

*ABA and The Combinatorics of Morphological Features

Jonathan David Bobaljik*
*University of Connecticut &
Leibniz-Zentrum Allgemeine
Sprachwissenschaft (ZAS)*
Storrs, CT, USA & Berlin, Germany

Uli Sauerland†‡
*Leibniz-Zentrum Allgemeine
Sprachwissenschaft (ZAS)*
Berlin, Germany

Revised version, August 2017

Abstract In several three cell paradigms, it has been observed that one logically conceivable pattern – ABA under some arrangement of cells – is unattested. Existing approaches assume that such *ABA generalizations provide evidence for feature inventories which are restricted to features that stand in containment relations, and are thus subject to Pāṇinian rule order. We present a novel approach to *ABA generalizations that derives from general properties of feature-based morphology. To this end, we develop a formal account of the widespread view that morphological paradigms derive from rules that relate abstract features from an inventory to morphological exponents. We demonstrate that the feature-based view restricts the space of typological patterns even without any further assumptions. We show furthermore that the feature-based theory derives *ABA as a special case of a broader class of generalizations if the number of features in the inventory must be minimal, and that these generalizations arise under a variety of general assumptions about feature-algebras (extrinsically ordered or Pāṇinian and with or without feature intersection). We discuss which explanation might be correct for actual cases of *ABA constraints, and we explore the consequences of the feature-based general approach for a range of paradigm sizes including those with more than three cells.

Keywords: features, morphology, combinatorics, syncretism, typology

Todo list

* jonathan.bobaljik@uconn.edu

† uli@alum.mit.edu

‡ Authors are listed alphabetically – both authors are equally first authors.

1 Introduction

One of the most interesting and difficult questions in research on language lies in formally characterizing the class of possible grammars. One aspect of this challenge asks whether there are constraints on grammars of a general, abstract nature, and in turn, whether these constraints are specific to language or instantiations of even broader, domain-general constraints on cognitive systems, with manifestations observable elsewhere. For example, some progress has been made in syntax on the basis of *Formal Language Theory* and the Chomsky Hierarchy (Chomsky 1956) for the analysis of sets of string sequences. We aim to contribute to the development of a similarly general perspective for morphology, particularly with respect to morphological features, i.e. the features that underlie the variation in how different concepts are grouped across languages as evidenced by exponence by the same form (syncretism). The architecture of feature-based morphological systems predicts that only certain patterns of variation are possible. In this paper, we address *ABA generalizations from this perspective. We show that a class of *ABA-type generalizations can be derived from the feature-based architecture in conjunction with a minimality assumption. We furthermore argue that such a derivation may be plausible for some cases of an *ABA generalization, but not for others.

The term **ABA generalization* refers to morphological patterns in which, given some arrangement of the relevant forms in a structured sequence, the first and third may share some property “A” only if the middle member shares that property as well. If the middle member is distinct from the first, then the third member of the sequence must also be distinct. Bobaljik (2012) demonstrates that a *ABA generalization holds for adjectival suppletion in the sequence positive-comparative-superlative: across a large cross-linguistic sample, one finds ABB patterns such as *good-better-best*, where the comparative and superlative share a root *be(t)*- distinct from the positive, but what is not found is an ABA pattern: **good-better-goodest*, in which the positive and superlative share a root, distinct from the comparative. Similar *ABA effects have been noted in extensive studies of case syncretism (Caha 2009), suppletion for both case and number in pronouns (Smith et al. 2016), Germanic verbs and participles (see Wiese 2008 on German, and class material cited by Starke 2009 on English), and in other domains.

In one way or another, almost all existing accounts of these generalizations have argued that the *ABA effect arises as a result of nesting or containment relations among features, along with the assumption that linguistic rules are arranged such that a more specific rule takes precedence over (bleeds) a more

general one, the so-called Elsewhere or Pāṇinian ordering (Kiparsky 1973, 1979). For the example above, Bobaljik argues that the representation of the superlative properly contains the representation of the comparative, which in turn properly contains the basic form of the adjective, as in (1).

