqzusčiŋnen]
/Ø-taβol-aɬ-qzu-z-čiŋ-nen/
3.SG-embrace-DESID-ASP-PRES-CL.II-3>3.SG
'He is always wanting to embrace her'
(Volodin 1976:76)

These examples show some of the complex affixal allomorphy discussed above. To a first approximation, the prefixes in these languages are markers of the subject (agreement). As the minimal pairs in (14a) vs. (b) and (c) vs. (d) show, the suffixes (for transitive verbs) mark the object. Overlaying this is a complex pattern of allomorphy. The pattern is asymmetric: the object suffixes are conditioned by the person and number of the subject, while the prefixes show no sensitivity to features of the object. Thus, the 3rd person sg. object suffix is either *-čen*, *-in, or -nen,* depending (in part) on whether the subject is 1st , 2nd, or 3rd person (glossed as 1>3.SG, 2>3.SG, and 3>3.SG). A third morpheme, the inflection class marker, shows allomorphy (here *xk*∼*-ki-*∼*č(i)ŋ*) conditioned by both subject and object features as well as the inflection class of the verb stem (for additional allomorphs of the affixes, see Bobaljik 2000, Bobaljik and Wurmbrand 2001.)

The stepwise derivation is given in (15), with some relevant VI rules in (16). At each step, VI replaces morphosyntactic information (gray) with morphophonological information (black). Each line represents a stage of the derivation and shows what information is and is not available to serve as a context for insertion at that step.

[7] The surface forms are derived from the underlying representations by various regular (morpho-)phonological processes, including cluster simplification, fusion of /ɬ+z/ to [s] and infixation of the glottal stop. Descriptions of Itelmen vary between transcribing the root-medial consonant as *x*∼*x*^{w}∼ *ɸ*. We chose /ɸ/ to simplify the presentation.

(15) Cyclic vocabulary insertion in Itelmen, deriving /Ø-tɸɬ-z-čŋ-in/ [tɸsčŋin] 'you are bringing it' from (14e)

a. [S:2SG [[[[BRING] PRES] CL] O:3SG]]

b. [S:2SG [[[[tɸɬ]$_2$ PRES] CL] O:3SG]]

c. [S:2SG [[[[tɸɬ]$_2$ z] CL] O:3SG]]

d. [S:2SG [[[[tɸɬ]$_2$ z] čŋ] O:3SG]]

e. [S:2SG [[[[tɸɬ]$_2$ z] čŋ] in]]

f. [Ø [[[[tɸɬ]$_2$ z] čŋ] in]]

(16) Some Itelmen VIs

Subject	Root	Infl. Class	Object
S:1SG ⇔ t-	$\sqrt{\text{BRING}}$⇔tɸɬ$_2$	CL ⇔ / -čŋ]$_2$ ___] O:3, S:2SG, REALIS	O:3SG ⇔ -in / ___] S:2SG, REALIS
S:2 ⇔ Ø	$\sqrt{\text{EMBRACE}}$⇔taβol$_2$	CL ⇔ -xk /]$_2$ ___] O:[+PARTICIPANT]	O:3SG ⇔ -nen / ___]S:3
		CL ⇔ -k(i) /]$_2$ ___	O:3SG ⇔ -čen / ___]

The hypothesis of cyclic vocabulary insertion explains asymmetries in allomorphy: the class marker is closer to the verb root than both subject and object agreement affixes, and shows outwards-sensitive allomorphy conditioned by both of these. The object agreement marker is inserted next, and its form is conditioned by features of the more peripheral subject agreement. Last to be inserted is the subject agreement marker—since it is the most peripheral relevant morpheme, it may not show allomorphy for the morphosyntactic features of the less peripheral object marker, since those have already been replaced by phonological material (see (15f)).

Whether such a result will generalize to yield a restrictive theory of allomorphy, crucially incorporating structural conditions on sensitivity, remains an active research question (see Bonet and Harbour, 2012 and some of the contributions in Gribanova and Shih (2017); there are also alternative proposals for asymmetries like those in Itelmen, which treat the asymmetry as reflecting the workings of agreement rather than exponence—see Béjar, 2003).

4.2 Locality

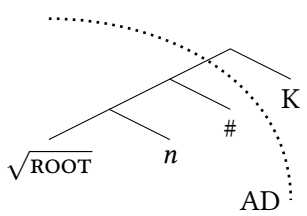
Related to the distinction in directionality of sensitivity is the issue of **locality**. The question is the same: are there structurally-defined locality conditions that restrict the possibilities for the features of one (abstract) morpheme to effect allomorphy of another? Many allomorphic alternations happen under linear and/or structural adjacency between target and trigger, and some version of an adjacency condition has long been entertained as a condition on allomorphy and other morphological interactions (Siegel, 1978, Allen, 1978, Embick, 2010). For example, if one assumes that the present tense in German is not merely phonologically null but rather radically absent, then the restriction of the readjustment rule in (10)) to the

present tense could be an effect of structural adjacency between agreement and the verb stem holding only in the present tense—in the past a null morpheme intervenes.

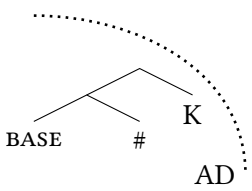
But not all instances of allomorphic alternations, not even all cases of suppletive alternations, happen under adjacency (see Moskal and Smith 2016, Kastner and Moskal 2018, Ostrove 2018). The Itelmen class marker in (14) is conditioned to occur with a limited set of verb stems, but is not always adjacent to the stem (the present tense marker intervenes in (14), as do aspectual markers). The recognition that adjacency as such appears to be too strict a condition on allomorphy, even on suppletive allomorphy, has led to a number of competing proposals for alternative characterizations of the relevant locality considerations. Work on suppletion in this vein has suggested a prominent role for structural conditions in understanding possible allomorphic patterns, in a manner consistent with DM (Bobaljik, 2012, Moskal, 2015b, Zompì, 2017, Smith et al., 2018).

Some of this work investigates whether there are structurally-defined absolute locality domains, reminiscent of phases in syntax. Moskal (2015a) defends this perspective in a proposal to understand why pronouns commonly show suppletive allomorphy for case and number, while nouns typically undergo suppletion, if at all, for number, but not (or very rarely) for case (Bybee, 1985). Moskal proposes to understand this as a reflex of the proposal that content words such as nouns include a root and a category-defining node (in this case *n*, as in (17)), whereas pronouns, as functional elements, lack the category-defining node (see (18)). On the assumption (see Embick 2010:§2.2.2 and our §5.1) that category nodes introduce domain boundaries (and with some ancillary assumptions), Moskal argues that the domain of features that may trigger suppletive root allomorphy—AD (accessibility domain) in the trees below—extends higher in pronouns (encompassing K=case), but is limited by the *n* node to a point that excludes K for lexical nouns. Specifically, Moskal argues that the AD extends one node beyond the categorizing node, when one is present. The presence of *n* in nouns allows K to be in the same AD as the base.[8]

(17)



(18)



As the relative recency of the works cited in this section indicates, the investigation of locality and other structural conditions on allomorphy is a live and rapidly evolving research area. The goal, of course, is a set of conditions that are flexible enough to accommodate observed examples, yet restrictive enough to

[8]Moskal assumes domains in morphology, as in syntax, are contextually determined—thus merger of # and *n* must happen before it is determined if n (a potential domain head) is actually a domain head. From a bottom-up structure-building perspective, #, but not K, is thus part of the representation at the point that rules such as vocabulary insertion apply to the root in (17).

provide an understanding of why certain patterns of allomorphy appear to be universally unattested, over large samples. As mentioned above, this area is exciting, in that understanding whether or not there are such structural conditions on allomorphy offers the potential to shed light not merely on the analysis of various phenomena, but on foundational or architectural points of current debate among approaches to morphology. We return to the relationship between phonological and morphosyntactic locality in §5.1 and §6.1.2.

5 Approaches to Suppletion in DM

At this point, we focus more narrowly on Vocabulary Insertion—the source of suppletive allomorphy, both of roots/stems and affixes. Formalizing the process can be done in various ways, and competing proposals within the general DM architecture lead to somewhat different predictions, at least in principle, about the nature of allomorphy. We review some of the major current proposals here.

