Cyclicity and Extended Exponence

Jelena Grofulović & Gereon Müller (Universität Leipzig) February 2, 2023

Abstract

The main goal of this paper is to derive, in a principled way, the Partially Superfluous Extended Exponence Generalization, according to which the more general one of two morphological exponents whose specifications for morpho-syntactic features are in a subset relation must always precede the more specific one at the base level of morphological organization. This issue has been addressed in optimality-theoretic approaches to morphology, where it has been argued that an account of the generalization requires a stratal or derivational approach to optimization (Caballero & Inkelas (2013), Stiebels (2015), Müller (2020)). In the present paper, we show that the generalization can also be derived without further ado in a Distributed Morphology approach, given that extended exponence requires feature copying (enrichment), and morphological realization obeys cyclicity: At the point where the derivation in which the more general exponent comes second could generate the required copy of a feature without violating cyclicity, the feature is already gone, due to prior insertion of the more specific exponent.

1. The Phenomenon

The concept of extended (or multiple) exponence as an issue in morphology goes back to Matthews (1972; 1974). Extended exponence refers to cases of morphological realization where a single morpho-syntactic feature seems to be expressed by more than one exponent. Thus, number is realized twice in the Archi nouns in (1-ad), with a plural exponent (*um* or *or*) that directly follows the stem accompanied by an ergative case exponent (*čaj*) that is also specified for plural (as shown in (1-be), the singular number realization is *li* or *i*). Furthermore, extended exponence of number is obligatory here; even though *čaj* realizes both plural and ergative case, the pure plural marker cannot be left out (cf. (1-cf)). Exactly the same pattern shows up with number and case exponents in dative plural inflections of German nouns in (2): In (2-ad), a pure plural exponent (*er* or *e*) is followed by a dative plural exponent *n* (the singular has a different realization as Ø or, in slightly archaic style, *e*; see (2-cf)); this does not render the pure plural exponent superfluous (cf. (2-be)).

(1) *Number in Archi Nouns*

a. gel-um-čaj
cup-PL-ERG.PL

b. gel-li
cup.SG-ERG

c. *gel-čaj
cup.SG-ERG.PL

d. qIinn-or-čaj
bridge-PL-ERG.PL

e. qIonn-i
bridge.SG-ERG

f. *qIonn-čaj
bridge.SG-ERG.PL

(2) *Number in German Nouns*:

a. Kind-er-n
child-PL-DAT.PL

b. *Kind-n
child.SG-DAT.PL

c. Kind-(e)
child.SG-DAT

d. Tisch-e-n
table-PL-DAT.PL

e. *Tisch-n
table.SG-DAT.PL

f. Tisch-(e)
table.SG-DAT

There are different types of extended exponence. A taxonomy going back to Caballero & Harris (2012) distinguishes between partially superfluous extended exponence, overlapping extended exponence, and fully superfluous extended exponence. Two morphological exponents /a/, /b/ co-occuring in a word exhibit partially superfluous extended exponence when their feature specifications are in a subset relation, as schematically depicted in (3).

(3) *Partially superfluous extended exponence*:
The feature specifications associated with two exponents are in a proper subset relation.

a. /a/ ↔ [f_1]
b. /b/ ↔ [f_1,f_2]

This is the scenario that shows up with multiple exponence of number in Archi (see (1)) and German (see (2)). Another example of such a pattern is instantiated by Meskwaki person agreement marking on verbs; see (4-ab) (cf. Dahlstrom (2000)). As observed by Caballero & Harris (2012), the subject person information is provided twice in the verb – both by the more general exponent (here the prefixes *ne* and *ke*), and by the suffixes (*pena* and *pwa*), which are more specific as they also encode subject number information in addition.

(4) *Person in Meskwaki*:

a. ne-nowi:-pena
1-go.out-1.PL
'We (excl.) go out'
b. ke-nowi:-pwa
2-go.out-2.PL
'You (pl.) go out'

A fourth and final example of partially superfluous extended exponence comes from Mari (see Alhoniemi (1993)). In Mari, the standard marker for second person singular in all tense/mood-combinations is *t*; see (5-a). In past contexts, in which a pure tense exponent *šə̑* shows up, the second person singular exponent *t* is replaced by *č́*; see (5-b). The restriction to past contexts (cf. (5-c)) implies that *č́* is specified for both person/number and tense.

(5) *Tense in Mari*:

a. kole-t
die-2SG
'You die'
b. kolə̑-šə̑-č́
die-PST-2SG.PST
'You died'
c. *kole-č́
die-2SG.PST
'You die'

Next, with overlapping extended exponence, two exponents in a word share a morpho-syntactic feature, but they also each have some morpho-syntactic feature that is not shared. A schematic illustration is given in (6); here, both abstract exponents realize a separate piece of morpho-syntactic information on the verb (viz., features [f_2] and [f_3], respectively).

(6) *Overlapping extended exponence*:
Two exponents share some morpho-syntactic feature, but their morpho-syntactic features are not in a subset relation.

a. /a/ ↔ [f_1,f_2]
b. /b/ ↔ [f_1,f_3]

One of the examples discussed by Caballero & Harris (2012) comes from Filomeno Mata Totonaco (cf. Inkelas et al. (2006), McFarland (2009)). As can be seen from (7-b), in second person singular progressive contexts, Totonacan morphology realizes second person subject information on the stem, on the progressive marker, and on the number marker; thus, there is a massive overlap of second person exponence. Still, none of the feature sets associated with the three morphological exponents is a subset of the feature set of any other exponent.

(7) *Person in Totonacan*:

a. min-maa
come-PROG
'he is coming'

b. tan-paa-ti
come.2SUBJ-PROG.2SUBJ-2SUBJ.SG
'you are coming'

c. *min-maa-ti
come-PROG-2SUBJ.SG

Finally, Caballero & Harris (2012) recognize fully superfluous extended exponence as a third pattern. As shown in (8), here the sets of morpho-syntactic features associated with two (or more) exponents in a word are identical.

(8) *Fully superfluous extended exponence*:
Two exponents have identical feature specifications.

a. /a/ $\leftrightarrow$ [f_1,f_2]
b. /b/ $\leftrightarrow$ [f_1,f_2]

However, for many of the relevant examples, closer inspection reveals that they do not instantiate extended exponence after all, e.g., because what at first sight looks like two separate exponents actually qualifies as a single discontinuous exponent, or because the features that are involved are not in fact identical, or because copying (of the whole exponent, i.e., including the form) is involved (see Stiebels (2015; 2016) and Müller (2020) for some case studies).

Causative formation in Sinhala (see Fenger & Weisser (2023)) might be an instance of discontinuous exponence. (9-a) shows the base form of a verb. In (9-b), causativization has applied, and a causative exponent shows up. With a verb of this type, the causative exponent takes the form of a zero item that triggers gemination of the stem-final consonant (plus schwa epenthesis), leading to *də*. However, in addition, there is also a second verb class that handles causative formation differently, viz., by adding a causative affix *wə*. Importantly, this segmental exponent can also be added to the non-segmental exponent of the first verb class without a change in meaning; see (9-c). This might then instantiate a scenario of the type in (8). Alternatively, and this is the analysis that we would like to adopt here, the (optional) co-occurence of a segmental and a suprasegmental marking in (9-c) can be understood in the same way as, say, plural markers like *e* and *er* in German (see (2) above), which may in addition trigger Umlaut on a preceding syllable (cf., e.g., *Schaf-e* vs. **Schä-f-e* ('sheep-pl') and **Hand-e* vs. *Händ-e* ('hand-pl')). In both the Sinhala and the German case, the most straightforward analysis would presumably postulate a single segmental exponent that can or must be accompanied by abstract supra-segmental information that subsequently phonologically modifies the stem (see, e.g., Wiese (2000) and Trommer (2011)).

(9) *Causative in Sinhala*:

a. adi-nə-wa
pull-NPST-IND
'pull'

b. ad-də-nə-wa
   pull-CAUS-NPST-IND
   'make somebody pull'

c. ad-də-wə-nə-wa
   pull-CAUS-CAUS-NPST-IND
   'make somebody pull'

Assuming that this result can be generalized, we will postulate, here and henceforth, that fully superfluous extended exponence does not in fact exist. It can also be noted that such a phenomenon would be unexpected under many theories of inflectional morphology; and the approach that we will develop also has this property. This leaves overlapping extended exponence and partially superfluous extended exponence as explananda for morphological theories. While overlapping extended exponence turns out to be unproblematic under most theories of inflectional morphology, the case is different with partially superfluous extended exponence, which raises problems for various restrictive theories of inflectional morphology: Essentially, the question is why the availability (and presence) of a more specific exponent like /b/ $\leftrightarrow$ [f_1,f_2] does not block a more general exponent /a/ $\leftarrow$ [f_1]; given the availability of /b/, the occurrence of /a/ looks redundant, and might be expected to be blocked for this reason. It is the primary goal of the present paper to give a principled answer to this question on the basis of Distributed Morphology (cf. Halle & Marantz (1993)).

We will proceed as follows. In section 2, we discuss the nature of this problem, and some solutions that have been advanced, against the background of a morphological theory in which it shows up in a particularly obvious way, viz., Optimality Theory. All existing solutions have a common core, which centers around what we call the Partially Superfluous Extended Exponence Generalization: The more general exponent must be closer to the stem than the more specific exponent. In section 3, we then turn to Distributed Morphology. We show that, as it stands, the Partially Superfluous Extended Exponence Generalization cannot yet be derived under any existing approach based on Distributed Morphology; in fact, the only such approach that might have anything to say about the phenomenon (viz., Bobaljik (2000)) turns out to make predictions that are diametrically opposed to the ones covered by the generalization. However, we show that by clarifying the nature of context features, by treating every instance of extended exponence as a consequence of a post-syntactic feature copy operation ('enrichment'; Müller (2007)), and, most importantly, by invoking cyclicity (the Cyclic Principle and the Strict Cycle Condition), the generalization can be derived. Finally, section 4 discusses some empirical challenges to the generalization and the Distributed Morphology account from which it follows.

2. Optimality Theory

Optimality Theory (cf. Prince & Smolensky (2004)) highlights the general problem with partially superfluous extended exponence because economy of representation is straightforwardly derived in this approach: Every grammatical operation automatically incurs some violation (e.g., of a faithfulness constraint), and this implies that, ceteris paribus, if all relevant constraints can be satisfied without this operation, it will be precluded. Hence, in a partially superfluous extended exponence scenario like (3), the more general exponent /a/ should always be blocked by the more specific exponent /b/: The presence of /a/ is per se costly; it invariably violates some (low-ranked) constraint. Therefore, it seems that /a/ has nothing to contribute that could not be obtained with /b/ alone: The constraint profile of a candidate with /a/ and /b/ must be worse than the constraint profile of a candidate with just /b/.