- (1) a. Positive: [ADJECTIVE]
 b. Comparative: [[ADJECTIVE] COMPARATIVE]
 c. Superlative: [[[ADJECTIVE] COMPARATIVE] SUPERLATIVE]

If a language has a rule of suppletion such as $\text{GOOD} \mapsto be(t)\text{-} / \text{ ___ } \text{COMPARATIVE}$, that rule will block the basic root *good* in both the comparative and the superlative, in virtue of being the most specific rule compatible with the context. Nothing forces the comparative and superlative to share a root – Latin uses an ABC pattern (*bonus-melior-optimus*) with a distinct root in each of the three grades, but the containment relation in (1) ensures that the ABA pattern is underivable (except as a case of accidental homophony).¹

In this paper, we discuss some results of an ongoing project studying the combinatorial properties of rule systems that describe syncretism in morphological paradigms. Although that project did not set out to examine *ABA patterns per se, it turns out that *ABA-like restrictions emerge as a quite general prediction as a consequence of the assumption that Universal Grammar selects the minimal feature inventories needed to generate a paradigm of a given size. We present both a general and a narrow version of this restriction. The narrow, more specific prediction arises if we assume that feature intersection is permitted in the formulation of rules of exponence. We believe this narrow result is particularly interesting, since the *ABA restriction emerges without the containment/nesting hypothesis that characterizes other accounts. Intuitively, *ABA emerges when a three-element sequence is the product of two overlapping features and their intersection: in the sequence (“paradigm”) x, y, z , if x and y share a feature, and y and z share a feature, but x and z do not share a feature, then even without a total containment relation among the features, it follows that the patterns ABC, ABB, and AAB are generable, but ABA is excluded. Most of the paper is devoted to showing that this state of affairs is not only formally possible, but is in fact forced in some contexts by plausible minimal assumptions about feature logics. While this approach seems implausible for some *ABA patterns (we think there are good reasons independent of suppletion to assume that superlatives contain

¹ Recently, Graf (2017) presents a novel account that derives *ABA and other constraints from an abstract order of cells, i.e. $\text{cell-}x < \text{cell-}y < \text{cell-}z$, rather than containment relations among features. See note 19 for some more discussion.

comparatives), we wish to bring this to the table as a possible alternative in other instances. The broader result is that the assumption of minimality, with or without feature intersection, and with or without a limitation to Pāṇinian ordering, yields a class of restrictions on the distribution of paradigm types, of which *ABA is a special case.

Although we have identified the *ABA result as an important point of contact with other current theoretical morphosyntax work, a significant portion of this paper will be devoted to presentation of a framework where classes of morphological models can be formally discussed, and where the effects of individual assumptions can be explicitly computed, for example, in terms of their restrictiveness. Alongside the *ABA result, we also discuss the effect of imposing Pāṇinian ordering on feature models, and show that its effects are comparatively weak in certain classes of model.²

2 Paradigms: Partitions, Features, and Sequences

2.1 Paradigms

Paradigms represent information about the pairing of grammatical properties and linguistic forms. Thus, a paradigm can be seen as a list of cells representing an inventory linguistic forms $\langle x, y, z, \dots \rangle$, each paired with a unique property or combination of properties or features (Stump 2016). As already mentioned in the introduction above, we explore an approach where the order that the cells are presented in plays no role in the morphology. That is our proposal applies to any set of cells regardless of whether a one-dimensional linear order of the cells is assumed or some multidimensional arrangement of the cells. For presentational purposes, we assume there is a conventional linear order for a given inventory of cells, and thus for the corresponding paradigms (lists of forms). As a simple illustration, partial case paradigms for selected German pronouns can be given as in (3), where the cases are presented in the order in (2):

- (2) NOMINATIVE, ACCUSATIVE, DATIVE, ...
- (3) a. 1SG: $\langle ich, mich, mir \rangle$
 b. 1PL: $\langle wir, uns, uns \rangle$
 c. 3PL: $\langle sie, sie, ihr \rangle$

² See also remarks in [Pertsova \(2011\)](#) on the possibly limited role of Pāṇinian ordering in explaining cross-linguistic patterns of syncretism.