In the original conception of DM, vocabulary insertion applies only at terminal nodes, i.e., the leaves of the tree assembled by the syntax. The derivations presented above are consistent with that model. For example, the German past tense *gab-Ø-en* has at least three morphemes structurally: ROOT-TENSE-AGREEMENT, with rules that apply at each node, including a rule of zero exponence of the past tense. While this accurately describes the pattern, under this presentation, there is no inherent reason why the special zero allomorph of past should co-occur with the special root allomorph—whether listed or derived by readjustment rule. Synchronically, these are treated as two independent facts, despite a rather widespread intuition that these facts should be related. Indeed, this criticism of the use of a zero allomorph, rather than having the vowel change itself spell out past, was made as early as Hockett (1954). On the other hand, forms such as *dach-te-n* support the claim that root allomorphy is independent of the exponent of the past tense, since the past tense receives its regular, default exponent in this context. As various authors have noted (Corbett 2007, Wurzel 1985), this redundancy is quite typical of root allomorphy: even though the root allomorph unambiguously signals that it is occurring in the context of some feature X (here past), the regular exponent of X frequently cooccurs with the root allomorph. We return to this issue in §6.2.1–6.2.3, where we consider alternatives to insertion at terminal nodes within DM and related frameworks.

5.1 Insertion into trees, or linearization?

There is an idealization in the claim that the input to Vocabluary Insertion is the syntactic tree: early DM authors proposed various mechanisms that alter the hierarchical structure generated in the syntax prior to vocabulary insertion, including fusion and fission (see Halle and Marantz 1993). But there are also proposals that include a different mapping mechanism between trees and linear strings. Thus, Embick (2010) proposes an 'concatenate' operation that maps the terminal nodes X and Y in the hierarchical structure to a linear string X⌢Y. This linear representation can undergo *pruning*, which removes any morphemes whose exponents are null and allows adjacent non-null morphemes to interact. It is then possible for morphemes to condition the realization of other morphemes even when they are not structurally adjacent to them in the tree—provided the intervening nodes do not have overt exponents, and provided the morphemes are

spelled out in the same cycle. Thus, the root of a verb can condition the exponence of tense and agreement in English and German (recall (6)), even though they are structurally remote. It is important that the verbalizing head *v* is null in these cases, and that all of these morphemes are linearized and spelled out in the same cycle.

(19) Linearized string for past tense of [dax-t-ə] 'thought' (following Embick 2010)

$\sqrt{\text{THINK}} \frown \text{v}, \text{v} \frown \text{T}^\circ, \text{T}^\circ \frown \text{Agr}^\circ$

Embick's model makes various interesting predictions, e.g., that all verbs with an overt *v* (e.g., *-ize* or *-ify* in English) should not display allomorphy of tense based on what the root is. Kastner (2019) shows that the prediction holds for allomorphy of vowels (which express voice distinctions) in templatic verbs in Hebrew. But Embick's model, read strictly, also has the consequence that the representation over which vocabulary insertion is defined has linear but not hierarchical structure in any given cycle—an aspect of the model we do not consider here. Note that Embick defends cyclic outwards vocabulary insertion, suggesting the representation is indeed hierarchical, as in canonical DM.

5.2 Locality conditions on morpho-phonological rules

Another open question in the study of allomorphy is whether morpho-phonological (readjustment) rules are subject to the same locality conditions as suppletion, or whether they are more akin to purely phonological rules, which can sometimes be nonlocal. If there is a systematic distinction between these types of allomorphy as regards locality, then this supports a theoretical architecture, such as DM, which draws the distinction. This issue has received less attention than the general issue of locality. In part, this may be because there is a difficult line to draw in distinguishing suppletion from readjustment: is the unproductive vowel change in *give/gave* suppletive or the result of an unproductive morphophonological rule? What about the class of *-ought* pasts (*think, buy, catch, ...* which retain their onset, but replace their rime with *[at]* —is that rule-governed? (Recall that we treated the German *denken/dachte* alternation as suppletive in §3.1, but this is by no means the only possible analysis. We return to the general issue of how and whether to state alternations as rules in §6.1.1).

The vast majority of phonotactically conditioned rules are conditioned locally, usually by adjacent segments or prosodic positions. There are non-local alternations such as vowel harmony and consonant dissimilation, but they appear to be limited in such a way that it is possible to analyze them as local at an abstract level of representation (McCarthy 1988) or because of a special relationship that skips over intervening segments (Rose and Walker 2004). Even for non-local phonological interactions, there are string-based locality generalizations: interaction at longer distances implies interaction across a single intervening segment (Suzuki 1998, Gouskova and Gallagher to appear), and phonological interactions rarely exceed the domain of the phonological word.[9]

[9] Stress can be assigned on the basis of units larger than the word, and vowel harmony (especially ATR harmony) is occasionally reported in larger domains—see, e.g., Hall and Hall 1980. We are setting aside speech errors such as spoonerisms, where it is not uncommon for phonological switches to involve remote positions within an utterance.

These observations apply to phonological interactions, but morpho-phonological rules do not always obey them. For example, in Celtic mutation, the trigger does not have to be adjacent to the undergoer (Green 2006, Hannahs and Tallerman 2006, Wolf 2007). In Zulu (see (20)), the labial-palatal alternation in the root is conditioned by the passive suffix even across an intervening causative morpheme (see Herbert 1990, Embick 2010):

(20) Zulu passives: labial palatalization conditioned on non-adjacent morphemes

ACTIVE	PASSIVE	CAUS. ACTIVE	CAUS. PASSIVE	
bamb-a	banj-w-a	bamb-is-a	banj-is-w-a	'catch'
boph-a	bosh-w-a	boph-is-a	bosh-is-w-a	'tie'

It is likely not accidental that non-locally conditioned alternations of this sort involve phonological changes but do not look like normal phonological rules. Celtic mutation is difficult to characterize in unified phonological terms (Gnanadesikan 1997, Green 2006, Wolf 2007). Palatalization of labials in Zulu is typologically unusual—it is far more normal for coronals and velars to palatalize (see Herbert 1990 on the history of the rule). Wolf (2007) suggests that Celtic mutation is really a reflex of an agreement morpheme with several floating feature allomorphs, and that the morpheme is positioned in the morphosyntax rather than phonologically. We return to Zulu in §6.1.2, where we suggest that certain cases of non-local readjustment rules are probably better analyzed as floating feature morphemes, whose distribution is governed by phonology.

5.3 Competition for insertion: subset principle vs. extrinsic ordering

According to Halle and Marantz (1993), the stored allomorphs of a morpheme compete for insertion into a node. An allomorph is selected for insertion according to the Elsewhere Principle: the best candidate for insertion matches the feature values in the node, and it expresses the most features available. We can illustrate this with an example from Hindi (see (21)). Hindi nouns show *diagonal* syncretism: the NOM.SG and OBL.PL suffixes are *-aa* and *-õ*, with NOM.PL and OBL.SG being realized as *-e*. Adjective paradigms have L-shaped syncretism, with only one cell distinct (the NOM.SG, *-aa*), and the rest realized as *-e* (see (21)):

(21) Hindi masculine class II noun stems and masculine adjectives (Arsenault 2007)

'boy (noun)'	NOM	OBL	'bad (adjective)'	NOM	OBL
SG	laḍk-aa	laḍk-e		baḍ-aa	baḍ-e
PL	laḍk-e	laḍk-õ		baḍ-e	baḍ-e

An analysis is sketched in (22): there are specific insertion rules for the NOM SG and the nominal OBL PL morphemes,[10] and *-e* is not specified for either case or number features. Since *-aa* and *-õ* express a superset of features compared to *-e*, they win the insertion competition.

[10] In order for this analysis to capture the difference between nominal and adjectival oblique plurals, the rule would have additional information in its application context, which we abstract away from.

(22) Vocabulary insertion rules for Hindi

MASC.NOM.SG ↔ -aa

MASC.OBL.PL ↔ -õ / ___

MASC ↔ -e

But the Elsewhere Principle does not help in the event of a tie. In a precursor to the Distributed Morphology framework, Halle (1990) proposes that some VI rules are extrinsically ordered.[11] Similarly, Halle and Marantz (1993) use extrinsic ordering in their analysis of Georgian, where the Elsewhere Principle fails to provide one. They do, however, leave other possibilities open, and various options have been explored in subsequent DM work.

For a sense of the issues involved and how they interact with other analytical choice points, consider the Russian example in (23). In Russian third person pronouns, gender is marked with three different suffixes in the singular, but the plurals of all three genders just have *-i*. (These same suffixes are used in participles, short-form (predicate) adjectives, and past tense verbs.) The obvious generalization here is that gender is not distinguished in the plural—a situation that is common but by no means universal.

(23) Russian Nominative suffixes

3RD PERS PRON	SG	PL
FEM	on-a	on-i
MASC	on-Ø	on-i
NEUT	on-o	on-i

There are many ways one could analyze this syncretism, but two analyses are sketched in (24).