This reasoning is illustrated for the competition underlying extended exponence of number with Archi nouns (recall (1)) in the tableau in (10). For the purposes of the present discussion, let us make the following assumptions about optimality-theoretic morphology:[1] First, suppose that the input for morphological exponence is a stem, together with a fully specified set of morpho-syntactic features that need to be realized by exponents – in the case at hand, I_1 has a [+pl] number feature and a case specification [–obl(ique),+gov(erned)] that represents the ergative. Second, the competing output candidates O_{11}, O_{12}, etc., have carried out morphological realization to different degrees, and with different exponents.[2] Third, faithfulness constraints derive the compatibility and specificity requirements that are stipulated as parts of a constraint like the Subset Principle (cf. Halle (1997)) or Panini's Principle (cf. Stump (2001)) in other morphological theories, like Distributed Morphology or Paradigm Function Morphology. More specifically, ID(ENT)-F(EATURE) ensures compatbility (i.e., exponents have feature specifications that are subsets of the target specification of the input), and MAXNUM and MAXCASE demand realization of number and case features of the input by output exponents. Finally, a low-ranked *STRUC(TURE) is here assumed to stand for whatever constraints ensure that no grammatical operation comes for free (including, of course, morphological exponence), and that thus derive the general economy effect in optimality theory. On this basis, the competition in (10) makes it clear that the intended winner (i.e., O_{14}, which exhibits partially superfluous extended exponence) has no chance to ever become optimal (signalled here by ☆); it will always be blocked by a more economical output candidate (viz., O_{11}, which dispenses with the gratuitous more general exponent that realizes number but not case); the wrong winner is indicated by ☛ here.

(10) *Extended exponence as a problem* (standard parallel optimality theory):

I_1: /[$_N$ gel:[+pl,–obl,+gov]]/	ID-F	MAXNUM	MAXCASE	*STRUC
☛O_{11}: gel[+pl,–obl,+gov]-čaj[+pl,–obl,+gov]				*
O_{12}: gel[+pl,–obl,+gov]-um[+pl]			*!*	*
O_{13}: gel[+pl,–obl,+gov]-li[–obl,+gov]		*!		*
☆O_{14}: gel[+pl,–obl,+gov]-um[+pl]-čaj[+pl,–obl,+gov]				**!

To the best of our knowledge, there is no solution to this problem in standard parallel optimality theory, as devised in Prince & Smolensky (2004). However, various solutions have been proposed that rely on versions of optimality theory that either invoke strata, or that are inherently derivational; stratal and derivational approaches have in common that they presuppose that grammatical operations (like morphological exponence, in the case at hand) can be ordered with respect to one another, such that an operation can become *opaque* (cf. Kiparsky (1973)), in the sense that it would be bled by another operation but is not factually bled because that other operation applies too late (i.e., counter-bleeding takes place).

More specifically, the solutions to the problem in (10) that have been suggested in Caballero & Inkelas (2013), Stiebels (2015), and Müller (2020) all take the following form: Assuming that there are several optimization procedures, which are organized sequentially, these approaches converge on the assumption that for the first optimization procedure in the

[1] What follows is an amalgamation and simplification of various different optimality-theoretic approaches to morphology; see, e.g., Grimshaw (2001), Trommer (2001), Don & Blom (2006), Ortmann (2004), and Stiebels (2006), among many others.

[2] I.e., a *realizational*, rather than *incremental*, approach to morphology is adopted; cf. Stump (2001). In fact, the problems with licensing partially superfluous extended exponence are exacerbated in an incremental approach, according to which there are no morpho-syntactic features in a word except for those contributed by morphological exponents; see, e.g., Wunderlich (1997).

scenario in (3), only /a/ $\leftrightarrow$ [f_1] is available, satisfying the constraint demanding realization of [f_1] (but not the constraint demanding realization of [f_2]). Subsequently, /b/ $\leftrightarrow$ [f_1,f_2] becomes available, and is selected so as to also satisfy the constraint demanding realization of [f_2]. The systems are *myopic*: An earlier selection of /b/ would have made selection of /a/ impossible (bleeding), but since /b/ is not initially available, selection of /a/ is *counter-bled* by subsequent selection of /b/.

While these stratal/derivational approaches all share a common core, the concrete implementations differ substantially; most importantly, the answer given to the question of *why* a more specific /b/ $\leftrightarrow$ [f_1,f_2] is not initially available, so that less specific /a/ $\leftrightarrow$ [f_1] can become optimal at an early stage, is addressed in diverging ways. To begin with, the approach proposed by Caballero & Inkelas (2013) presupposes *strata* (see Kiparsky (1982)): On this view, /a/ belongs to stratum 1, /b/ belongs to stratum 2, and optimization in stratum one (where /b/ is not yet available) precedes optimization in stratum 2. Second, Stiebels' (2015) analysis makes use of *f-seq* (see Wunderlich (1997), Starke (2001)): The order of exponent selection follows the functional sequence of grammatical categories. If [f_2] outranks [f_1] on f-seq, the exponent /b/ that (also) realizes [f_2] must come after the exponent /a/ that (only) realizes [f_1]. Finally, the approach developed in Müller (2020) relies on a constraint *Minimize Satisfaction* that is independently designed to capture effects like those covered by Chomsky's (2001) Merge over Move constraint. *Minimize Satisfaction* is an overarching, inviolable constraint demanding (non-zero) minimization of new constraint satisfactions in a derivational version of optimality theory (*harmonic serialism*; cf. McCarthy (2016)). In this approach to inflectional morphology, partially superfluous extended exponence is possible since the more general exponent /a/ (yielding fewer new constraint satisfactions) is optimal at an early stage but selection of more specific /b/ is both required and permitted at a later stage (because it improves the constraint profile, and there is no exponent left that would do so with fewer new constraint satisfactions at this point).

All these optimality-theoretic approaches relying on derivational order derive the generalization in (11).

(11) *The Partially Superfluous Extended Exponence Generalization*:
If there are two exponents /a/ $\leftrightarrow$ [f_1] and /b/ $\leftrightarrow$ [f_1,f_2] in a word, /a/ is realized closer to the stem than /b/.

Based on a preliminary investigation of the typological record, Stiebels (2015) ventures the hypothesis that this prediction is corroborated by the empirical evidence in the world's languages. In what follows, we will postulate that this is indeed the case (we will address some pieces of apparent counter-evidence in section 4). Given this state of affairs, the question arises of whether the Partially Superfluous Extended Exponence Generalization can also be derived in other, non-optimality-theoretic approaches to morphology. The prospects would seem to be bleak for any theory that is (a) non-derivational, and that (b) intrinsically permits unlimited feature realization by multiple exponents, like Paradigm Function Morphology (see Stump (2001)) or Network Morphology (see Brown & Hippisley (2012)). However, things might be different with Distributed Morphology, which is derivational in nature, and which associates each morpho-syntactic feature with a designated functional head.

3. Distributed Morphology

3.1. State of the Art

Can the Partially Superfluous Extended Exponence Generalization in (11) also be derived in Distributed Morphology? As a first step towards an answer, it can be noted that the phenomenon of extended exponence is typically addressed by recourse to *secondary, contextual features* in Distributed Morphology. Such contextual features are associated with exponents just like primary, "core" features, but in contrast to the latter, they are usually put in brackets. Thus, instead of the two exponents /a/ $\leftrightarrow$ [f_1] and /b/ $\leftrightarrow$ [f_1,f_2] in a partially superfluous extended exponence scenario (cf. (3)), we get /a/′ $\leftrightarrow$ [f_1] (as before) and /b/′ $\leftrightarrow$ [f_2] ([f_1]), where [f_2] counts as primary and ([f_1]) counts as secondary. Furthermore, a contextual feature of an exponent like /b/′ is not matched by the functional head X into which /b/′ is inserted; rather, the matching [f_2] feature is located on some other functional head Y in the vicinity of X. Thus, on this view, extended exponence qualifies as *contextual allomorphy*.

For this reason, one might expect that restrictions for contextual allomorphy as they have been proposed in Distributed Morphology have some bearing on patterns of extended exponence. There are two relevant concepts. The first one is that of a *morphological phase* (see Marvin (2002), Embick (2010), and Bermúdez-Otero (2011), among others). Morphological phases act as locality domains for morphological realization, and thus ensure that a secondary, contextual feature of an exponent must be matched within this domain. However, irrespective of the exact definition of morphological phases, it seems clear that due to their size, they cannot systematically restrict patterns of partially superfluous extended exponence in an interesting way, let alone derive the Partially Superfluous Extended Exponence Generalization. The second relevant concept that has been argued to govern morphological exponence is that of *cyclicity*. And indeed, as argued by Bobaljik (2000), subjecting morphological exponence to cyclicity potentially can successfully restrict extended exponence.

Bobaljik's (2000) approach to contextual allomorphy rests on three basic assumptions. First, there is what he calls *separation*: Morphology interprets syntactic structures, rather than feeding them; i.e., a realizational approach is adopted (see footnote 2), as it is standardly assumed in Distributed Morphology. Second, the operation of vocabulary insertion that brings about morphological realization in Distributed Morphology implies feature *discharge* (or 'rewriting', in Bobaljik's terminology) in the target functional head (cf. Noyer (1997) and Trommer (1999)): Matched features are used up by vocabulary insertion and no longer a part of the representation. Third, morphological realization is subject to *cyclicity*, in the sense that vocabulary insertion proceeds root-outwards.

This system makes the following two predictions, one for morpho-syntactic features, and a contrary one for morpho-phonological diacritic features. First, outwards-sensitivity to morpho-syntactic features on functional heads in a complex word is possible because the heads hosting those features have not yet been subject to vocabulary insertion (and, hence, feature discharge); however, inwards-sensitivity to such morpho-syntactic features is not possible because, given cyclicity, they have already been discharged as a consequence of earlier vocabulary insertion. Second, the inwards-sensitivity to morpho-phonological diacritic features (like inflection class) on vocabulary items in a complex word is possible (because these features, by assumption, were originally brought into the structure by the root vocabulary item, in minimal violation of a the tenets of a realizational approach); in contrast, outwards-sensitivity to such diacritics is not possible (because the items that introduce them into the structure are not yet present, given cyclicity).

Unfortunately, this approach does not derive the Partially Superfluous Extended Exponence Generalization in (11); in fact, it predicts more or less the opposite of what is covered by it: Given the generalization, it should be the case that the outer (more specific) exponent /b/′ $\leftrightarrow$ [f_2] ([f_1]) requires the contextual presence of the morpho-syntactic feature [f_1] matched by the inner (more general) exponent /a/′ $\leftrightarrow$ [f_1]; however, once /a/′ has been inserted into a head X, [f_1] is gone from X, and subsequent insertion of /b/′ $\leftrightarrow$ [f_2] ([f_1]) in an outer head Y will be impossible because the latter exponent does not find the contextual feature ([f_1]) in X anymore that it needs to satisfy the compatibility ('subset') requirement of the Subset Principle. Consequently, a form like *Kind-er-n* ('child-PL-DAT.PL') (or any other instance of partially superfluous extended exponence in the above examples) can never be generated under these assumptions: The exponent /er/ $\leftrightarrow$ [+pl] is inserted first (given cyclicity) into a functional head X (which one may consider a number head #), thereby discharging (and removing) the number feature [+pl] from X, and subsequent insertion of /n/ $\leftrightarrow$ [+obj,+obl] ([+pl]) into a higher functional head Y (which we may identify as K, for case, with [+obj,+obl] standing for dative) will fail because ([+pl]) does not find a matching feature in the syntactic representation anymore.