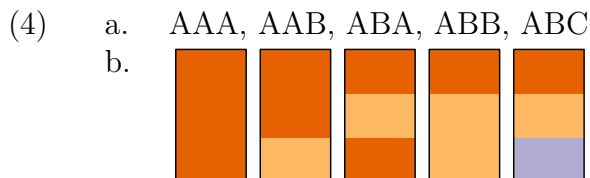
The German pronouns further illustrate a property that is central to the study of paradigms, namely, *syncretism*, the many-to-one mapping from features or properties (like case) to *exponents* (phonological forms) seen in (3b)-(3c). Where the first person singular pronoun is characterized by a three-way contrast, the first and third person plural forms each show only a two-way distinction in form, corresponding to a three-way distinction in grammatical properties. Describing these patterns as syncretic constitutes a claim that the identity of form is represented grammatically, and is thus distinct from *accidental homophony* (see Harbour 2008; Sauerland & Bobaljik 2013). In the notation of the previous section, a syncretic pattern ABB, as in (3b), has only two listed forms (A=*wir*, B=*uns*), and the grammar codes the fact that the B form occupies the second and third cells of the paradigm. Accidental homophony is the state of affairs where the grammar lists three forms (ABC), but two of those forms simply happen to have the same phonology. Our investigation concerns only syncretism, although we recognize that it is in practice a notorious analytical challenge to identify a sharp dividing line for the analysis of specific data points.³

It is widely held that syncretism is central to investigating the nature and inventory of features as in (2). That the accusative and dative forms of the 1PL pronoun are syncretic suggests that those two cells share a feature, a property not represented in (2).⁴ In this manner, comparing the range of attested and unattested syncretisms over some sufficiently large sample may reveal the underlying inventory and organization of the relevant features.

When we abstract away from the details of particular forms, the study of syncretic patterns is at its core the study of partitions of a set, a well-defined mathematical concept. The number of distinct partitions for an n -celled paradigm is the Bell number: B_n . For a three-celled paradigm, the $B_3 = 5$ distinct partitions are listed in (4a). The same information is displayed graphically in (4b) to emphasize that the absolute values (A, B, etc.) are not relevant; all that matters is sameness or difference of cell contents:

³ One criterion for the division is that abstract patterns of syncretism that recur across unrelated languages—our main interest here—are more likely to constitute syncretism than convergent accidental homophony. By this criterion, the patterns in (3b-c) are consistent with a cross-linguistically robust pattern (Smith et al. 2016) and are likely to constitute syncretism. A plausible example of accidental homophony is the 2PL suffix *-t* and the 3SG.PRESENT suffix *-t* in German verbal inflection, which are normally treated as formally distinct, homophonous elements (Albright & Fuß 2012). The discussion in this paper holds for cases of syncretism and not accidental homophony.

⁴ Or equivalently, for example, that features have internal hierarchical structure, or a geometry.



The Bell number grows very fast. There are $B_8 = 4,140$ logically possible partitions of an 8-cell paradigm, and a 10-celled paradigm space (arguably, the number of cases in Russian, see Corbett 2008) has 115,975 possible partitions. In medium- to large-scale studies of syncretism (Cysouw 2003; Bobaljik 2012; Baerman et al. 2005; Harbour 2016), it is commonly observed that only a subset, often only a very small subset, of the theoretically distinct partitions are attested. For example, Cysouw (2003, 2011) considers a sample of person paradigms drawn from more than 250 languages, characterized as an 8-cell paradigm space,⁵ but finds only 60-some-odd distinct partitions from among the more than 4,000 logical possibilities. The *ABA generalizations, mentioned above, make the same point: over some sizeable range of data, where 5 patterns are possible, only four are actually found in the world’s languages: AAA, AAB, ABB, and ABC, but not ABA. Typically, studies of syncretism seek explanations for such typological patterns — i.e. develop theories that predict only a subset of partitions to be possible. In what follows, we address exactly this problem but one level of generality higher — we investigate how general assumptions about morphological analysis restrict which subsets of partitions can arise as typological predictions. For example, we show that a restriction to the partition set {AAA, ABB, ABC} cannot be derived solely within our general assumptions, while the *ABA condition can be derived. In this, we hope here to make new contributions in the formal investigation of the restrictiveness of competing models.