(24) Two Analyses of Russian Nominative suffixes

	Analysis I		Analysis II
a.	[+PLURAL] ↔ -i	a'.	PLURAL ↔ -i
b.	FEM.[-PL] ↔ -a	b'.	FEM ↔ -a
c.	MASC.[-PL] ↔ -Ø	c'.	MASC ↔ -Ø
d.	NEUT.[-PL] ↔ -o	d'.	[ ] ↔ -o

These analyses differ in the features assigned to the various vocabulary items that occur in the singular. In analysis I (following Halle 1990, with some simplifications re: gender/declension class interaction), each of the suffixes is positively specified for either [-PL] or [+PL]). In Analysis I, there is no crucial ordering among the exponents—no context can ever be simultaneously singular and plural, or masculine and feminine, and thus the unique correct exponent will be chosen in all contexts, with no reference to order. In Analysis II, each exponent is associated with (at most) a single feature. In this analysis, the order of

[11] An echo of this approach is to enforce realization with specific morphs via arbitrarily ranked constraints in an OT-like framework; see, e.g., Pertsova (2004) or MacBride (2004).

the rules is important. Rule (24d') is effectively last by the Elsewhere Condition.[12] But under Analysis II, the contexts of [PLURAL,FEMININE] constitutes a tie: the Elsewhere Condition does not resolve the choice between rule (24a') and rule (24b'), or likewise between rules (24a') and (24c') in the [PLURAL, MASCULINE] context. For examples of this general sort (though not for this particular case), Halle suggested the rules are crucially ordered as in (24a–d): gender is not distinguished in the plural in Russian because it is a property of Russian that the exponent of plural (24a) is more highly ranked than the exponent of feminine. In related languages such as Serbian, gender is distinguished in the plural, so the rules would specify gender in the plural and in the singular.

Since extrinsic ordering is by nature stipulative, there have been many attempts to derive certain aspects of ordering. For example, Noyer (1997) proposes a universal hierarchy of features: 1>2>PLURAL>DUAL>FEM. In the proposal, the higher the feature is in the hierarchy, the more priority it gets in insertion. If such a hierarchy is universal, and in particular the expression of NUMBER takes priority over that of GENDER, then the Russian order in Analysis II respects this universal hierarchy and is no longer encoded as a parochial (language-specific) stipulation. Relatedly, Analysis II requires no extrinsic ordering (i.e., no ordering other than the Elsewhere Condition) if cross-linguistic generalizations, such as having fewer genders in the plural than in the singular, are expressed via Impoverishment operations (X REF TO IMPROVERISHMENT CHAPTER): if number and gender simply do not cooccur on individual terminal nodes in Russian, then there are no ties to resolve, and order among the vocabulary items in Analysis II is not needed.

Yet another family of analyses derives the ordering of VI rules from feature geometries (Arsenault 2007, Bonet 1991, Harley 1994, Harley and Ritter 2002). Both universal hierarchies and universal feature geometries are designed to capture implicational universals, so the true test for these theories is to look for typological generalizations about morphemes that act as defaults.

Other cases are more complex. In both the Russian and the Hindi examples just considered, we see that analytical choices regarding the features interact in non-trivial ways with the question of rule ordering. In Analysis I in (24), no ordering is needed, but reference is made to both [+PL] and [-PL]. This analysis is not compatible with an approach in which [PLURAL] is a privative feature, and where rules such as vocabulary insertion are unable to make reference to the absence of a feature. Similarly in Hindi, the simple characterization of the L-shaped pattern of syncretism in (21) seems easy enough to characterize as in (22), but here too, the analysis treats associated the most marked (i.e., most highly specific context) vocabulary item with what might on other grounds be considered the least marked feature values (masculine, singular, and nominative), all of which are, in some theories, the absence of feature specifications.[13]

Luckily, it is not our place to adjudicate among the many possible analyses here. As the discussion

[12] We assume for argument's sake that neuter is the default, which is a debatable point but for which there are some arguments in Russian—e.g., inanimate indeclinables are neuter, as are abstract adjective-derived nouns (*neznakomoje* 'the unknown'). Dative subject verbs have neuter agreement, *-o*, rather than agree with the subject. Other arguments suggest that the masculine is the default for animates, however (e.g., verbs in "who" questions are masculine, vs. "what", neuter), so the status of default gender is controversial.

[13] Christopoulos and Zompì (2019) observe that patterns like (21) are widespread. From this, they argue that one cannot maintain both that (a) 'unmarked' categories like nominative case signify the absence of the relevant feature, and (b) that rules may not refer to the absence of features (unless one allows for disjunctive contexts in rules). They argue for a representation of markedness that preserves the insights associated with 'containment' representations such as Caha (2009a) but allows reference to the 'unmarked' value in order to describe L-shaped paradigms. See Smith et al. (2018) for a related proposal for number.

has shown, while some DM proposals stipulate rule ordering in Vocabulary Insertion, there are typically alternative analyses, which make no recourse to these stipulations. Thus, the issue cannot be decided independent of other assumptions—for example, about the nature of features and representations.

5.4 Do root allomorphs compete?

VI is the mechanism in DM that formally encodes suppletive allomorphy, whether of roots or affixes. Our discussion would then not be complete without touching on a further point of debate within DM, namely whether lexical roots show suppletion at all. Bobaljik (2012) uses VI rules such as (25) to characterize suppletion in adjectival gradation.

(25) English comparatives analyzed as root suppletion

a. $\sqrt{\text{GOOD}} \rightarrow$ bɛ(t)- /___ CMPR

b. $\sqrt{\text{GOOD}} \rightarrow$ gʊd

These two rules provide exponents for the abstract root $\sqrt{\text{GOOD}}$ and their competition is regulated by the Elsewhere Principle. Despite the obvious ease with which root suppletion can be characterized in this way, one line of thought within DM, notably as expressed in Marantz (1996), holds that such suppletive alternations may characterize elements of the grammatical vocabulary (function words), but that open-class, lexical roots do not undergo suppletion in this way. The empirical issue is that in many languages, apparent cases of root suppletion (like *good∼better, go∼went, person∼people* in English) are both rare and amenable to reanalysis. There are at least three possibilities consistent with Marantz's proposal: (i) the forms do not constitute suppletive pairs, but are distinct lexemes with substantial overlap in meaning—note that apparent suppletive forms sometimes exist in doublets with a regular form, like *persons* (recall §2). (ii) the alternation is derived by powerful readjustment rules, or (iii) the items in question are part of the functional vocabulary, not the lexical vocabulary—e.g., perhaps *go* is essentially a light verb.

The claim that roots do not supplete has subsequently been challenged (notably in Harley, 2014), and is no longer (if it ever was) widely held within the DM literature, but we provide here a brief review of the relevant issues [X-REF TO HARLEY CHAPTER]. The central issue is whether roots are individuated in the syntax, and if so, then how. Recall from the introduction that one part of the argument for Separation was that differences in the phonological realization of a given morpheme generally play no role in the syntax (cf. the Principle of Phonology-Free Syntax, Zwicky and Pullum, 1986). This state of affairs is explained if phonological information is inserted 'Late', i.e., post-syntactically—syntactic representations simply do not contain phonological information, so no syntactic operation can make reference to such information. Marantz notes that analogous considerations hold of some dimensions of lexical semantics. Distinctions such as mass versus count and animate versus inanimate are known to play a role in (morpho)-syntax in various languages, and thus need to be a part of the syntactic representation, but within the (sub)-categories defined by such features, finer-grained semantic differences that distinguish individual roots are not syntactically relevant: the syntax does not care about the meaning difference between, say, *cats* and *bats*, any more than it cares about the phonological difference in the first consonant. Thus a conceptually

'clean' approach might hold, as Marantz did, that meaning differences that are not characterizable in terms of grammatical features are also not a part of the syntactic representation, and are also inserted 'late'. If roots sharing the same grammatical features are not differentiated from one another, then rules such as (25) are unstatable.

This view is challenged by the many apparent cases of suppletion of elements that are poor candidates for being part of the functional vocabulary. A particularly striking example given by Corbett (2007) is the Archi pair (from Kibrik 1977) meaning 'corner of a bag', singular: *bič'ní*, plural: *božd*ó.[14] Harley (2014) also develops arguments that lexical roots which alternate for participant number in Hiaki are truly examples of lexical suppletion. Working backwards, if it is accepted that lexical roots do undergo suppletion, and thus that rules like (25) are to be countenanced, then challenging questions arise concerning the representation of roots in the pre-insertion lexicon. Root suppletion shows that the phonological matrix cannot be the key to identifying a root, and Harley (2014) argues that many roots do not have a consistent, stable, identifiable semantics (i.e., roots may show *allosemy* just as they show *allomorphy*). Pursuing these issues in any depth would take us far afield from the topic of this contribution, but we note here that what seems like a simple, empirical question: Is there suppletion of lexical roots? turns out to have deeper architectural consequences that must be grappled with.