This problem can in principle be solved by postulating that morpho-syntactic features that are discharged by exponent insertion into a functional head are *deleted* (and thus not accessible for direct insertion of another vocabulary item into the same head), but not fully *erased* (see Chomsky (1995)), so that subsequent reference by more specific exponents is still permitted. Such a modification of the concept of discharge would imply that all existing cases of extended exponence can be covered, but since restrictions on exponence are now weakened, it does not come as a surprise that it would not get us any closer to deriving the Partially Superfluous Extended Exponence Generalization in (11). In fact, the resulting approach would be hardly distinguishable from one where all features are available everywhere (modulo phases), all the time, as in Paradigm Function Morphology or Network Morphology (see above). The only remaining restriction would be that outwards-sensitivity to morpho-phonological diacritic features would still predicted not to be possible. In view of this, we take it that there is every reason to develop a new approach to extended exponence in Distributed Morphology that captures the generalization in (11). We lay out such an approach in the next section.

3.2. A New Approach

The new approach relies on six assumptions. The first assumption concerns *disjunctive blocking*, and it is a standard one made in Distributed Morphology (see, e.g., Halle & Marantz (1993)): Only one vocabulary item can be inserted into a given functional head. As a consequence, two morphological exponents that participate in extended exponence in a word must have been inserted into two separate functional heads.

The second assumption is shared with Bobaljik's approach: Morphological exponence involves *discharge* (see Noyer (1997) and Trommer (1999)): The insertion of an exponent with a matching feature discharges, and thereby deletes, a feature in the locally accessible domain (the morphological phase).

We refer to the third assumption as *feature uniqueness*: There is no distinction between "primary" and "secondary/contextual" features on morphological exponents; all morpho-syntactic features of a vocabulary item are of the same kind. It follows from this assumption that the specific exponents in a partially superfluous extended exponence scenario must look as in (12-a), and cannot take a form like the one in (12-b).

(12) a. $\surd$/b/ $\leftrightarrow$ [f_1,f_2]
     b. */b/′ $\leftrightarrow$ [f_2]([f_1])

The conclusion that there can be no meaningful ontological differences among the morpho-syntactic features characterizing a morphological exponent has been most forcefully defended in Stump (2001, ch. 5) (also cf. Müller (2020, ch 3.)). Problems with contextual features pointed out by Stump include *ambiguity* (How can it be that one and the same exponent may qualify as a primary exponent of a given morpho-syntactic property in one environment, and as a secondary exponent of the same morpho-syntactic property in another environment?); *learnability* (How can a child acquiring a language decide whether a given feature on some vocabulary item is a primary or a secondary (contextual) one?); *specificity* (To what extent do contextual features count for the specificity of exponents that bear them?);[3] and *locality* (How far away can a contextual feature on a functional head be located from the exponent that shares it?). We take the first two of these problems to be decisive.

Fourth, we assume that extended exponence never comes for free; it must always be brought about by *enrichment* (see Müller (2007)). Enrichment is a post-syntactic feature copying operation that is a mirror image of post-syntactic impoverishment; enrichment generates *local copies* of morpho-syntactic features on a functional head, which can then give rise to extended exponence because there are now two (or more, if enrichment applies more than once) identical features that can be separately subject to discharge.[4]

Fifth, these four assumptions require a clarification of the core concepts of *compatibility and specificity* governing vocabulary insertion. A modified Subset Principle (cf. Halle (1997)) is called for that is made sensitive to contextual features in syntactic representations, i.e., features of the syntactic contexts which are accessed by morphological exponents that bear them, but that are not located in the functional head into which the exponent is inserted. We adopt the following version of the Subset Principle.

(13) *Subset Principle*:
A vocabulary item *V* is inserted into a functional morpheme *M* contained in a morphological phase *P* (thereby discharging all matching features in *P*) iff (a) and (b) hold:

a. *Compatibility*:
V realizes a feature of *M*, and the morpho-syntactic features of *V* are a subset of the morpho-syntactic features of *P*.

b. *Specificity*:
Among the vocabulary items that satisfy (a), there is no V′ that realizes more features of *M* than *V*.

Thus, an exponent can only be inserted into a given head if it shares a feature with it, and if all other features that the exponent may be equipped with are available for insertion (and, consequently, discharge) on some head (which may or may not be the same head) in the local domain (the morphological phase). If there is more than one exponent that satisfies this compatibility requirement, specificity selects the one(s) that realize(s) most features in the head into which insertion takes place. Note that it is only the features of this head *M* that play a role for specificity, and *not* the features in the syntactic context *P*. As we will see, this may in principle lead to situations where more than one vocabulary item could be inserted in accordance with the Subset Principle; and this not only if the feature specifications two exponents are identical (cf. Hein (2008) and Driemel (2018)), but also if they differ. However, such optionality is not ususally found in morphological paradigms; it is precluded by the final assumption to be mentioned here.

Sixth and finally, the present approach incorporates *cyclicity*. Cyclicity manifests itself in two different but related constraints, both of which can be shown to be required in derivational systems based on cyclic application of operations.[5] The first constraint is the Cyclic

[3] However, see Arregi & Nevins (2012) and Hanink (2018) for some suggestions.
[4] Feature copying is a well-established operation in Distributed Morphology; see Halle & Marantz (1993), Embick & Noyer (2007), and Norris (2014), among many others.
[5] Cf. Müller (2023) for an overview of the arguments for this claim; also see below.

Principle in (14) (see Perlmutter & Soames (1979) for the formulation; and Adger, Béjar & Harbour (2003), Embick (2010), and Kalin & Weisser (2021), next to Bobaljik (2000), for applications in morphology).[6]

(14) *Cyclic Principle*:
When two operations can be carried out, where one applies to the cyclic domain D_x and the other applies to the cyclic domain D_{x-1} included in D_x, then the latter is applied first.

Assuming the most restrictive concept where every projection in a tree qualifies as a cyclic domain, the Cyclic Principle ensures that post-syntactic morphological operations like vocabulary insertion and enrichment (i.e., feature copying) apply root-outwards, exactly as in Bobaljik's approach.

The second cyclicity constraint that we will adopt is the Strict Cycle Condition (cf. Chomsky (1973; 1995; 2015)); a simple version of the constraint is given in (15). As before, it can be assumed that every projection qualifies as a cyclic domain.

(15) *Strict Cycle Condition*:
Once a cyclic domain D_x has been affected by an operation, no subsequent operation may exclusively affect a cyclic domain D_{x-1} that is a proper subdomain of D_x.

It remains to be shown how this set of assumptions derives the Partially Superfluous Extended Exponence Generalization in (11). In a nutshell, the underlying logic will be as follows. If vocabulary insertion is to lead to extended exponence, it is clear that feature copying (enrichment) is required; this is the only way how two exponents bearing this feature can show up in a word where there is initially one one occurrence of the feature. Thus, feature copying and vocabulary insertion are two post-syntactic operations that will invariably interact; and the interaction of the two processes is governed by the the cyclicity constraints on the one hand, and by the compatibility and specificity requirements of the Subset Principle on the other hand. As we will see, the interaction ensures that the more *general vocabulary item* only has a chance to show up in a word if it is inserted *early*, in a position close to the root: At the point where the derivation could generate the required copy of the feature for a more general exponent that comes second without violating cyclicity, the feature is already gone as a consequence of insertion of the more specific exponent. If, on the other hand, the more general exponent comes first, no such problem arises: Enrichment can apply early to the crucial feature that is shared by the two exponents. Therefore, the more *specific vocabulary item*, which has more options to satisfy the Subset Principle since it is equipped with more features that permit more insertion sites, can be inserted *later*.

3.3. Deriving Partially Superfluous Extended Exponence

Let us assume, as before, that there are two vocabulary items, the more general exponent /a/ $\leftrightarrow$ [f_1] and the more specific exponent /b/ $\leftrightarrow$ [f_1,f_2]; and that, furthermore, X and Y are functional categories (here assumed to be suffixal) with morpho-syntactic features in need of realization by vocabulary insertion, where X hosts one of these features, and Y hosts the other feature, and X is closer to the root than Y.[7] Then, two basic scenarios need to be

[6] Also note that the widely adopted Earliness Principle (see Pesetsky & Torrego (2001)) and Featural Cyclicity (see Richards (2001) and Preminger (2018)) are basically just versions of the Cyclic Principle for syntactic derivations.

[7] However, everything that follows can be generalized to scenarios where more features than just two are involved.

considered: In the first one, [f_1] is in X, and [f_2] is in Y. In the second scenario, it is the other way round: [f_2] is in X, and [f_1] is in Y. Thus, it follows that /a/ ↔ [f_1] can only ever have a chance to be inserted into X in the first scenario, and into Y in the second; a priori, there is no such restriction for /b/ ↔ [f_1,f_2], which is equipped with both features.

On this basis, let us start with the first scenario (which we will henceforth also refer to as "a→X"): The more general exponent /a/ bears a feature that shows up in the hierarchically lower head X. The abstract derivation in (16) shows how this scenario can lead to a successful instance of partially superfluous extended exponence, in accordance with (11).[8]

(16) *Scenario a→X, derivation 1:* √

a. Initial structure:
$[_Y [_X [_c \sqrt{\ } c]\ X_{[f_1]}]\ Y_{[f_2]}]$

b. Root lexicalization:
$[_Y [_X [_c \sqrt{\alpha}\ c]\ X_{[f_1]}]\ Y_{[f_2]}]$

c. Feature copying on X cycle:
$[_Y [_X [_c \sqrt{\alpha}\ c]\ X_{[f_1],[f_1]}]\ Y_{[f_2]}]$

d. Vocabulary insertion on X cycle:
$[_Y [_X [_c \sqrt{\alpha}\ c]\ [_{X_{[f_1]}}$ /a/]] $Y_{[f_2]}$]

e. Vocabulary insertion on Y cycle:
$[_Y [_X [_c \sqrt{\alpha}\ c]\ [_X$ /a/]] $[_Y$ /b/]]

Starting with the complex head in (16-a) (formed by earlier head movement of $\sqrt{\ }$ to c, of c to X, and of X to Y, in the syntax or at the beginning of post-syntax), the root is lexicalized first (indicated by α) in (16-b) (given cyclicity); this completes the c cycle. Next, the derivation moves to the X cycle. Suppose that the first operation that applies is enrichment in (16-c); such feature copying is effected by a designated enrichment rule like (17).

(17) Ø → [f_1]/[f_1]__.

Now there are two features [f_1] available in X in (16-c). In the next step in (16-d), the more general exponent /a/ ↔ [f_1] is inserted into X, thereby discharging one of the two [f_1] features there. Note that this does not violate the Subset Principle in (13).[9] Finally, the derivation reaches the Y cycle in (16-e), and inserts /b/ ↔ [f_1,f_2] in Y, in accordance both with the Subset Principle ([f_2] on /b/ is matched and discharged in Y, and [f_1] on /b/ is matched and discharged in X), the Cyclic Principle, and the Strict Cycle Condition (insertion of /b/ affects the embedded domain X by discharging [f_1], but it does not do so exclusively since it also affects the Y domain).

The derivation in (16) thus gives rise to partially superfluous extended exponence. It essentially underlies all instances of the phenomenon discussed in section 1. This is shown for the case of dative plural nouns in German (cf. (2)); (18) parallels (16) in all relevant respects.