2.2 Features

In order to make headway on these issues, we propose to start by presenting a largely theory-neutral means for representing features. Our notation allows us to express any of numerous competing assumptions about features and feature-logics, allowing us to then compare them directly. We start by recognizing that at its most basic, a feature is a name for individual cells or sets of cells in a paradigm. With reference to an n -celled paradigm, we write a feature as f indexed with a binary vector, where 1 indicates the cell or cells that feature

⁵ Four persons (1,2,3 and inclusive) \times two numbers; see Harbour (2016) for more discussion in particular of other number values. See also appendix A below.

names. Thus, one way of naming features that generate a 3-celled paradigm is as in (5), with a unique feature naming each cell.

- (5) a. f_{100}
 b. f_{010}
 c. f_{001}

We define a *model* of a given paradigm as having two components: an inventory of features, and rules of exponence, which relate features to form. Alongside the simple feature inventory in (5), we may state the rules of exponence in (6):

- (6) a. $f_{100} \mapsto A$
 b. $f_{010} \mapsto B$
 c. $f_{001} \mapsto C$

Each model is a grammar (fragment), generating one paradigm. In the trivial example just considered, the model comprising (5) and (6) generates an ABC partition—a three-celled paradigm that is *maximally differentiated*, i.e., in which each cell has a distinct form. Any number of examples of such an approach can be found in morphological descriptions. The description of the German 1SG pronouns in (3a) could be expressed in these terms. Three unanalyzable case features are assumed: f_{100} = “nominative”, f_{010} = “accusative”, etc., and each one is associated with exactly one exponent, yielding the maximally differentiated paradigm. Analysis of personal pronouns that use three unanalyzed features like “first person” (f_{100}), “second person” (f_{010}), and “third person” (f_{001}) also instantiate this schema.⁶ Any maximally differentiated paradigm can be expressed in these terms. In fact, from the inventory in (5), the maximally differentiated partition ABC is the only *complete* partition that may be generated. (By complete, we mean that a phonological form is assigned to every cell of the paradigm space.) Using only rules of exponence of the format in (6), only maximal differentiation is possible, because the cells share no features in common.

But as we have already seen, maximal differentiation is by no means the only way in which an n -celled paradigm space may be partitioned. Thus a feature inventory that is restricted to generating the maximally differentiated

⁶ Appeal to a “default” form implicitly invokes an additional feature, shared by all the cells: f_{111} , and there is no such feature in (6). We return to the status of the default feature in Section 3.2.2 below, as well as implications of claims of the sort that third person is the absence of a feature (Benveniste 1956), etc.

partition provides no purchase for an account of syncretic patterns, such as ABB, AAB and the like, other than via accidental homophony.⁷

Characterizing syncretic partitions such as the ABB pattern seen in (3b) thus requires features that name (contain) more than one cell of the paradigm such as f_{011} . Consider, from this perspective, the inventory in (7), which represents the standard approach to *ABA generalizations in the notation used here:

- (7) a. f_{001}
 b. f_{011}
 c. f_{111}

This encodes the same relationship among paradigm cells as (1). One feature is shared by all three cells ((7c)—this constitutes the default), one by two, and one is unique to a single element. On the assumption that the feature inventory in (7) remains constant across languages, but that the rules of exponence may vary from language to language or even from lexeme to lexeme, a variety of different paradigms (partitions) may be generated from this single, shared inventory of features. Rules of exponence for two models sharing the inventory in (7) are given in (8) and (9).