6 Refinements and Alternatives

6.1 Phonological issues and alternatives

6.1.1 Readjustment rules or morphosyntactically conditioned suppletion?

Readjustment rules are designed to capture regularities that are distinct from phonological and suppletive allomorphy. Consider Halle's (1990) analysis of the English plural, sketched in (2). There are some nouns that take a zero affix and show no stem changes (*deer*). Others, suffixed with *-us*, lose that suffix and gain *-i*. A small set of nouns (*childr-en, ox-en*) have the *-en* suffix, sometimes with stem vowel changes. Finally, the default is *-z* with no vowel changes, which applies to most nouns. Halle's analysis separates vowel changes from the affixes, allowing for a cross-classification of nouns as shown in (26).

(26) Halle (1990): analysis of English plural allomorphy

	PL →Ø	PL→-i	PL →-en	PL →-z
no readj. rules	deer	syllab-us/syllab-i	ox-en	dog-s
V→[-bk,-lo]/___PL	foot/feet	—	childr-en	—

Readjustment rules have been controversial for several reasons. First, they mix phonological and morphosyntactic information, which some see as unrestrictive (Siddiqi 2009, Gribanova 2015). Another crit-

[14] Corbett raises this example against the common view that suppletion applies only to high frequency lexemes, rather than in terms of Marantz's lexical/functional contrast. Although this is frequently raised as an example of suppletion of a low-frequency lexical root, Moskal (2015b) notes the similar pair meaning 'pier of a bridge'–singular: *biqˀní*, plural: *boʁdó*, and suggests this is an instance of readjustment, rather than true suppletion.

icism is that unification of certain morphophonological changes requires a highly abstract analysis, with some of the rules having such a narrow scope as to apply to one item in the language. The trade-off between gaining a general explanation for *feet* and *children* undergoing the same rule is that *children* requires several additional rules of no generality at all (and even *foot/feet* and *goose/geese* type cases are not following the exact same rule; see §6.1.2). The alternative to readjustment rules would treat idiosyncratic changes as suppletion, despite the family resemblance between the allomorphs.

The problem of "applies to just one item" has to be seen as part of a larger question of when it is appropriate to posit a rule. It is not uncommon to characterize allomorphy as suppletion when only one or two morphemes fall into a subpattern; e.g., in English indefinite allomorphy there is a phonological resemblance between *a/an*, but no other morpheme has a [n]∼Ø alternation depending on the following consonant/vowel context. The [n]∼Ø alternation has a productivity of 1 in English, but because the indefinite is so frequent in usage, it is not in doubt that the alternation between *a/an* is quite productive.[15] The analysis of such cases is controversial—it is debated whether they are suppletive or derived by lexically idiosyncratic phonological rule. According to some, rules are supposed to capture aspects of speakers' knowledge that are generalizable; if a rule cannot be applied to new items, it is not a rule (Yang 2016). But experimental research on morpho-phonological productivity suggests that it is gradient, and even rules that do not appear to be productive in the wild can become so in an experiment (Albright and Hayes 2003, Ernestus and Baayen 2003, Becker and Gouskova 2016, Gouskova and Becker 2013). This kind of experimentation has to be coupled with explicit computational models of rule learning and generalization, and this too is an exciting area of current and future research.

6.1.2 Readjustment rules vs. Floating features

For a fair number of cases, there is an alternative to readjustment rules: floating features (McCarthy 1983, Akinlabi 1996, Zoll 1996, Wolf 2007). Floating feature morphemes are a natural consequence of the view that phonological representations are hierarchically structured (Goldsmith 1976, Selkirk 1978, McCarthy 1986); the idea is that some morphemes consist of unattached phonological features in addition to (or instead of) full segmental content. Just like templatic morphemes, which consist of syllables or feet (McCarthy and Prince 1986), these morphemes are slightly representationally defective. The floating feature account is especially intuitive when applied to allomorphy consisting of a feature change on the stem-final consonant, as in *belief* (n) vs. *believe (*v*)* or *hou[s]e* (n) vs. *hou[z]e* (v) (see (27a)), or the vowel changes in English plurals such as *f[ʊ]t*∼ *f[i]t* (see (27b)).

(27) Some examples of floating feature analyses for allomorphy in English

a. $v \leftrightarrow$ [+voice] / X_n]___ where X ∈ {bəlif, hæf, haus. . .}
b. PL $\leftrightarrow$ [+high,-back] / X X ∈ {fʊt, gus, . . .}

[15] Though of course some speakers have *an* before certain h-initial words such as *historical,* and there are some dialects that generalize *a* to the pre-vocalic context, too (see Gouskova et al. 2015, Pak 2016 for recent discussion).

Just as templatic morphemes can consist of a mix of templatic and segmental material, a floating morpheme can also consist of a mix of segments and floating features: for example, as the velar palatalization in the Russian diminutive in *volk/volʧ-onok* 'wolf' could be represented as /[+strident], -onok/.

In a sufficiently rich theory of phonology, floating features can account for a wide range of alternations (see especially Wolf 2007), and with some additional assumptions can even be extended to patterns such as subtractive morphology (Trommer and Zimmermann 2014, Zimmermann 2017). These analyses raise many interesting phonological questions, but their relevance to DM is that they introduce a different set of locality issues. The docking of floating phonological features in these proposals is conditioned by phonological rather than morphological domains. One of the best-known examples of floating features respecting phonological logic comes from Chaha (McCarthy 1983), whose various verbal distinctions are realized as labialization, palatalization, or both. Tonal morphemes, which are usually treated as floating elements, can also dock nonlocally—for example, Kawahara and Wolf (2010) describe a Japanese suffix that causes an accent to appear on the word-initial syllable. Wolf's (2007) phonological theory of floating feature docking includes a provision that might explain why a feature might dock nonlocally from its trigger in a case like Zulu labial palatalization: he proposes a constraint that requires docking to be non-vacuous. If the Zulu passive suffix *-w* comes with a coronal feature and is subject to the non-vacuous docking prohibition, then it would not be able to dock on the causative suffix *-is*, and would skip over to the nearest palatalizable labial. This analysis of Zulu predicts that palatalization should operate even at longer distances, and in the closely related Xhosa, it does indeed skip over as many as three coronal-containing morphemes (shown in (28); we follow source in using the orthography.[16]

(28) Xhosa labial palatalization, triggered by passive *w-*, across three morphemes (Buell 2005:38–39)

a. cham- a
√urinate FV
'to urinate'

b. chany- is- w- an- a
√urinate CAUS- PSV- RECIP- FV
'to cause/help each other to urinate (pass.)'

c. khuph- is- an- a
√*compete* CAUS- RECIP- FV
'compete'

d. khutsh- is- an- el- w- a
√*compete* CAUS- RECIP- APPL- PSV- FV
'be competed for'

Finally, although floating features may offer a certain intuitive appeal for some critics of readjustment rules, they are equally vulnerable to criticisms. One of the appealing aspects of the floating feature view is that it allows one to have a completely internally consistent item-and-arrangement[17] view of morphology: everything is an affix. Things can quickly get complicated, however. For example, in English verb formation, voicing is sometimes but not always accompanied by vowel changes (*bath* [bæθ] /*bathe* [beɪð], *life/live*), and it is not always possible to unite the vowel changes in plural formation into a neat

[16] This example bears on Embick's (2010) Readjustment Activity Hypothesis, whereby the undergoer of a readjustment rule must be in the same cycle as the trigger/context morpheme. It is likely that there is at least one spellout cycle intervening between the passive in (28) and the root—on Buell's analysis, there are two instances of little v in the structure. Note that Kastner (2019) analyzes the inability of the passive to condition vowels inside templates in Hebrew in cyclic terms.

[17] The terms *item-and-arrangement* and *item-and-process* come from Hockett (1954).

featural description (*man/men, mouse/mice* do not involve the same vowel changes as *foot/feet*). Some refinements are needed to capture the location of docking of the floating features in longer plurals with schwa in the second syllable (*e.g., woman/women* [wʊ̲mən/wɪ̲mən]; the vowel [ʊ] moreover does not tense as in *foot/feet*). Worse still, while a few roots show voicing alternations in both plurals and verb formation (*half/to halve/the halves, shelf/to shelve/shelves*), others do only in plurals (*leaf/leaves*, but *to leaf through*), and still others only in verb formation (*belief/to believe*, but *beliefs* in the plural) or in plurals and deverbal nominals (*thief/thieves/thievery/thieving*, but *?to thieve*). A full analysis of all these cases would require either massive suppletion of *v* and PL or an extensive role for abstraction in the phonological component. As is often the case, then, the appeal of the reductionist move to treat everything as an affix is diminished when encountering reality.