(18) *A well-formed derivation for dative plural nouns in German*:

[8] Some remarks on notation. Here and in what follows, the current cyclic domain is rendered in black, and the domain which is not yet affected by some operation in gray. $\sqrt{\ }$ designates the root, c the categorizing head, and α the root vocabulary item inserted into $\sqrt{\ }$.

[9] Does this imply that there could be optionality of /a/ in /b/ in a minimally different system where there is no enrichment, and hence no extended exponence? This is not the case if there is an overarching requirement demanding every head's morpho-syntactic features to be discharged if possible; also, in such a system, there would be arguably no good evidence for two separate heads X and Y in the morphology to begin with.

Exponents:
(i) /er/ $\leftrightarrow$ [+pl]
(ii) /n/ $\leftrightarrow$ [+pl,+obl,+gov]

a. Initial structure:
$[_{\rm Y}\ [_{\rm X}\ [_{\rm c} \sqrt{\ }\ {\rm c}\ ]\ {\rm X}_{[+{\rm pl}]}\ ]\ {\rm Y}_{[+{\rm obl},+{\rm gov}]}\ ]$

b. Root lexicalization:
$[_{\rm Y}\ [_{\rm X}\ [_{\rm n}\ {\rm Kind}\ \text{Ø}\ ]\ {\rm X}_{[+{\rm pl}]}\ ]\ {\rm Y}_{[+{\rm obl},+{\rm gov}]}\ ]$

c. Feature copying on X cycle:
$[_{\rm Y}\ [_{\rm X}\ [_{\rm n}\ {\rm Kind}\ \text{Ø}\ ]\ {\rm X}_{[+{\rm pl}],[+{\rm pl}]}\ ]\ {\rm Y}_{[+{\rm obl},+{\rm gov}]}\ ]$

d. Vocabulary insertion on X cycle:
$[_{\rm Y}\ [_{\rm X}\ [_{\rm n}\ {\rm Kind}\ \text{Ø}\ ]\ [_{{\rm X}_{[+{\rm pl}]}}\ {\rm er}\ ]]\ {\rm Y}_{[+{\rm obl},+{\rm gov}]}\ ]$

e. Vocabulary insertion on Y cycle:
$[_{\rm Y}\ [_{\rm X}\ [_{\rm n}\ {\rm Kind}\ \text{Ø}\ ]\ [_{\rm X}\ {\rm er}\ ]]\ [_{\rm Y}\ {\rm n}\ ]]$

As a matter of fact, it turns out that derivation 1 in scenario a→X is the *only* derivation that can give rise to partially superfluous extended exponence. Consider, e.g., a derivation that is minimally different from derivation 1 in that the order of the two insertion operations is reversed, and /b/ is inserted into Y before /a/ is inserted into X; cf. (19). As illustrated in (19-d), premature insertion of /b/ on the Y cycle violates the Cyclic Principle (since insertion of /a/ on the X cycle is skipped); furthermore, final insertion of /a/ would then also violate the Strict Cycle Condition (so this is an environment where the two cyclicity constraints make the same predictions).

(19) *Scenario a→X, derivation 2: **

a. Initial structure:
$[_{\rm Y}\ [_{\rm X}\ [_{\rm c} \sqrt{\ }\ {\rm c}\ ]\ {\rm X}_{[{\rm f}_1]}\ ]\ {\rm Y}_{[{\rm f}_2]}\ ]$

b. Root lexicalization:
$[_{\rm Y}\ [_{\rm X}\ [_{\rm c} \sqrt{\alpha}\ {\rm c}\ ]\ {\rm X}_{[{\rm f}_1]}\ ]\ {\rm Y}_{[{\rm f}_2]}\ ]$

c. Feature copying on X cycle:
$[_{\rm Y}\ [_{\rm X}\ [_{\rm c} \sqrt{\alpha}\ {\rm c}\ ]\ {\rm X}_{[{\rm f}_1],[{\rm f}_1]}\ ]\ {\rm Y}_{[{\rm f}_2]}\ ]$

d. Vocabulary insertion on Y cycle **Cyclic Principle*:
$[_{\rm Y}\ [_{\rm X}\ [_{\rm c} \sqrt{\alpha}\ {\rm c}\ ]\ {\rm X}_{[{\rm f}_1]}\ ]]\ [_{\rm Y}\ /{\rm b}/\ ]]$

e. Vocabulary insertion on X cycle **Strict Cycle Condition*:
$[_{\rm Y}\ [_{\rm X}\ [_{\rm c} \sqrt{\alpha}\ {\rm c}\ ]\ [_{\rm X}\ /{\rm a}/\ ]]\ [_{\rm Y}\ /{\rm b}/\ ]$

Another derivation that is doomed to fail in the a→X scenario is given in (20). The first three steps are as before; however, in the fourth step in (20-d), it is /b/ $\leftrightarrow$ [f_1,f_2] (rather than /a/ $\leftrightarrow$ [f_1]) that is inserted into the X node. This, as such, is in accordance with both the Subset Principle and the cyclicity constraints: Insertion of /b/ affects both the X cycle and the Y cycle (the latter via deletion of [f_2]), but this is unproblematic. However, subsequent insertion of /a/ $\leftrightarrow$ [f_1] in (20-d) will now be impossible because of the Subset Principle (Y only hosts an incompatible [f_2] feature to begin with, and earlier insertion of /b/ has removed this feature in any event).

(20) *Scenario a→X, derivation 3: **

a. Initial structure:
$[_{\rm Y}\ [_{\rm X}\ [_{\rm c} \sqrt{\ }\ {\rm c}\ ]\ {\rm X}_{[{\rm f}_1]}\ ]\ {\rm Y}_{[{\rm f}_2]}\ ]$

b. Root lexicalization:
$[_{\rm Y}\ [_{\rm X}\ [_{\rm c} \sqrt{\alpha}\ {\rm c}\ ]\ {\rm X}_{[{\rm f}_1]}\ ]\ {\rm Y}_{[{\rm f}_2]}\ ]$

c. Feature copying on X cycle:

$[_{Y}\ [_{X}\ [_{c}\ \sqrt{\alpha}\ c\ ]\ X_{[f_1],[f_1]}\ ]\ Y_{[f_2]}\ ]$

d. Vocabulary insertion on X cycle:
$[_{Y}\ [_{X}\ [_{c}\ \sqrt{\alpha}\ c\ ]\ [_{X_{[f_1]}}\ /b/\ ]]\ Y\ ]$

e. Vocabulary insertion into Y on Y cycle **Subset Principle*
$[_{Y}\ [_{X}\ [_{X}\ [_{c}\ \sqrt{\alpha}\ c\ ]\ [_{X_{[f_1]}}\ /b/\ ]]\ [_{Y}\ /a/\ ]]$

Further derivations based on an a→X scenario are also excluded. In particular, derivations in which, on a given cycle where it can apply, enrichment does not precede feature-removing vocabulary insertion can never give rise to extended exponence.[10]

Thus, as an interim conclusion regarding a→X scenarios, it can be noted that the more general exponent /a/ can occur before the more specific exponent /b/ as a result of one derivation (viz., derivation 1); other derivations of an /a/-/b/ sequence fail (cf. derivation 2), and the reverse /b/-/a/ order cannot be generated in an a→X scenario for very basic reasons (derivation 3). It now remains to be shown that an /a/→Y scenario, where the feature $[f_1]$ of the more general exponent /a/ is matched by the outer head Y (i.e., where /a/ realizes a hierarchically higher feature) cannot lead to a successful derivation under present assumptions – in particular, a /b/-/a/ order that would violate the Partially Superfluous Extended Exponence Generalization in (11) must not be generated.

As before, suppose that the enrichment rule in (17) is active, and that /a/ ↔ $[f_1]$ and /b/ ↔ $[f_1,f_2]$ are as before; the only difference is that the lower head X now has $[f_2]$, and the higher head Y has $[f_1]$. As shown by derivation 4 in (21), if the more specific exponent /b/ is inserted on the X cycle in (21-c), it will remove both $[f_2]$ from X *and* $[f_1]$ from the higher head Y.[11] When the derivation subsequently moves to the Y cycle, there are no features left for carrying out enrichment, and feature copying will not apply (cf. (21-d)). As a consequence, the more general exponent cannot be inserted: There are no features left for morphological exponence of Y.

(21) *Scenario a→Y, derivation 4: **

a. Initial structure:
$[_{Y}\ [_{X}\ [_{c}\ \sqrt{\ }\ c\ ]\ X_{[f_2]}\ ]\ Y_{[f_1]}\ ]$

b. Root lexicalization:
$[_{Y}\ [_{X}\ [_{c}\ \sqrt{\alpha}\ c\ ]\ X_{[f_2]}\ ]\ Y_{[f_1]}\ ]$

c. Vocabulary insertion on X cycle:
$[_{Y}\ [_{X}\ [_{c}\ \sqrt{\alpha}\ c\ ]\ [_{X}\ /b/\ ]]\ Y\ ]$

d. Feature copying on Y cycle *cannot apply*:
$[_{Y}\ [_{X}\ [_{c}\ \sqrt{\alpha}\ c\ ]\ [_{X}\ /b/\ ]]\ Y\ ]$

e. Vocabulary insertion into Y on Y cycle **Subset Principle*:
$[_{Y}\ [_{X}\ [_{c}\ \sqrt{\alpha}\ c\ ]\ [_{X}\ /b/\ ]]\ [_{Y}\ /a/\ ]]$

The only possible option for Y to be realized by /a/ would be to have feature copying *preceding* morphological realization of X by /b/, so that $[f_1]$ on Y can be used to generate a

[10] It has indeed been suggested that operations that manipulate morpho-syntactic features in post-syntactic representations, like impoverishment and, under present assumptions, enrichment, are always ordered before vocabulary insertion in any cyclic domain, due to the nature of the material that they affect; cf. Arregi & Nevins (2012, ch. 6).

[11] Note that this reasoning presupposes that an insertion operation affecting a lower head takes place in the cyclic domain defined by this head, even if, as a consequence of this insertion, a feature on a higher head is ultimately also discharged. This follows naturally if vocabulary insertion is viewed as a complex operation consisting of two separate suboperations, viz., (i) insertion under feature matching followed by (ii) feature discharge.

second $[f_1]$ before it is removed. However, as shown in derivation 5 in (22), this is impossible because of cyclicity. Feature copying takes place early in (22-c); but since this operation affects $[f_1]$ on Y, this operation takes place on the Y cycle. An alternative third step of the derivation would have been to carry out vocabulary insertion of /b/; as we have seen, this would have applied on the X cycle. Thus, the derivation has skipped X in (22-c), in violation of the Cyclic Principle, and the derivation crashes.