- (8) a. $f_{001} \mapsto C$
 b. $f_{011} \mapsto B$
 c. $f_{111} \mapsto A$
- (9) a. $f_{011} \mapsto B$
 b. $f_{111} \mapsto A$

As the reader may verify, the model in (7) + (8) derives a maximally differentiated, ABC paradigm, one in which each cell is distinct from the others. The model consisting of (7) + (9) derives an ABB paradigm, with syncretism of the last two cells.⁸ In these models, rules of exponence are ordered sequentially (read by convention from top to bottom), and disjunctively – the first rule of exponence specified for any given cell must apply to that cell, and once one rule has applied, no other rule may apply. In both (8) and (9), the final

⁷ Or by using more powerful algebras, such as the curly brackets in SPE notation (Chomsky & Halle 1968) representing the disjunction of two distinct rules; see McCawley's (1974) critical remarks on this device; see also section 3.2 below.

⁸ The variables A,B,C etc. are understood to be unique in any given model. Syncretism arises when a single rule of exponence provides the form for multiple cells. Accidental homophony is the term for instances where two or more distinct rules of exponence yield phonological forms that happen to be the same.

exponent (A) is the default – in principle it is compatible with all three cells – but it does not appear in any but the first cell because the rule introducing the default is ‘blocked’ by the application of the more specific rules.

Returning to our German pronoun example, the feature inventory in (7) (unlike the one in (5)) could then support the description of each of the pronouns in (3); the 1SG pronoun using rules of exponence corresponding to (8) and the 1PL to (9). We may represent this outcome more compactly, by listing the features realized by rules of exponence in the feature-based morphological analysis of a given data set as a *sequence* of features (ordered left-to-right, rather than top-to-bottom for compactness). (10a) represents the ordered rules in (8) as a sequence and (10b) that in (9) ((10c) derives (3c) from the same feature inventory). The string to the right of each sequence characterizes the partition defined by that sequence.⁹

- (10) a. $\langle f_{001}, f_{011}, f_{111} \rangle$: ABC
 b. $\langle f_{011}, f_{111} \rangle$: ABB
 c. $\langle f_{001}, f_{111} \rangle$: AAB

The presentation in (10) expresses the fact that several different partitions are derivable from the common feature inventory in (7), by invoking different sequences of features in the rules of exponence.

2.3 Partition Sets

This now gives us the tools we need to introduce the main object of inquiry, namely *partition sets*:

- (11) For any feature inventory I , the *Partition Set* of I , PS_I , is the set of all partitions that may be generated from I .

As we have seen, the partition set of (5) is trivial: $PS_{(5)} = \{ABC\}$. That is, the inventory in (5) will generate all and only maximally differentiated partitions. The partition set of (7) is more interesting, and (10) represents only a subset. To see this, consider all the sequences generable from (7). Since there are 3 features, there are $3! = 6$ (total) sequences to consider, as in (12) (we explain the use of blue font presently):

⁹ Unlike our list notation for paradigms, where order is arbitrary (see section 2.1), the order in a sequence is meaningful, since it represents the bleeding relationships among disjunctively ordered rules.

- (12) a. $\langle f_{001}, f_{011}, f_{111} \rangle$: ABC
 b. $\langle f_{001}, f_{111}, f_{011} \rangle$: AAB
 c. $\langle f_{011}, f_{001}, f_{111} \rangle$: ABB
 d. $\langle f_{011}, f_{111}, f_{001} \rangle$: ABB
 e. $\langle f_{111}, f_{001}, f_{011} \rangle$: AAA
 f. $\langle f_{111}, f_{011}, f_{001} \rangle$: AAA

Collecting the derivable partitions, we find that $PS_{(7)} = \{AAA, ABB, AAB, ABC\}$. Notably, of the $B_3 = 5$ possible partitions of the three-celled space, one is missing: ABA is not included in the partition set of (7).

A different feature inventory may yield a different partition set. For example, adding the default feature f_{111} to the inventory in (5) renders the inventory unrestrictive: any logically possible partition may be derived.¹⁰

In this way, we see the general logic that relates typological generalizations to conclusions about features in Universal Grammar. The data we have are the attested partition sets in some domain—the range of partitions that are (un)attested cross-linguistically. The explanans is then the feature inventory: Domains in which the *ABA generalization holds are domains in which one logically possible partition is not attested. The gap is explained if Universal Grammar admits only the feature inventory in (7)—as (12) shows, the unattested partition is not in the partition set of this inventory.