6.1.3 Is selection ever phonological?

It is well-known that it is often possible to predict the choice of a suppletive affix from the phonology of the stem (Siegel 1974, Carstairs 1988, and many others). One of the controversies in phonology has concerned the nature and extent of this predictability. On the one hand, some approaches (Kager 1996, Mascaró 2007, Smith 2015) argue that the choice between suppletive allomorphs is determined in the phonology, by regular markedness constraints. This explains why suppletive allomorphy sometimes mimics Jakobsonian syllable structure universals (e.g., Korean nominatives alternate between *-i* and *-ka* in a way that ensures perfect CV syllables, *mom-i* 'body NOM' vs. *kʰo-ka* 'nose NOM', Nevins 2011:2363).

On the other hand, typological surveys suggest that it is not possible to reduce all suppletive alternations to phonological markedness (Paster 2006, Wolf 2015, Nevins 2011). Some suppletion is predictable from phonological context but cannot be viewed as phonologically optimizing (e.g., English comparatives and superlatives favor mono- and disyllabic stems, but homophonous suffixes such as the *-er* in *malingerer* do not impose such restrictions). Recognizing that some listing is inevitable, Gouskova et al. (2015) propose that even apparently phonologically conditioned suppletion is really lexically conditioned. The phonological generalizations are extracted through phonotactic learning over lists associated with each VI rule. This approach can be extended to semantic generalizations about lexical classes, as well, such as those in Bantu languages (see Corbett 1991 for an overview). Since both phonotactic and semantic generalizations about allomorphy are often gradient and admit exceptions, a stochastic learner should be able to extract generalizations even in the face of counterexamples (see Hayes and Wilson 2008, Hayes et al. 2009, Gagliardi and Lidz 2014).

6.2 Morphosyntactic issues and alternatives

6.2.1 Portmanteaux and Non-terminal insertion

The English comparatives *good→bett-er* and *bad→worse* are both suppletive, and thus characterized in terms of competing vocabulary items, yet they differ in an interesting way. *Better* evidently shows suppletion of the root, but retention of the regular comparative suffix *-er*. Thus, it is easily described as the result of the VI rule in applying to the root node:

(29) GOOD → *bett-* / ___]CMPR

But *worse* requires something more, since the regular comparative suffix *-er* is missing. If VI applies uniquely to terminal nodes, and there is no restructuring of the output given by syntax, then portmanteau elements like *worse*, where a single overt exponent appears to correspond to two syntactic terminal nodes, requires a kind of mutual conditioning, wherein the adjectival root shows allomorphy conditioned by the comparative, and in turn the comparative shows allomorphy conditioned by the specific root (as in (30 a)-(30 b)).

(30) (a) BAD → *worse* / ___]CMPR

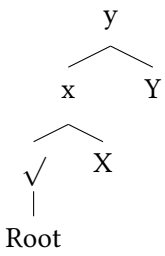
(b) CMPR → Ø/ *worse*]___]

Such mutual conditioning is consistent with the tenets of DM, and for the basic cases, empirically adequate. An alternative which has found various proponents is to relax the terminal condition on VI, and to allow for VI at non-terminal nodes (Neeleman and Szendrői 2007, Caha 2009b, Starke 2009, Radkevich 2010, Bobaljik 2012):

(31) [[BAD]CMPR] → *worse*

This rule would take care of both the root suppletion and the absence of the productive affix in one fell swoop. Permitting non-terminal insertion raises technical questions, requiring additional conditions to prevent overapplication, discussed in the literature cited. Whether there are empirical benefits depends on other assumptions. Arguments in favour of non-terminal insertion in DM include the following: Radkevich (2010) argues that in the domain of complex locative case morphology, there is a unique, universal hierarchical grouping of the elementary features of locative cases relative to which the distribution of attested portmanteau morphemes corresponds to those that constitute nodes in her structures, where sequences of features that are do not correspond to a constituent never occur as portmanteau morphemes. Allowing insertion to apply to any node, not just terminal nodes, thus derives the right cross-linguistic generalizations in this case, but these generalizations would appear to be accidental if mutually-conditioned allomorphy is unconstrained (or, Radkevich contends, if Spanning is permitted). Additionally, both Radkevich (2010) and Bobaljik (2012) argue from locality that in structures like (32), suppletion of the root conditioned by a feature at Y is possible only if ROOT+X form a portmanteau (like *worse*).

(32)



Latin provides an illustration where X is the comparative and Y the superlative *-imus*. Many superlatives contain the formative *-iss-* at position X, a historical reflex of the comparative, but the suppletive superlative *opt-imus* lacks the *-iss-* exponent. This is consistent with the view that contextual allomorphy is limited

to adjacent nodes. If VI rules are limited to terminal nodes, a rule like (34) would be inadmissable, as the root and trigger are not adjacent, whereas if non-terminal insertion is allowed, a portmanteau rule like (33) would satisfy the condition of adjacency with its trigger.

(33) [GOOD] → *opt* / ___]...SPRL

(34) [[GOOD]CMPR] → *opt* / ___]SPRL

Note that the argument for treating Latin in these terms rests on an adjacency (or contiguity) condition on root suppletion, an assumption that has recently faced an increasing array of challenges (Moskal and Smith, 2016, Smith et al., 2018, Kastner and Moskal, 2018).

6.2.2 Spanning

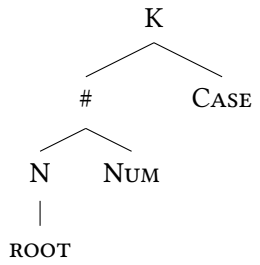
Radkevich and Bobaljik's proposals to admit non-terminal VI within DM retain the idea that VI rules apply to constituents, i.e., nodes in the complex X^0. That too has been challenged, and a strand of work in frameworks neighbouring DM proposes that vocabulary items may express a contiguous *span* of nodes, which need not form a constituent. The term is introduced in Williams (2003) and popularized within one version of Nanosyntax (Svenonius, 2012, Merchant, To appear). A portmanteau in inflection is a case of many: one mapping of features to exponents (i.e., cumulative exponence in the terminology of Matthews (1972)). In an inflectional language like Russian, nominal endings express case, number and gender (or more accurately, declension class), but are typically internally unsegmentable (see (35)).

(35) Russian portmanteaux suffixes

	stol 'table' (M)		kniga 'book' (F)	
	SG	PL	SG	PL
NOM	stol	stol-i	knig-a	knig-i
ACC	stol	stol-i	knig-u	knig-i
GEN	stol-a	stol-ov	knig-i	knig
DAT	stol-u	stol-am	knig-e	knig-am
INSTR	stol-om	stol-ami	knig-oj	knig-ami
PREP	stol-e	stol-ax	knig-e	knig-ax

It is a reasonable question whether the word-internal syntax of Russian involves only a simple structure such as: [[ROOT] INFL], in which the inflectional features are 'bundled' together on a single terminal node, or whether number and case morphemes are projected as separate heads, as in (36):

(36)



If there are arguments for the structure in (36) (which conforms to the same cross-linguistic generalization as seen transparently in Tamil, Moskal and Smith 2016), then the inflectional exponents in (35) are portmanteaux. Canonical DM, which restricts VI to terminal nodes, might then describe the Russian paradigm via mutual conditioning of exponents and liberal use of zeros.[18]. But proponents of spanning would contend that the affixes in (35) each represent a span of two nodes: number and case, even though these nodes do not form a constituent in (35).