(22) *Scenario a→Y, derivation 5: **

a. Initial structure:
   $[_{Y}\ [_{X}\ [_{c}\ \sqrt{\ }\ c\ ]\ X_{[f_2]}\ ]\ Y_{[f_1]}\ ]$

b. Root lexicalization:
   $[_{Y}\ [_{X}\ [_{c}\ \sqrt{\alpha}\ c\ ]\ X_{[f_2]}\ ]\ Y_{[f_1]}\ ]$

c. Feature copying on Y cycle **Cyclic Principle*:
   $[_{Y}\ [_{X}\ [_{c}\ \sqrt{\alpha}\ c\ ]\ X_{[f_2]}\ ]\ Y_{[f_1],[f_1]}\ ]$

d. Vocabulary insertion into X on Y cycle:
   $[_{Y}\ [_{X}\ [_{c}\ \sqrt{\alpha}\ c\ ]\ [_{X}$ /b/]] $Y_{[f_1]}$]

e. Vocabulary insertion into Y on Y cycle:
   $[_{Y}\ [_{X}\ [_{c}\ \sqrt{\alpha}\ c\ ]\ [_{X}$ /b/]] $[_{Y}$ /a/]]

An interesting question arising at this point is whether subsequent vocabulary insertion of /b/ into X in (22-d) violates the Strict Cycle Condition. Here the exaxt wording of the constraint becomes relevant. On the one hand, recall the premise that vocabulary insertion into X *applies to* the cyclic domain X, even if as a consequence eventually some feature beyond X (i.e., in Y) is affected (i.e., discharged); so, for the purposes of the Cyclic Principle in (14), such insertion is an operation on the X cycle. On the other hand, for the Strict Cycle Condition in (15), the question is whether a vocabulary insertion into X *affects* the cyclic domain X or the cyclic domain Y if it ultimately gives rise to feature discharge in Y. If this is the case, the Strict Cycle Condition will not be violated by the step in (22-d); if insertion into X in the case at hand does not affect (in the technical sense of (15)) the cyclic domain of Y, the Strict Cycle Condition will be violated. Given that the first option is arguably the more plausible one, this means that we now have a further argument for keeping the two concepts of cyclicity apart: The Cyclic Principle can exclude some sequences of operations as counter-cyclic that may be compatible with the Strict Cycle Condition.

These considerations notwithstanding, it can be concluded that cyclicity plays a major role in deriving the Partially Superfluous Extended Exponence Generalization since it ensures that feature copying in a higher domain (which is required for the presence of the more general exponent in an outer position) cannot take place before vocabulary insertion in a lower domain, which may bleed it. As with the a→X scenario, there are further derivations to consider, but they are all ruled out for obvious reasons. For instance, any derivation in which /a/ is inserted before feature copying takes place will never give rise to extended exponence. More generally, then, the question of how the existence of partially superfluous extended exponence, and the Partially Superfluous Extended Exponence Generalization in (11), can be derived in Distributed Morphology has received an answer: The Cyclic Principle and the Subset Principle can only be both satisfied in derivation 1: If the more general (proper subset) exponent realizes a hierarchically lower feature, it must come first; if it realizes a hierarchically higher feature, there is no good output because feature copying is bled by cyclic vocabulary insertion.

3.4. *Overlapping Extended Exponence*

At this point, the question arises of how overlapping extended exponence can be accounted for under the present system of assumptions; as noted above, this type of extended exponence is much less of a challenge for many theories of moprhology. Indeed, overlapping extended exponence is also predicted to be possible in the approach under consideration – but, as we will see, there is a caveat.

Suppose that there are two exponents /a/ $\leftrightarrow$ $[f_1,f_2]$ and /b/ $\leftrightarrow$ $[f_1,f_3]$ that share a feature $[f_1]$, as in (6); and that there is a functional head X initially bearing the features $[f_1]$, $[f_2]$, and a functional head Y that is equipped with the feature $[f_3]$. In addition, the enrichment rule (17) is active in the language, as before. Under these assumptions, the derivation in (23) gives rise to overlapping exponence. Importantly, feature copying applying to $[f_1]$ on X must take place early, on the X cycle; see (23-c). After that, /a/ is inserted into X, thereby discharging $[f_2]$ and one copy of $[f_1]$ in X; see (23-d). Finally, /b/ is inserted into Y, which gives rise to a discharge of $[f_3]$ in Y, and of the other copy of $[f_1]$ in X; see (23-e).

(23) *Overlapping extended exponence, derivation 1:* √

a. Initial structure:
$[_{Y}\ [_{X}\ [_{c}\ \sqrt{\ }\ c\ ]\ X_{[f_1],[f_2]}\ ]\ Y_{[f_3]}\ ]$

b. Root lexicalization:
$[_{Y}\ [_{X}\ [_{c}\ \sqrt{\alpha}\ c\ ]\ X_{[f_1],[f_2]}\ ]\ Y_{[f_3]}\ ]$

c. Feature copying on X cycle:
$[_{Y}\ [_{X}\ [_{c}\ \sqrt{\alpha}\ c\ ]\ X_{[f_1],[f_1],[f_2]}\ ]\ Y_{[f_3]}\ ]$

d. Vocabulary insertion on X cycle:
$[_{Y}\ [_{X}\ [_{c}\ \sqrt{\alpha}\ c\ ]\ [_{X_{[f_1]}}$ /a/ $]\ Y_{[f_3]}\ ]$

e. Vocabulary insertion on Y cycle:
$[_{Y}\ [_{X}\ [_{c}\ \sqrt{\alpha}\ c\ ]\ [_{X}$ /a/ $]\ [_{Y}$ /b/ $]]$

Thus, overlapping extended exponence is predicted to be possible. However, as noted, there is a caveat. Overlapping extended exponence is in fact ceteris paribus predicted to be *impossible* if the shared feature is not on the lower head, as in (23), but on the higher head, as in the minimally different derivation 2 in (24). Here, the shared feature $[f_1]$ is not on X, but on Y; see (24-a). Since $[f_1]$ is not present on the X cycle, feature copying cannot apply here, and the derivation inserts /a/ on the X cycle, which removes both $[f_2]$ from X, which is unproblematic, and $[f_1]$ from Y, which is fatal because now there is no $[f_1]$ feature left that enrichment on the Y cycle could apply to; cf. (24-c). Subsequent insertion of /b/ on the Y cycle will therefore have to violate the Subset Principle; see (24-d).

(24) *Overlapping extended exponence, derivation 2:* *

a. Initial structure:
$[_{Y}\ [_{X}\ [_{c}\ \sqrt{\ }\ c\ ]\ X_{[f_2]}\ ]\ Y_{[f_1],[f_3]}\ ]$

b. Root lexicalization:
$[_{Y}\ [_{X}\ [_{c}\ \sqrt{\alpha}\ c\ ]\ X_{[f_2]}\ ]\ Y_{[f_1],[f_3]}\ ]$

c. Vocabulary insertion on X cycle:
$[_{Y}\ [_{X}\ [_{c}\ \sqrt{\alpha}\ c\ ]\ [_{X}$ /a/ $]\ Y_{[f_3]}\ ]$

d. Vocabulary insertion on Y cycle **Subset Principle*:
$[_{Y}\ [_{X}\ [_{c}\ \sqrt{\alpha}\ c\ ]\ [_{X}$ /a/ $]\ [_{Y}$ /b/ $]]$

As we have seen with the analogous issue for partially superfluous extended exponence, any attempt at solving the problem in (24) by applying feature copying to $[f_1]$ on Y earlier will invariably lead to a violation of the Cyclic Principle; see (25).

(25) *Overlapping extended exponence, derivation 3: **

a. Initial structure:
$[_Y\ [_X\ [_c \sqrt{\ }\ c\ ]\ X_{[f_2]}\ ]\ Y_{[f_1],[f_3]}\ ]$

b. Root lexicalization:
$[_Y\ [_X\ [_c \sqrt{\alpha}\ c\ ]\ X_{[f_2]}\ ]\ Y_{[f_1],[f_3]}\ ]$

c. Feature copying on Y cycle **Cyclic Principle*:
$[_Y\ [_X\ [_c \sqrt{\alpha}\ c\ ]\ X_{[f_2]}\ ]\ Y_{[f_1],[f_1],[f_3]}\ ]$

Thus, there is an unresolvable problem with a derivation where the shared feature in overlapping exponence is on the higher head: Insertion of a specific exponent discharges the shared feature before it can be copied for the other exponent in accordance with the Cyclic Principle. At present, we take it to be an open question whether this prediction might be called into question by empirical evidence; pursuing this issue in detail, based on the available empirical evidence, is beyond the scope of the present paper.

3.5. Fully Superfluous Extended Exponence

Fully superfluous exended exponence is predicted to be impossible: Either there is a problem with the compatibility (i.e., the Subset Principle), or there is a cyclicity problem. Thus, suppose that there are two exponents /a/ $\leftrightarrow$ [f_1,f_2] /b/ $\leftrightarrow$ [f_1,f_2] with identical feature specifications, as in (8); by assumption, the features [f_1] and [f_2] are located on two separate heads X and Y (if they were to show up on only one head, extended exponence would trivially be excluded). Given these assumptions, a schematic derivation illustrating the compatibility problem problem is given in (26). Feature copying applies early here (cf. (26-c)), but subsequent insertion of /a/ (or, for that matter, /b/) in X leads to discharge of [f_2] on Y, so that the remaining exponent can never be inserted in Y (cf. (26-d)).

(26) *Fully Superfluous extended exponence, derivation 1: **

a. Initial structure:
$[_Y\ [_X\ [_c \sqrt{\ }\ c\ ]\ X_{[f_1]}\ ]\ Y_{[f_2]}\ ]$

b. Root lexicalization:
$[_Y\ [_X\ [_c \sqrt{\alpha}\ c\ ]\ X_{[f_1]}\ ]\ Y_{[f_2]}\ ]$

c. Feature copying on X cycle:
$[_Y\ [_X\ [_c \sqrt{\alpha}\ c\ ]\ X_{[f_1],[f_1]}\ ]\ Y_{[f_2]}\ ]$

d. Vocabulary insertion on X cycle:
$[_Y\ [_X\ [_c \sqrt{\alpha}\ c\ ]\ [_{X_{[f_1]}}$ /a/]] Y]

e. Vocabulary insertion into Y on Y cycle **Subset Principle*:
$[_Y\ [_X\ [_c \sqrt{\alpha}\ c\ ]\ [_{X_{[f_1]}}$ /a/]] $[_Y$ /b/]]

The cyclicity problem arising with the alternative derivation where feature copying on the Y cycle precedes vocabulary insertion on the X cycle is shown in (27).

(27) *Fully Superfluous extended exponence, derivation 2: **

a. Initial structure:
$[_Y\ [_X\ [_c \sqrt{\ }\ c\ ]\ X_{[f_1]}\ ]\ Y_{[f_2]}\ ]$

b. Root lexicalization:
$[_Y\ [_X\ [_c \sqrt{\alpha}\ c\ ]\ X_{[f_1]}\ ]\ Y_{[f_2]}\ ]$

c. Feature copying on X cycle:
$[_Y\ [_X\ [_c \sqrt{\alpha}\ c\ ]\ X_{[f_1],[f_1]}\ ]\ Y_{[f_2]}\ ]$

d. Feature copying on Y cycle **Cyclic Principle*:
$[_{\mathrm{Y}}\ [_{\mathrm{X}}\ [_{\mathrm{c}}\ \sqrt{\alpha}\ \mathrm{c}\ ]\ \mathrm{X}_{[\mathrm{f}_1],[\mathrm{f}_1]}\ ]\ \mathrm{Y}_{[\mathrm{f}_2],[\mathrm{f}_2]}\ ]$

e. Vocabulary insertion into X on Y cycle:
$[_{\mathrm{Y}}\ [_{\mathrm{X}}\ [_{\mathrm{c}}\ \sqrt{\alpha}\ \mathrm{c}\ ]\ [_{\mathrm{X}_{[\mathrm{f}_1]}}\ /\mathrm{a}/\ ]]\ \mathrm{Y}_{[\mathrm{f}_2]}\ ]$

f. Vocabulary insertion into Y on Y cycle:
$[_{\mathrm{Y}}\ [_{\mathrm{X}}\ [_{\mathrm{c}}\ \sqrt{\alpha}\ \mathrm{c}\ ]\ [_{\mathrm{X}_{[\mathrm{f}_1]}}\ /\mathrm{a}/\ ]]\ [_{\mathrm{Y}}\ /\mathrm{b}/\ ]]$

3.6. Convergence

Arguably, from a more general point of view, the present approach to partially superfluous extended exponence based on Distributed Morphology captures the same underlying core idea as the approaches based on derivational versions of Optimality Theory (strata, f-seq, harmonic serialism) that were discussed in section 2: In the present approach, the more general exponent needs to find a matching feature in the syntactic head into which it is supposed to be inserted (because of the Subset Principle), and the interaction of copying and insertion governed by cyclicity ensures that there will not be such a feature if the more general exponent comes too late. So, all these approaches share the common core that the more general exponent can only show up if it shows up early; once the more specific exponent is part of the structure, it will block the more general exponent.[12]

4. Empirical Issues

4.1. Exceptions

There are exceptions to the Partially Superfluous Extended Exponence Generalization, i.e., cases where the more general (proper subset) exponent can or must show up further away from the stem than the more specific (superset) exponent. To name just a few examples: There is partially superfluous extended exponence of negation in past contexts in Swahili (see Stump (2001)), in forms like *ha-tu-ku-taka* (NEG-1.PL-NEG.PAST-want; 'We did not want'), where the pure, general negative prefix *ha* shows up outside of the more specific negative past prefix *ku*; there is also partially superfluous extended exponence of third person in plural contexts in Ojibwe (see Oxford (2019)), in forms like *waapam-ikw-waa-t-pan* (see-INV-

[12] An issue that the preceding approach has so far remained silent on is extended exponence involving roots and their categorizing heads; such phenomena are usually discussed under rubrics like *root suppletion* or *stem allomorphy*. Thus, suppose that the combination $\sqrt{\ }$-c of an abstract root morpheme $\sqrt{\ }$ and a categorizing head c (which, to simplify matters, we will treat as a primitive item here) is equipped with a feature $[\mathrm{f}_1]$, and the next higher functional head X has the feature $[\mathrm{f}_2]$. Now, if the actual root vocabulary item α is characterized as $[\mathrm{f}_1,\mathrm{f}_2]$ (such that α is expected to give rise to $[\mathrm{f}_2]$-conditioned suppletion), then there will be a problem because α-insertion will discharge $[\mathrm{f}_2]$ on X, and a vocabulary item realizing X will not be insertable anymore; furthermore, this consequence is independent of whether we are dealing with partially superfluous extended exponence (as in this scenario) or overlapping extended exponence (if the vocabulary item for X also has some other feature). (Thanks to Elango Kumaran for noticing this.)
There are at least two ways to address this issue under present assumptions. One option would be to assume that the $\sqrt{\ }$-c stem is special in that it does not in fact qualify as a cyclic domain, in contrast to what we have assumed so far. As a consequence, $[\mathrm{f}_2]$ on X can be copied before α is inserted in the root position. Another option would be to stipulate that the insertion of root vocabulary items is special in that it does not lead to feature discharge. The two options differ with respect to the predictions for *non-local stem allomorphy*, where α bears a feature $[\mathrm{f}_3]$ that is located on some yet higher head Y (either instead of $[\mathrm{f}_2]$, or in addition to $[\mathrm{f}_2]$): In the first approach, more must be said to permit such non-local extended exponence; in the second approach, it can be derived without problems (as long as Y is still part of the same morphological phase). In view of the fact that non-local stem allomorphy appears to be a marked phenomenon in the world's languages, and requires additional assumptions (and, often, additional tools, like spanning, hyper-contextual realization rules or buffers) in all existing derivational approaches to morphology (see Merchant (2015), Moskal & Smith (2016), Weisser (2017), Kastner & Moskal (2018), and Božič (2019)), we will refrain from deciding this question.

3.PL-3-PRET; 'The other saw them'), where the more specific third person plural exponent *waa* is closer to the root (*waapam*) than the more general bare plural exponent *t*; and several more of such examples can be found in the literature. Clearly, if these apparent exceptions are taken at face value, this would imply that there is no interesting generalization to be made about partially superfluous extended exponence after all. What is more, from the perspective of grammatical theory, there would be no principled answer left to the question of why the phenomenon exists in the first place (since all available approaches that have something to say about this question derive the Partially Superfluous Extended Exponence Generalization).

In view of this state of affairs, it seems to us that the most promising strategy is to account for apparent exceptions to the Partially Superfluous Extended Exponence Generalization in a way that leaves the generalization (and its explanation) intact. Accordingly, we would like to contend that if the more general exponent shows up outside of the more specific exponent in a partially superfluous extended exponence scenario, this is either (i) due to exponent movement (either in the morphological component or in the phonological component), or can (ii) be shown to be compatible with the generalization after all, due to a reanalysis of the data.

As for exponent movement (i), the assumption is that the more general exponent is first inserted into the word in a position that is closer to the root than the position of the more specific exponent but subsequently moves to a position outside the domain of this latter marker. Note that this is in all relevant respects identical to what happens in syntactic derivations; for instance, an object must be base-generated closer to the verb than the subject (cf. *Mary often [reads books]*), but may, as a consequence of movement, eventually come to be placed outside of the domain of the subject at the end of the derivation (cf. *What does [Mary read]?*). In Müller (2020) and Gleim et al. (2021; 2022), arguments are presented for the existence of word-internal movement of morphological exponents that is triggered by alignment constraints; in line with this, an example like Swahili *ha-tu-ku-taka* is argued to involve morphological movement of *ha*, triggered by a morphological alignment constraint that requires left-alignment of negative items in a word (whereas a higher-ranked alignment constraint on all tense exponents forces *ku* to stay in situ). An alternative to morphological movement of exponents is phonological movement of exponents, which is triggered by purely phonological requirements (this is analogous to the concept of PF movement in syntactic derivations; cf., e.g., Chomsky (1995), Truckenbrodt (1995), Agbayani et al. (2015)).

As for reanalyses of the data (ii), we suggest that closer inspection of apparent exceptions to the Partially Superfluous Extended Exponence Generalization will often reveal that there is in fact no extended exponence to begin with (in the sense that a single morphosyntactic feature justified for syntactic reasons is realized by two or more exponents in the morphology). There are various possibilities as to how such a configuration can come about. One possibility is that there are two independently motivated occurrences of the same feature in a given word from the start (i.e., as a consequence of what happens in the syntax), which are then separately targetted by morphological realization without requiring enrichment (see, e.g., Sells (2004) and Alexiadou et al. (2021) for relevant discussion, also with respect to larger grammatical units). Another possibility is that an independently motivated decomposition of seemingly primitive morpho-syntactic features provides more targets for morphological realization by exponents – if, say, a feature [±A] that, at first sight, seems to be realized by two exponents α and β, is to be decomposed into a combination [±b,±c], it may be the case that α realizes only [±b], and β only [±c] (cf. Stiebels (2016) and Caha (2021) for analyses along these lines).

In what follows, we will discuss two relevant cases in a bit more detail, viz., phonologically triggered movement of more general morphological exponents in Huave (see Kim (2010)), and the distribution of ϕ-features and case features on class markers in Itelmen (see Bobaljik (2000)).

4.2. *Partially Superfluous Extended Exponence and Phonological Exponent Movement in Huave*

As noted in Grofulović et al. (2021), there are patterns of partially superfluous extended exponence in San Francisco del Mar Huave that seem to contradict the order predicted by the Partially Superfluous Extended Exponence Generalization in (11). The inflected verb in (28) exhibits partially superfluous extended exponence: *s* (here realized as [ʃ]) is a general first person marker, and *n* is a more specific exponent realizing first person and subordination (SB). Since both exponents are prefixes of the transitive verb *a-ʰʧ*, the Partially Superfluous Extended Exponence Generalization is contradicted on the surface.

(28) *ʃ*-i-*n*-a-ʰʧ
1-FT-1SB-TV-give
'I will give'

However, there is evidence that the positions of *s* and *n* are motivated by phonological requirements, and do not necessarily represent the positions occupied by the exponents when vocabulary insertion takes place in the morphology. The crucial observation is that San Francisco del Mar Huave exhibits the phenomenon of so-called mobile affixation (see Kim (2010), Zukoff (2021)): Depending on phonological constraints, one and the same morphological exponent may show up in different positions in a word. The phenomenon of mobile affixation in San Francisco del Mar Huave is illustrated in (29) (cf. Kim (2010)).[13]

(29) a. *t*-a-ʰʧ-ju-s
CP-TV-give-1
'I gave'

b. paʰk-a-*t*-u-s
face.up-V-CP-ITR-1
'I lay face up'

In San Francisco del Mar Huave, the completive aspect (CP) exponent *t* is one of several exponents that are 'mobile' in the sense that they can show up either as a suffix (as in (29-b)) or as a prefix (as in (29-a)). The placement is regulated by the phonotactic constraints of the language. If the CP exponent *t* occurs with a verbal base starting with a consonant and ending in a vowel, it is realized as a suffix; however, if the verbal base starts with a vowel and ends in a consonant, the consonant cluster that would result if *t* were to be realized as a suffix, and the ensuing vowel epenthesis that would take place as a repair, are avoided by realizing the exponent as a prefix instead.

Returning to the problematic case of partially superfluous extended exponence in (28), it can be noted that the general first person marker *s* and the more specific person/subordination marker *n* are both mobile affixes; e.g., (30) illustrates that *n* shows up as a suffix under the right phonotactic conditions.

(30) *ʃ*-i-ʧut-u-*n*
1-FT-sit-V-1SB
'I will sit'

Approaches to mobile affixation (like Kim (2010) and Zukoff (2021)) typically do not pos-

[13] The examples are rendered in IPA, based on the convention in Zukoff (2021) and the glossing rules given in Kim (2010).

tulate that movement of any type is actually involved; rather, the assumption normally is that the exponents in question are directly placed in a word according to the demands of syllable structure constraints, with morphology and phonology intermingled. However, as argued in Grofulović et al. (2021), the available evidence is fully compatible with an approach that respects modularity by separating vocabulary insertion in underlying morphology from the phonologically conditioned placement of an exponent in actual output forms (see also Kalin & Rolle (2021) for independent arguments for such an approach). On this view, the two exponents in a partially superfluous extended exponence relation in (28) and (30) are first inserted in the order required by the Partially Superfluous Extended Exponence Generalization, with the more general exponent coming first and the more specific one coming second; and subsequently, phonologically driven movement takes place, leading to an opaque surface representation.