6.2.3 Vocabulary Insertion in Nanosyntax

Within DM, non-terminal insertion, including spanning, is perhaps a minority view: the Vocabulary is generally held to be a list of atomic elements, each of which has no internal syntactic structure.[19] Non-terminal insertion is, however, a major component of treatments of VI in Nanosyntax, a framework that shares with DM a commitment to a sub-word syntactic arrangement of abstract morphemes as the input to VI. The intuition here trades on the partial inter-translatability of rules like (30 a) and (31): a feature that conditions a rule may occur either in the rule's structural description or as a part of an additional contextual condition. Thus, if DM is amended to allow insertion at non-terminals, then there is a concern about restrictiveness, in the sense that there are at least two possible analyses for words like *worse*: mutually-conditioned contextual allomorphy as in (30 a)–(30 b), or portmanteau insertion as in (31). Canonical DM would resolve this in favour of mutually-conditioned allomorphy, since (31) is unavailable.[20] A recent set of proposals within Nanosyntax, notably Caha et al. (2018) goes the other way, proposing to exclude contextual (root) allomorphy entirely, and treating all allomorphy as ostensibly context-free VI—leaving (31) as the only analysis.[21] An immediate question is raised by cases like *good/bett-er*, assuming the *-er* is synchronically segmentable as the comparative. For a solution to this, Caha et al. (2018) look to Czech, a language in which the comparative suffix has both a regular long form *-jejš-* (orthographically: *-ějš-*) and an irregular short form: *-š-* (shown in (37); the -í/-ý suffixes are concord markers):

[18] See Fenger (2018) on Agreement portmanteaux. As we have seen above, number, but not case, governs suppletion in Russian (of roots and of the baby-diminutive). A structural explanation of that asymmetry along the lines of Moskal (2015b) would seem to require a structure like (36), in turn requiring that a stand be taken on the derivation of portmanteau exponents.

[19] The treatment of idioms, which relate chunks of structure to non-compositional meaning, is given over to the Encyclopedia (Halle and Marantz 1993, X-REF ANOTHER CHAPTER IN BOOK)

[20] If *fusion* is countenanced, then canonical DM still has two possible analyses.

[21] Although they propose to exclude all contextual allomorphy, they invoke it in their treatment of some affixal allomorphy in the paper, thus the proposal remains to some extent programmatic.

(37) Czech: allomorphy of the comparative suffix

POS	CMPR	gloss
červen-ý	červen-ějš-í	'red'
hloup-ý	hloup-ějš-í	'stupid'
star-ý	star-š-í	'old'
bohat-ý	bohat-š-í	'rich'

Caha et al. (2018) observe that suppletive comparatives in Czech consistently take the short form affix only, thus *lep-š-í*, rather than **lep-ějš-í*:

(38) Czech comparatives: root suppletion

POS	CMPR	gloss
dobr-ý	lep-š-í	'good'
špatn-ý	hor-š-í	'bad'

Caha et al. (2018) reason as follows. Following a venerable tradition in the Slavicist/Indo-Europeanist literature, they assume that the longer comparative is bi-morphemic: *-jej-š-*, and they suggest that all comparatives abstractly have the bi-morphemic structure in (39):

(39) [[[[ROOT] CMPR1] CMPR2] INFL]

The comparative of 'red' in (37) shows individual exponents corresponding to each of the four morphemes. This expansion of the comparative allows them to treat the suppletive comparatives as portmanteaux, like English *worse*, as in (40 a).

(40) (a) [[GOOD] CMPR1] ⇔ lep-

(b) GOOD ⇔ lep- / ___] CMPR

This bleeds expression of the lower comparative morpheme, but allows expression of the higher part of the comparative, explaining why suppletive comparatives take the *-š-* suffix, but cannot combine with the longer affix, much in the same way that Latin, as discussed above, fails to express the comparative reflex *-iss-* in suppletive superlatives on the analysis in Bobaljik (2012). Programmatically, they resolve the theoretical indeterminacy by rejecting context-sensitive rules of VI (like (40 b) categorically, proposing that non-terminal insertion is the only means for capturing (suppletive) contextual allomorphy (at least of roots). Returning to English *better*, they must then assume that English, like Czech, has a bimorphemic comparative, and that *-er* is the outer comparative suffix. The inner one, needed in order to treat *bett-* as a portmanteau, is always null. In this framework, suppletion is a diagnostic for the presence of an additional node in the structure.

As with other proposals discussed above, the interesting question is whether the analysis of Czech generalizes. Treating all apparent contextual allomorphy (at least of roots) as portmanteau VI in this

manner derives in a strong manner a structural adjacency requirement between what would otherwise be described as the target and trigger of suppletion (allomorphy). Decomposing heads like the comparative (or PLURAL as would be needed for the Russian baby-diminutive facts considered in §3.2) into multiple heads, allows for the characterization of apparent examples of contextual allomorphy where the trigger remains overt, but the trigger should always be contiguous to the expressed element. In a strong form, their proposal thus derives a conjecture put forward in Bobaljik (2012) (with an antecedent in Radkevich (2010)):

(41) If an exponent X expresses $\sqrt{}$root + $F_1 \ldots F_n$, for some features $F_1 \ldots F_n$, then $F_1 \ldots F_n$ must be adjacent to (or contiguous with) $\sqrt{}$root.

Just as apparent non-local/non-adjacent allomorphy (e.g., Tamil and Kayardild, Moskal and Smith 2016) is problematic for theories that stipulate an adjacency requirement, such examples are all the more challenging for theories from which this follows from basic principles (see Christopoulos In prep). One particularly striking case (brought to our attention by Christos Christopoulos) comes from Basque comparative adverbs, as shown in (42).

In Basque, comparatives of adjectives are derived by adding the suffix *-ago*. The adjective meaning 'good' *on* has a suppletive comparative *hobe*, which may occur with or without the comparative suffix: *hobe-(ago)*. Adverbs may be derived from adjectives by means of suffixes *-ki*∼*-gi* or *-to* (the latter not shown here). Like adjectives, adverbs may also have a comparative grade. When the adverb-forming and comparative suffix co-occur, it is the comparative suffix which is the more peripheral of the two. Now, in the comparative adverbial corresponding to *ongi* 'well', the suppletive comparative form *hobe* is used, just as it is in the adjectival comparative. Just as in the adjectival comparative, the suffix *-ago* is optional. But, when the suffix is overt, it is not adjacent to the suppleting adjectival root.

(42) Basque comparatives (de Rijk 2008)

ADJ	CMPR	ADV	ADV.CMPR	gloss
berri	berri-ago	berri-ki	berri-ki-ago	'new, recent'
on	hobe-(ago)	on-gi	hobe-ki-(ago)	'good, well'

This presents a challenge to any theory that requires adjacency for suppletive allomorphy, since there is no motivation to posit even a piece of the comparative closer to the adjectival root than the adverbial-forming affix *-ki*. The challenge can be technically circumvented in various ways (obviously, by positing an always null CMPR1 beneath the adverb-forming suffix), but it remains open whether those circumventions weaken the predictive power of an approach to contextual allomorphy built on portmanteau insertion. At a more general level, the key issue here is an empirical one: to what extent must we countenance non-adjacent interactions in allomorphy, and if there are such examples, can a structure-based theory provide the most restrictive characterization of locality consistent with the observed facts.

7 Conclusion

As we have shown, DM studies allomorphy as an interplay of rules that mediate between various grammatical components. Some rules relate syntactic structures to lexically stored morphemes. Others modify those morphemes' phonology in syntactic or lexical contexts. Every aspect of this system requires making non-trivial decisions about the nature of the syntax-lexicon-phonology relationship, the content of syntactic and lexical features, the representation of morphemes, and restrictions on rules—their targets and contexts, cyclic application, and locality. Being explicit about the relationship has allowed the theory to ask new questions about the nature of allomorphy, and to make testable empirical predictions that distinguish it from other theories of morphology. We have seen that some of these predictions have already been explored in detail, while others remain open for future exploration.

References

Akinlabi, Akinbiyi. 1996. Featural affixation. *Journal of Linguistics* 32:239–289.

Albright, Adam, and Bruce Hayes. 2003. Rules vs. analogy in English past tenses: A computational/experimental study. *Cognition* 90:119–161.

Allen, Margaret. 1978. Morphological Investigations. Doctoral Dissertation, University of Connecticut.

Anderson, Stephen R. 1992. *A-Morphous Morphology*. Cambridge: Cambridge University Press.

Arsenault, Paul. 2007. Marking the unmarked: Exceptional patterns of syncretism in English and Hindi. In *Proceedings of the 2007 annual conference of the Canadian Linguistic Association.*

Becker, Michael, and Maria Gouskova. 2016. Source-oriented generalizations as grammar inference in Russian vowel deletion. *Linguistic Inquiry* 47:391–425.

Béjar, Susana. 2003. Phi-Syntax: A theory of agreement. Doctoral Dissertation, University of Toronto.

Blevins, James P. 2016. *Word and paradigm morphology*. Oxford University Press.

Bobaljik, Jonathan. 2000. The ins and outs of contextual allomorphy. In *University of Maryland working papers in linguistics*, ed. K. K. Grohmann and Caro Struijke, volume 10, 35–71. College Park, MD: University of Maryland.

Bobaljik, Jonathan. 2012. *Universals in comparative morphology*. Cambridge, MA: MIT Press.

Bobaljik, Jonathan David, and Susi Wurmbrand. 2001. Seven Prefix-Suffix Asymmetries in Itel'men. CLS 37: The Panels. Papers from the, 205–219.

Bonet, Eulàlia, and Daniel Harbour. 2012. Contextual Allomorphy. In *The Morphology and Phonology of Exponence*, ed. Jochen Trommer, 195–235. Oxford: Oxford University Press.

Bonet, M. Eulalia. 1991. Morphology after Syntax: Pronominal Clitics in Romance. Doctoral Dissertation, MIT, Cambridge, MA.

Borowsky, Toni. 1986. Topics in the Lexical Phonology of English. Doctoral Dissertation, University of Massachusetts, Amherst.

Buell, Leston Chandler. 2005. Issues in Zulu verbal morphosyntax. Doctoral Dissertation, UCLA, Los Angeles, CA.

Bybee, Joan. 1985. *Morphology: A Study of the Relation between Meaning and Form.* Amsterdam: John Benjamins.

Bybee, Joan. 2001. *Phonology and language use.* Cambridge: Oxford University Press.

Bye, Patrik, and Peter Svenonius. 2012. Non-concatenative morphology as epiphenomenon. In *The Morphology and Phonology of Exponence*, ed. Jochen Trommer, 427–495. Oxford University Press.

Caha, Pavel. 2009a. The nanosyntax of case. Doctoral dissertation, University of Tromsø.

Caha, Pavel. 2009b. The Nanosyntax of case. Doctoral Dissertation, CASTL Tromsø.

Caha, Pavel, Karen DeClercq, and Guido Vanden Wyngaerd. 2018. The Fine Structure of the Comparative. *Studia Linguistica* Https://ling.auf.net/lingbuzz/003790.

Carstairs, Andrew. 1987. *Allomorphy in Inflexion.* Croom Helm Linguistics Series. London: Croom Helm.

Carstairs, Andrew. 1988. Some implications of phonologically conditioned suppletion. In *Yearbook of morphology 1988*, ed. Geert Booij and Jaap van Marle, 67–94. Dordrecht: Kluwer.

Chomsky, Noam, and Morris Halle. 1968. *The Sound Pattern of English.* New York: Harper and Row.

Christopoulos, Christos. In prep. Span Adjacency effects without Span Adjacency Remarks on Merchant (2015). Unpublished manuscript Harvard and UConn.

Christopoulos, Christos, and Stanislao Zompì. 2019. Allomorphy Patterns, Case Containment and the Directionality of Metasyncretism. Paper presented at Connecting roots and affixes, Brno.

Corbett, Greville. 1991. *Gender*. Cambridge Textbooks in Linguistics. Cambrdige, MA: Cambridge University Press.

Corbett, Greville G. 2007. Canonical Typolgy, Suppletion, and Possible Words. *Language* 8–42.

Embick, David. 2010. *Localism versus globalism in morphology and phonology*. Cambridge, MA: MIT Press.

Ernestus, Mirjam, and Harald Baayen. 2003. Predicting the Unpredictable: Interpreting Neutralized Segments in Dutch. *Language* 79:5–38.

Fenger, Paula. 2018. Multiple agreeing persons is not that special: Restrictions on person portmanteaux. *U. Penn Working Papers in Linguistics* 24:1–10.

Gagliardi, Annie, and Jeffrey Lidz. 2014. Statistical insensitivity in the acquisition of Tsez noun classes. *Language* 90:58–89.

Gnanadesikan, Amalia. 1997. Phonology with ternary scales. Doctoral Dissertation, University of Massachusetts at Amherst, Amherst, MA.

Goldsmith, John. 1976. Autosegmental Phonology. Doctoral Dissertation, MIT, Cambridge, MA. Published by Garland Press, New York , 1979.

Gouskova, Maria. 2012. Unexceptional segments. *Natural Language and Linguistic Theory* 30:79–133.

Gouskova, Maria, and Michael Becker. 2013. Nonce words show that Russian yer alternations are governed by the grammar. *Natural Language and Linguistic Theory* 31:735–765.

Gouskova, Maria, and Gillian Gallagher. to appear. Inducing nonlocal constraints from baseline phonotactics. *Natural Language & Linguistic Theory* URL `https://doi.org/10.1007/s11049-019-09446-x`.

Gouskova, Maria, Sofya Kasyanenko, and Luiza Newlin-Łukowicz. 2015. Selectional restrictions as phonotactics over sublexicons. *Lingua* 167:41–81. URL `http://ling.auf.net/lingbuzz/002673`.

Green, Anthony Dubach. 2006. The independence of phonology and morphology: the Celtic mutations. *Lingua* 116:1946–1985.

Green, Anthony Dubach. 2007. *Phonology limited*, volume 27 of *Linguistics in Potsdam*. Potsdam: Universitätsverlag Potsdam.

Gribanova, Vera. 2015. Exponence and morphosyntactically triggered phonological processes in the Russian verbal complex. *Journal of Linguistics* 51:519–561.

Gribanova, Vera, and Stephanie S Shih. 2017. *The morphosyntax-phonology connection: Locality and directionality at the interface*. Oxford University Press.

Hall, Beatrice L., and R.M.R. Hall. 1980. Nez Perce vowel harmony: an Africanist explanation and some theoretical questions. In *Issues in vowel harmony*, ed. Robert Vago, 201–236. Amsterdam: John Benjamins.

Halle, Morris. 1990. An approach to morphology. In *Proceedings of NELS 20*, volume 20, 150–184. Amherst, MA: GLSA Publications.

Halle, Morris, and Alec Marantz. 1993. Distributed Morphology. In *The View from Building 20. Essays in Honor of Sylvain Bromberger*, ed. Kenneth Hale and Samuel Jay Keyser. Cambridge, MA: MIT Press.

Hannahs, Stephen J, and Maggie Tallerman. 2006. At the interface: selection of the Welsh definite article. *Linguistics* 44:781–816.

Harley, Heidi. 1994. Hug a tree: Deriving the morphosyntactic feature hierarchy. *MIT working papers in linguistics* 21:289–320.

Harley, Heidi. 2014. On the identity of roots. *Theoretical linguistics* 40:225–276.

Harley, Heidi, and Elizabeth Ritter. 2002. A feature-geometric analysis of person and number. *Language* 78:482–526.

Hayes, Bruce, and Colin Wilson. 2008. A Maximum Entropy Model of Phonotactics and Phonotactic Learning. *Linguistic Inquiry* 39:379–440.

Hayes, Bruce, Kie Zuraw, Péter Siptár, and Zsuzsa Cziráky Londe. 2009. Natural and unnatural constraints in Hungarian vowel harmony. *Language* 85:822–863.

Herbert, Robert K. 1990. Labial palatalization in Nguni and Sotho languages: internal and external evidence. *South African Journal of African Languages* 10:74–80.

Hockett, Charles F. 1954. Two models of grammatical description. *Word* 10:210–234.

Jusczyk, Peter W, Paul Smolensky, and Theresa Allocco. 2002. How English-learning infants respond to markedness and faithfulness constraints. *Language Acquisition* 10:31–73.

Kager, René. 1996. On affix allomorphy and syllable counting. In *Interfaces in phonology*, ed. Ursula Kleinhenz, 155–171. Berlin: Akademie Verlag.

Kapatsinski, Vsevolod. 2010. Velar palatalization in Russian and artificial grammar: Constraints on models of morphophonology. *Laboratory Phonology* 1:361–393.

Kastner, Itamar. 2019. Templatic morphology as an emergent property. *Natural Language & Linguistic Theory* 1–49.

Kastner, Itamar, and Beata Moskal, ed. 2018. *Snippets 34: Special issue on non-local contextual allomorphy*.

Kawahara, Shigeto, and Matthew Wolf. 2010. On the existence of root-initial-accenting suffixes: An elicitation study of Japanese [-zu]. *Linguistics* 48:837–864.

Kibrik, Aleksandr E. 1977. *Opyt strukturnogo opisanija arčinskogo jazyka. II: Taksonomičeskaja grammatika*. Izdatel'stvo Moskovskogo Universiteta.

Kiparsky, Paul. 1973. 'Elsewhere' in phonology. In *A Festschrift for Morris Halle*, ed. Stephen R. Anderson and Paul Kiparsky, 93–106. New York: Holt, Rinehart and Winston.

MacBride, Alexander Ian. 2004. A constraint-based approach to morphology. Doctoral Dissertation, UCLA, Los Angeles, CA.

Marantz, Alec. 1996. 'Cat' as a phrasal idiom. Ms. MIT, Cambridge, Mass.