4.3. Partially Superfluous Extended Exponence Reanalyzed in Itelmen

As a second case study, let us look at the distribution of ϕ-features and case features on agreement exponents and inflection class markers in Itelmen, which Bobaljik (2000) takes to provide the core empirical evidence in support of his model of contextual allomorphy discussed in section 3.1 above. There are two patterns in Itelmen verb inflection that initially would seem to support the view that there is outwards-sensitivity to morpho-syntactic features on functional heads, but no inwards-sensitivity. From the present perspective, this implies that the more specific exponent is closer to the root than the more general exponent in a partially superfluous extended exponence scenario; consequently, the patterns in question pose potential challenges for the generalization in (11) and, more specifically, the present account of the generalization based on how the interaction of vocabulary insertion and feature copying (enrichment) is governed by cyclicity.

Following Bobaljik (2000), a simplified structure of Itelmen verbs containing the relevant functional morphemes in need of morphological realization is given in (31): AgrS is realized as a prefix, and Class and AgrO are realized by suffixal exponents.

(31) $[_{\text{AgrS}}$ AgrS $[_{\text{AgrO}}$ $[_{\text{Class}}$ V Class] AgrO]]

Two instances of (what at first sight looks like) partially superfluous extended exponence can be observed. First, as illustrated in (32-a), an object (accusative) agreement exponent in AgrO (here: *čeʔn*) can realize features of both the subject (nominative) and the object (or, in Bobaljik's terms, an object agreement exponent can be conditioned by subject agreement features); in addition, a more general subject (nominative) exponent shows up in AgrS. Second, in (32-b), the suffixal inflection class exponent *ki* is specified for ϕ-features of both the nominative and the accusative argument but clearly shows up closer to the root than at least the AgrO exponent.

(32) *Extended exponence of nominative/accusative and case/class exponents in Itelmen*:

a. $[_{\text{AgrS}}$ *t′* $[_{\text{AgrO}}$ $[_{\text{Class}}$ $[_{\text{V}}$ əŋkzu-s] Ø] *čeʔn*]]
1.SG.NOM help-PRES CL.I 1.NOM./3.PL.ACC
'I'm helping them.'

b. $[_{\text{AgrS}}$ *t* $[_{\text{AgrO}}$ $[_{\text{Class}}$ $[_{\text{V}}$ tϕ-s] *ki*]
1.SG.NOM bring-PRES CL.II.1.SG.NOM/3.PL.ACC
čeʔn]]
1.NOM./3.PL.ACC
'I'm bringing them.'

Given (32-a) and (32-b), it looks as though Itelmen exhibits two patterns where a more specific exponent shows up in a position that is closer to the stem than the position occupied by the more general exponent in a partially superfluous extended exponence configuration.[14] Assuming this to be the case, a potential problem arises for the generalization in (11) and the present analysis.

Turning to the subject and object agreement exponents in (32-a) first, it can be noted that the specific analysis that Bobaljik (2000), following Bobaljik & Wurmbrand (1997), suggests for examples like (32-a) does not actually instantiate extended exponence in the sense adopted througout this paper, as *one* morpho-syntactic feature resulting in realization by *two* (or more) exponents. More specifically, here are Bobaljik & Wurmbrand's (1997) assumptions about the AgrO slot: First, in third person object environments, there are no person features in AgrO. Second, there is an EPP-like requirement in Itelmen to have person features in this position. Third, to satisfy this requirement, Itelmen employs a general Agree-like copying mechanism where *all* the features of AgrS are copied onto AgrO (i.e., not just the missing person information, but the whole feature bundle). Fourth, Agree-like copying of subject features also takes place in intransitive environments. As a consequence of these assumptions, the ϕ-features and case features of subjects are available twice in the Itelmen verb in the relevant contexts, on AgrS and on AgrO; and assuming that the Agree-like copying operation is either syntactic or, at least, takes place very early in the post-syntactic component (a view which is supported by the fact that the operation is non-local and transfers feature bundles corresponding to entire categories), there will be no cyclicity restrictions of the type addressed in section 3 above. Thus, AgrS in (32) is realized by a subject agreement exponent, and AgrO in (32) is realized by a *portmanteau exponent*. Consequently, if this analysis is adopted, there is no problem with either the Partially Superfluous Extended Exponence Generalization or its derivation via cyclic application of feature copying and vocabulary insertion: There is no extended exponence, just faithful realization of two separate feature sets that are independently present in the structure.

Turning to the interaction of the subject and object agreement exponents with the inflection class exponents in (32-b) next, the situation is a bit more complex; but the overall conclusion will be the same: It is likely that there is no partially superfluous extended exponence involved.

For concreteness, there is evidence that sheds doubt on the existence of a separate functional morpheme hosting class exponents. Bobaljik (2000) observes that there are a number of class II markers: Next to *ki* in (32-b), there is *k*, there is *čiŋ*, there is *xk*, etc.; the choice among these is mainly determined by the ϕ-features of subject and object. These inflection class exponents always show up immediately adjacent to the AgrO exponent, not adjacent to V (tense exponents, e.g., intervene between V and the alleged inflection class exponent, as in (32-b)), which may already be regarded as somewhat suspicious. Accordingly, Georg & Volodin (1999) analyze strings like *ki-čeʔn* as *primitive*, non-decomposable AgrO exponents without any internal fine structure: *kičeʔn*. Interestingly, exactly the same string *kičeʔn*

[14] At least, this corresponds to Bobaljik's (2000) conclusion, based on the premise that the structure of inflected verbs in Itelmen looks as in (31), and that his assumptions about contextual allomorphy hold. Strictly speaking, however, the data in (32) only unequivocally show that more specific Class exponents are closer to the root than more general AgrO exponents. Bobaljik notes that independent evidence for the height of prefixes in an Itelmen verb is hard to come by; but Bobaljik & Wurmbrand (2001) provide a couple of arguments for the view that subject agreement prefixes are higher than object agreement prefixes in the language. At least for the sake of the argument, we will follow Bobaljik in assuming that not only are more specific Class exponents closer to the root than the AgrO exponents that they are "conditioned by", but more specific AgrO and Class exponents are also closer to the root than the AgrS prefix exponents they are "conditioned by".

also shows up in the other, unmarked inflection class I *in intransitive contexts*; and in this context, Bobaljik (2000) also assumes that *kičeʔn* is indeed a primitive, non-decomposable AgrO marker. Thus, in Bobaljik's (2000) system, *kičeʔn* is viewed as a concatenation of two morphological exponents in one environment, and as a single morphological exponent in another, closely related environment. This looks like a generalization is being missed.

In view of all this, we would like to suggest that the functional morpheme AgrO in Itelmen is subject to *fission*, in the sense of Noyer (1997) and Trommer (1999): Vocabulary insertion discharges features in a fissioned morpheme, but the remaining features can trigger a new vocabulary insertion operation affecting the same functional morpheme.[15] At this point, the question arises of how an analysis of strings like *kičeʔn* that is based on subanalysis of exponents, i.e., fission, can avoid potential problems for the present account. There are at least two possible answers.

First, given fission, the various AgrO exponents are not hierarchically distinct; they are all inserted into one and the same functional morpheme, and are therefore all part of one and the same morphological cycle: [$_{\text{AgrO}}$ *ki–če–ʔn*], not *[[[*ki*]–*če*]–ʔn]. Therefore, features can be copied early, before they are discharged by vocabulary insertion. On this view, there is extended exponence (of the partially superfluous or overlapping type), but it is entirely unproblematic: Given the absence of discriminating structure in a fissioned morpheme, the feature copying that is required for extended exponence is not required by cyclicity to come too late to feed exponence.

Second, a closer analysis of the morphological system reveals that the feature sets realized by the class marker and the object agreement exponent could in fact emerge as *complementary*: It looks as though it might be possible to maintain the view that the class exponent realizes inflection class and subject agreement features in AgrO, and the object agreement exponent realizes only object agreement features in AgrO. Under such an analysis, there would be no extended exponence in the system. For reasons of space and coherence, we will not try to advance a full-fledged analysis of the whole paradigm of verb inflection in Itelmen here, and decide for one of the two options; suffice it to say that on either view, the problem that the Itelmen data in (31) might initially pose for the Partially Superfluous Extended Exponence Generalization (and its derivation based on cyclicity) disappears.

Acknowledgements

[to be added]

References

Adger, David, Susana Béjar & David Harbour (2003): Directionality of Allomorphy: A reply to Carstairs-McCarthy, *Transactions of the Philological Society* 101(1), 109–115.

Agbayani, Brian, Chris Golston & Toru Ishii (2015): Syntactic and Prosodic Scrambling in Japanese, *Natural Language and Linguistic Theory* 33, 47–77.

Alexiadou, Artemis, Imke Driemel, Johannes Heinand Fabienne Martin & Yining Nie (2021): Multiple Exponence in Child Language. Ms., ZAS Berlin.

Alhoniemi, Alho (1993): *Grammatik des Tscheremissischen (Mari)*. Helmut Buske Verlag, Hamburg.

[15] This implies that the *disjunctive blocking* assumption from section 3.2 must be qualified for fissioned morphemes. Independent evidence for fission of AgrO heads in Itelmen comes from the distribution of *partial syncretism*; there are several such cases in the paradigm; cf., e.g., *če-n* (1.NOM./3.SG.ACC) vs. *če-ʔn* (1.NOM./3.PL.ACC).

Arregi, Karlos & Andrew Nevins (2012): *Morphotactics: Basque Auxiliaries and the Structure of Spellout*. Springer, Heidelberg.

Bermúdez-Otero, Ricardo (2011): Cyclicity. In: M. van Oostendorp, C. Ewen, E. Hume & K. Rice, eds., *The Blackwell Companion to Phonology*. Vol. 4, Wiley-Blackwell, Malden, MA, pp. 2019–2048.

Bobaljik, Jonathan (2000): The Ins and Outs of Contextual Allomorphy. In: K. Grohmann & C. Struijke, eds., *University of Maryland Working Papers in Linguistics*. Vol. 10, University of Maryland, College Park, pp. 35–71.

Bobaljik, Jonathan & Susi Wurmbrand (1997): Preliminary Notes on Agreement in Itelmen. In: B. Bruening, Y. Kang & M. McGinnis, eds., *PF: Papers at the Interface*. MIT Workinpg Papers in Linguistics, MIT, Cambridge, Mass., pp. 395–423.

Bobaljik, Jonathan & Susi Wurmbrand (2001): Seven Prefix-Suffix Asymmetries in Itel′men. In: M. Andronis, C. Ball, H. Elston & S. Neuvel, eds., *Proceedings from the Panels of the Chicago Linguistic Society's Thirty-Seventh Meeting*. Chicago Linguistic Society, Chicago, pp. 205–219.

Božič, Jurij (2019): Constraining Long-Distance Allomorphy, *The Linguistic Review* 36, 485–505.

Brown, Dunstan & Andrew Hippisley (2012): *Network Morphology*. Cambridge University Press, Cambridge.

Caballero, Gabriela & Alice Harris (2012): A Working Typology of Multiple Exponence. In: F. Kiefer et al., eds., *Current Issues in Morphological Theory: (Ir)Regularity, Analogy, and Frequency*. Benjamins, Amsterdam, pp. 163–188.

Caballero, Gabriela & Sharon Inkelas (2013): Word Construction: Tracing an Optimal Path Through the Lexicon, *Morphology* 23, 103–143.