Mascaró, Joan. 1996. External allomorphy as emergence of the unmarked. In *Current Trends in Phonology: Models and Methods*, ed. Jacques Durand and Bernard Laks, 473–483. Salford, Manchester: European Studies Research Institute, University of Salford.

Mascaró, Joan. 2007. External allomorphy and lexical representation. *Linguistic Inquiry* 38:715–735.

Matthews, P.H. 1972. *Inflectional Morphology: A theoretical study based on aspects of Latin verb conjugation.* Cambrdige: Cambridge University Press.

McCarthy, John J. 1983. Consonantal morphology in the Chaha verb. In *The Proceedings of the West Coast Conference on Formal Linguistics 2*, ed. M. Barlow, D. Flickinger, and M. Wescoat, 176–188. Stanford: Stanford Linguistic Association.

McCarthy, John J. 1986. OCP Effects: Gemination and antigemination. *Linguistic Inquiry* 17:207–263.

McCarthy, John J. 1988. Feature geometry and dependency: A review. *Phonetica* 43:84–108.

McCarthy, John J., and Alan Prince. 1986. Prosodic Morphology. Technical Report #32, Rutgers University Center for Cognitive Science, version of 1996.

Mel'čuk, Igor A. 1976. On suppletion. *Linguistics* 14:45–90.

Mel'čuk, Igor A. 1994. Suppletion: toward a logical analysis of the concept. *Studies in Language. International Journal sponsored by the Foundation "Foundations of Language"* 18:339–410.

Merchant, Jason. To appear. How much context is enough? Two cases of span-conditioned stem allomorphy. *Linguistic Inquiry* .

Mohanan, K. P. 1982. Lexical Phonology. Doctoral Dissertation, MIT, Cambridge, MA. Distributed by IULC Publications.

Moskal, Beata. 2015a. A case study in nominal suppletion. *Linguistic Inquiry* 46:376–387.

Moskal, Beata. 2015b. Domains on the border: between morphology and phonology. Doctoral Dissertation, University of Connecticut, Storrs, CT.

Moskal, Beata, and Peter W. Smith. 2016. Towards a theory without adjacency: Hyper-contextual VI-rules. *Morphology* 26:295–312.

Myler, Neil. 2017. Exceptions to the Mirror Principle and Morphophonological "Action at a distance": The role of "word"-internal Phrasal Movement and Spell Out. In *The structure of words at the interfaces*, ed. Heather Newell, Maire Noonan, Glyne Piggot, and Lisa Travis, 100–125. Oxford University Press.

Neeleman, Ad, and Kriszta Szendrői. 2007. Radical Pro Drop and the Morphology of Pronouns. *Linguistic Inquiry* 38:671–714.

Nevins, Andrew. 2011. Phonologically-conditioned allomorph selection. In *Blackwell companion to phonology*, ed. Colin Ewen, Elizabeth Hume, Marc van Oostendorp, and Keren Rice, 2357–2382. Wiley-Blackwell.

Noyer, Rolf. 1997. *Features, positions and affixes in autonomous morphological structure.* New York: Garland.

Ostrove, Jason. 2018. Stretching, spanning, and linear adjacency in Vocabulary Insertion. *Natural Language & Linguistic Theory* 36:1263–1289.

Padgett, Jaye. 2003. Contrast and postvelar fronting in Russian. *Natural Language and Linguistic Theory* 21:39–87.

Pak, Marjorie. 2016. How allomorphic is English article allomorphy? *Glossa: A journal of general linguistics* 1.

Paster, Mary. 2006. Phonological Conditions on Affixation. Doctoral Dissertation, UC Berkeley, Berkeley, CA.

Pertsova, Katya. 2004. Distribution of Genitive Plural Allomorphs in the Russian Lexicon and in the Internal Grammar of Native Speakers. Doctoral Dissertation, UCLA, Los Angeles. Linguistics.

Pierrehumbert, Janet. 2001. Exemplar dynamics: word frequency, lenition, and contrast. In *Frequency effects and the emergence of linguistic structure*, ed. Joan Bybee and Paul Hopper, 137–157. Amsterdam: John Benjamins.

Radkevich, Nina. 2010. On Location: The structure of case and adpositions. Doctoral Dissertation, University of Connecticut.

de Rijk, Rudolf P. G. 2008. *Standard Basque*. Cambrdige: MIT Press.

Rose, Sharon, and Rachel Walker. 2004. A typology of consonant agreement as correspondence. *Language* 80:475–532.

Selkirk, Elisabeth. 1978. On prosodic structure and its relation to syntactic structure. In *Nordic Prosody*, ed. T. Fretheim, volume 2, 111–140. Trondheim: TAPIR.

Siddiqi, Daniel. 2009. *Syntax within the word: Economy, allomorphy, and argument selection in distributed morphology*, volume 138. Amsterdam: John Benjamins Publishing Company.

Siegel, Dorothy. 1974. Topics in English Morphology. Doctoral Dissertation, MIT, Cambridge, MA.

Siegel, Dorothy. 1978. The Adjacency Constraint and the theory of Morphology. In *Proceedings of NELS 8*, ed. M. Stein, 189–197. Amherst: GLSA.

Simpson, Jane, and Meg Withgott. 1986. Pronominal Clitic Clusters and Templates. In *The syntax of pronominal clitics*, ed. Hagit Borer, volume 19 of *Syntax and Semantics*, 149–174. New York: Academic Press.

Smith, Brian W. 2015. Phonologically conditioned allomorphy and UR constraints. Doctoral Dissertation, UMass Amherst, Amherst, MA.

Smith, Peter, Beata Moskal, Ting Xu, Jungmin Kang, and Jonathan David Bobaljik. 2018. Case and number suppletion in pronouns. *Natural Language and Linguistic Theory* URL `https://doi.org/10.1007/s11049-018-9425-0`.

Starke, Michal. 2009. Nanosyntax: A short primer to a new approach to language. *Nordlyd* 36:1–6.

Steriopolo, Olga. 2008. Form and function of expressive morphology: a case study of Russian. Doctoral Dissertation, University of British Columbia, Vancouver, BC.

Stump, Gregory T. 2001. *Inflectional morphology: A theory of paradigm structure*, volume 93. Cambridge University Press.

Suzuki, Keiichiro. 1998. A Typological Investigation of Dissimilation. Doctoral Dissertation, University of Arizona, Tucson, AZ.

Svenonius, Peter. 2012. Spanning. Ms. Tromsø.

Trommer, Jochen, and Eva Zimmermann. 2014. Generalised mora affixation and quantity-manipulating morphology. *Phonology* 31:463–510.

Volodin, Aleksandr Pavlovich. 1976. *Itelmenskij Jazyk*. Leningrad: Nauka.

Williams, Edwin. 2003. *Representation Theory*. Cambridge: MIT Press.

Wolf, Matthew. 2007. For an autosegmental theory of mutation. In *UMOP 32: Papers in Optimality Theory III*, ed. Leah Bateman, Michael O'Keefe, Ehren Reilly, and Adam Werle, volume 315–404. Amherst, MA: GLSA.

Wolf, Matthew. 2008. Optimal interleaving: Serial phonology-morphology interaction in a constraint-based model. Doctoral Dissertation, University of Massachusetts, Amherst, Amherst, MA.

Wolf, Matthew. 2015. Lexical insertion occurs in the phonological component. In *Understanding allomorphy: Perspectives from optimality theory*, ed. M. Eulalia Bonet, Maria-Rosa Lloret, and Joan Mascaró, 361–407. London: Equinox. Available as ROA-912 on the Rutgers Optimality Archive, `http://roa.rutgers.edu`.

Wurzel, Wolfgang. 1985. Die Suppletion bei den Dimensionsadjektiven. *Studien zur Morphologie und Phonologie* A126:114–143.

Yang, Charles. 2016. *The price of linguistic productivity*. Cambridge: MIT Press.

Zimmermann, Eva. 2017. *Morphological length and prosodically defective morphemes*. Oxford University Press.

Zoll, Cheryl. 1996. Parsing below the Segment in a Constraint-based Framework. Doctoral Dissertation, University of California, Berkeley. Available as ROA-143 on the Rutgers Optimality Archive, http://roa.rutgers.edu.

Zompì, Stanislao. 2017. Case decomposition meets dependent-case theories. Master's thesis, Università de Pisa.

Zonneveld, Wim. 1978. *A formal theory of exceptions in generative phonology*. Lisse: Peter de Ridder Press.

Zwicky, Arnold M, and Geoffrey K Pullum. 1986. The principle of phonology-free syntax: introductory remarks 63–91.