Caha, Pavel (2021): No Multiple Exponence. Ms., Masarykova Univerzita.

Chomsky, Noam (1973): Conditions on Transformations. In: S. Anderson & P. Kiparsky, eds., *A Festschrift for Morris Halle*. Academic Press, New York, pp. 232–286.

Chomsky, Noam (1995): *The Minimalist Program*. MIT Press, Cambridge, Mass.

Chomsky, Noam (2001): Derivation by Phase. In: M. Kenstowicz, ed., *Ken Hale. A Life in Language*. MIT Press, Cambridge, Mass., pp. 1–52.

Chomsky, Noam (2015): Problems of Projection: Extensions. In: E. Di Domenico, C. Hamann & S. Matteini, eds., *Structures, Strategies and Beyond: Studies in Honour of Adriana Belletti*. Benjamins, Amsterdam, pp. 1–16.

Dahlstrom, Amy (2000): Morphosyntactic Mismatches in Algonquian: Affixal Predicates and Discontinuous Verbs. In: A. Okrent & J. Boyle, eds., *Proceedings from the Panels of the Chicago Linguistic Society's Thirty-Sixth Meeting*. Chicago Linguistic Society, Chicago, pp. 63–87.

Don, Jan & Elma Blom (2006): A Constraint-Based Approach to Morphological Neutralization. In: *Linguistics in the Netherlands 2006*. Benjamins, Amsterdam, pp. 78–88.

Driemel, Imke (2018): Person Mismatch Agreement. In: S. Hucklebridge & M. Nelson, eds., *Proceedings of NELS 48*. GLSA, Amherst, Mass.

Embick, David (2010): *Localism vs. Globalism in Morphology and Phonology*. MIT Press, Cambridge, Mass.

Embick, David & Rolf Noyer (2007): Distributed Morphology and the Syntax/Morphology Interface. In: *The Oxford Handbook of Linguistic Interfaces*. Oxford University Press, Oxford, pp. 289–324.

Fenger, Paula & Philipp Weisser (2023): What Umlaut Tells Us about the Underlying Morphological Structure of Verbs in Sinhala. Ms., Universität Leipzig. To appear in Proceedings of (F)ASAL.

Georg, Stefan & Alexander Volodin (1999): *Die itelmenische Sprache*. Harrasowitz, Wiesbaden.
Gleim, Daniel, Gereon Müller, Mariia Privizentseva & Sören Tebay (2021): Non-Local Phonology by Morphological Movement. In: A. Farinella & A. Hill, eds., *Proceedings of NELS 51*. University of Massachusetts, Amherst.
Gleim, Daniel, Gereon Müller, Mariia Privizentseva & Sören Tebay (2022): Reflexes of Exponent Movement in Inflectional Morphology: A Study in Harmonic Serialism. *Natural Language and Linguistic Theory* (Online first.).
Grimshaw, Jane (2001): Optimal Clitic Positions and the Lexicon in Romance Clitic Systems. In: G. Legendre, J. Grimshaw & S. Vikner, eds., *Optimality-Theoretic Syntax*. MIT Press, Cambridge, Mass., pp. 205–240.
Grofulović, Jelena, Gereon Müller & Sam Zukoff (2021): Mobile Affixes Undergo Movement: An Argument from Extended Exponence. Ms., Universität Leipzig and USC.
Halle, Morris (1997): Distributed Morphology: Impoverishment and Fission. In: B. Bruening, Y. Kang & M. McGinnis, eds., *Papers at the Interface*. Vol. 30, MITWPL, pp. 425–449.
Halle, Morris & Alec Marantz (1993): Distributed Morphology and the Pieces of Inflection. In: K. Hale & S. J. Keyser, eds., *The View from Building 20*. MIT Press, Cambridge, Mass., pp. 111–176.
Hanink, Emily (2018): Super Light-Headed Relatives, Missing Prepositions, and Span-Conditioned Allomorphy in German, *The Journal of Comparative Germanic Linguistics* 21, 247–290.
Hein, Johannes (2008): Verbflexion im Warembori: Eine Analyse im Rahmen der Distribuierten Morphologie. In: F. Heck, G. Müller & J. Trommer, eds., *Varieties of Competition*. Vol. 87 of *Linguistische Arbeits Berichte*, Universität Leipzig, pp. 49–63.
Inkelas, Sharon, Teresa McFarland & Anne Pycha (2006): The Flip Side of Blocking: Multiple Exponence in an Optimization Framework. Ms., University of California, Berkeley.
Kalin, Laura & Nicholas Rolle (2021): Deconstructing Subcategorization: Conditions on Insertion versus Conditions on Position. Ms., Princeton University and ZAS Berlin. To appear in *Linguistic Inquiry*.
Kalin, Laura & Philipp Weisser (2021): Minimalism, Morphology, and the Lexicon: Then and Now. Ms., Princeton University and Universität Leipzig.
Kastner, Itamar & Beata Moskal, eds. (2018): *Non-Local Contextual Allomorphy*. Vol. 34 of *Snippets*, Ledonline. Special Issue.
Kim, Yuni (2010): Phonological and Morphological Conditions on Affix Order in Huave, *Morphology* 20, 133–163.
Kiparsky, Paul (1973): Abstractness, Opacity and Global Rules. In: O. Fujimura, ed., *Three Dimensions in Linguistic Theory*. TEC, Tokyo, pp. 57–86.
Kiparsky, Paul (1982): Lexical Morphology and Phonology. In: I.-S. Yang, ed., *Linguistics in the Morning Calm*. Hanshin Publishing Company, Seoul, pp. 3–91.
Marvin, Tatjana (2002): Topics in the Syntax and Stress of Words. PhD thesis, MIT, Cambridge, Mass.
Matthews, Peter (1972): *Inflectional Morphology: A Theoretical Study Based on Aspects of Latin Verb Conjugation*. Cambridge University Press, Cambridge.
Matthews, Peter (1974): *Morphology*. Cambridge University Press, Cambridge.
McCarthy, John (2016): The Theory and Practice of Harmonic Serialism. In: J. McCarthy & J. Pater, eds., *Harmonic Grammar & Harmonic Serialism*. Equinox, Sheffield, pp. 47–87.
McFarland, Teresa (2009): The Verbal Morphology of Totonaco de Filomeno Mata. PhD thesis, University of California, Berkeley.

Merchant, Jason (2015): How Much Context is Enough? Two Cases of Span-Conditioned Allomorphy, *Linguistic Inquiry* 46, 273–303.

Moskal, Beata & Peter Smith (2016): Towards a Theory Without Adjacency: Hyper-Contextual VI-Rules, *Morphology* 26, 295–312.

Müller, Gereon (2007): Extended Exponence by Enrichment. Argument Encoding in German, Archi, and Timucua. In: T. Scheffler, J. Tauberer, A. Eilam & L. Mayol, eds., *Proceedings of the 30th Annual Penn Linguistics Colloquium*. Vol. 13.1 of *Penn Working Papers in Linguistics*, University of Pennsylvania, Philadelphia, pp. 253–266.

Müller, Gereon (2020): *Inflectional Morphology in Harmonic Serialism*. Advances in Optimality Theory, Equinox, Sheffield. http://home.uni-leipzig.de/muellerg/harm/imhs.pdf.

Müller, Gereon (2023): Challenges for Cyclicity. In: F. Andermann, G. Müller & M. Privizentseva, eds., *Cyclicity*. Vol. 95 of *Linguistische Arbeits Berichte*, Institut für Linguistik, Universität Leipzig, pp. 1–32.

Norris, Mark (2014): A Theory of Nominal Concord. PhD thesis, UCSC.

Noyer, Rolf (1997): *Features, Positions and Affixes in Autonomous Morphological Structure*. Garland Publishing, New York.

Ortmann, Albert (2004): A Factorial Typology of Number Marking. In: G. Müller, L. Gunkel & G. Zifonun, eds., *Explorations in Nominal Inflection*. Mouton de Gruyter, Berlin, pp. 229–267.

Oxford, Will (2019): Fission in Algonquian and the Status of Morphological Templates. In: D. Reisinger & R. Y.-H. Lo, eds., *Proceedings of the Workshop on the Structure and Constituency of Languages of the Americas*. UBCWPL, Vancouver, BC.

Perlmutter, David & Scott Soames (1979): *Syntactic Argumentation and the Structure of English*. The University of California Press, Berkeley.

Pesetsky, David & Esther Torrego (2001): T-to-C Movement: Causes and Consequences. In: M. Kenstowicz, ed., *Ken Hale. A Life in Language*. MIT Press, Cambridge, Mass., pp. 355–426.

Preminger, Omer (2018): How to Tell a Syntactic Phenomenon When You See It. Ms., University of Maryland.

Prince, Alan & Paul Smolensky (2004): *Optimality Theory. Constraint Interaction in Generative Grammar*. Blackwell, Oxford.

Richards, Norvin (2001): *Movement in Language*. Oxford University Press, Oxford.

Sells, Peter (2004): Syntactic Information and Its Morphological Expression. In: L. Sadler & A. Spencer, eds., *Projecting Morphology*. CSLI Publications, Stanford University, pp. 187–225.

Starke, Michal (2001): Move Dissolves Into Merge. PhD thesis, University of Geneva.

Stiebels, Barbara (2006): Agent Focus in Mayan Languages, *Natural Language and Linguistic Theory* 24, 501–570.

Stiebels, Barbara (2015): Multiple Exponence: Typology, Triggering Factors, and Theoretical Modelling. Ms., Universität Leipzig.

Stiebels, Barbara (2016): Multiple Exponence in a Lexical-Incremental Framework. Ms., Universität Leipzig.

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

Trommer, Jochen (1999): Morphology Consuming Syntax' Resources. In: *Proceedings of the ESSLI Workshop on Resource Logics and Minimalist Grammars*. University of Nijmegen, pp. 37–55.

Trommer, Jochen (2001): Distributed Optimality. PhD thesis, Universität Potsdam.

Trommer, Jochen (2011): Phonological Aspects of Western Nilotic Mutation Morphology. Habilitation thesis, Universität Leipzig.

Truckenbrodt, Hubert (1995): Extraposition from NP and Prosodic Structure. In: J. Beckman, ed., *NELS 25*. University of Massachusetts, GLSA, Amherst, pp. 503–517.

Weisser, Philipp (2017): Allomorphy Tracks Selection. Ms., Universität Leipzig.

Wiese, Bernd (2000): Warum Flexionsklassen? Über die deutsche Substantivdeklination. In: R. Thieroff, M. Tamrat, N. Fuhrhop & O. Teuber, eds., *Deutsche Grammatik in Theorie und Praxis*. Niemeyer, Tübingen, pp. 139–153.

Wunderlich, Dieter (1997): A Minimalist Model of Inflectional Morphology. In: C. Wilder, H.-M. Gärtner & M. Bierwisch, eds., *The Role of Economy Principles in Linguistic Theory*. Akademie Verlag, Berlin, pp. 267–298.

Zukoff, Sam (2021): A Parallel Approach to Mobile Affixation in Huave. In: R. Bennett, R. Bibbs, M. L. Brinkerhoff, M. Kaplan, S. Rich, N. Van Handel & M. Wax Cavallaro, eds., *Supplemental Proceedings of AMP 2020*. LSA, Washington, DC.