Our aim in this article is to attempt to approach the issues here from the other direction. In the following, we use the notation introduced here to explore the consequences of various kinds of formal restrictions one could conceivably apply to models of this sort (inventories of features and associated rules of exponence). We do so in the first instance entirely in the abstract, with no connection to substantive features or empirical data. Our goal is to better understand some of the formal properties of feature logics, and to compare the ways in which various intuitively plausible assumptions do and do not restrict the combinatorics.

That is, rather than starting with some observed partition sets and attempting to infer the feature inventory (a task for which there are often multi-

¹⁰ The following partial list of sequences, from among the $4! = 24$ possibilities, demonstrates this point:

- (i) a. $\langle f_{001}, f_{010}, f_{100}, f_{111} \rangle$: ABC
 b. $\langle f_{001}, f_{111}, f_{010}, f_{100} \rangle$: AAB
 c. $\langle f_{010}, f_{111}, f_{001}, f_{100} \rangle$: ABA
 d. $\langle f_{100}, f_{111}, f_{010}, f_{001} \rangle$: ABB
 e. $\langle f_{111}, f_{001}, f_{010}, f_{100} \rangle$: AAA

ple solutions), we investigate here ways in which general constraints on feature algebras do (or do not) restrict the hypothesis space. In other words, rather than arguing that something like (7) is a plausible feature inventory in some domain because it derives the observed facts, we will look instead for general reasons which may favour an inventory like (7) over other inventories on a priori grounds.

One reason to pursue this exercise is that, as our notation calls attention to, absent any prior assumptions about the content of features, the number of possible features that can be defined grows quickly. For a paradigm of n cells, there are $2^n - 1$ non-empty features that may be defined. For a three cell paradigm, the 7 definable features are these:

$$(13) \quad \begin{array}{l} f_{100} \\ f_{010} \\ f_{001} \\ f_{110} \\ f_{101} \\ f_{011} \\ f_{111} \end{array}$$

If features could be freely chosen to form inventories, then 128 distinct feature inventories could in principle be constructed from these features (including the empty set). For a four cell paradigm, there are correspondingly 15 features and 32,768 possible inventories to consider. As we have seen above, from each inventory, a number of distinct models can be constructed. That is, each inventory can be mapped to one or more sequences, thus yielding a variety of partition sets. If rule ordering is unconstrained, then from a single inventory with n features, there are $n!$ distinct sequences that may be so constructed.¹¹ The number of possible models (and thus grammars) thus quickly becomes astronomical, and we suggest it is therefore important to ask whether there may be some universal constraints that drastically restrict the classes of possible models to be considered. Thus, we will spend a fair part of the following discussing the combinatorics involved. We will approach this as follows: Using the understanding of features, paradigms, and models outlined above, we will set out to explore in quantitative terms various conditions that may be imposed, and show explicitly how they do and do not restrict the space of

¹¹ The number is even larger if partial sequences are admitted: the total number of arrangements of a set with n elements: $a(n) = n * a(n - 1) + 1$, $a(0) = 1$. Some of the feature sequences in this group would be incomplete though.

possibilities. Many of the numerical results are non-obvious, and we provide the code in on-line supplemental materials for this paper.

Before proceeding further, by way of a brief housekeeping remark, we note that some of the features in some sequences are redundant. The redundant features in (12) are indicated in blue. Because sequences are ordered, once each cell has been assigned an exponent, all further features will have no effect in characterizing the partition. That is, a feature is redundant in a sequence if it is bled by earlier rules of exponence, and thus in principle cannot be expounded. Eliminating a redundant feature from a sequence is indistinguishable from the sequence with that feature (compare the notion of *inessential feature* in Kracht 1997; Pullum & Tiede 2010). The sequence (12d), for example, is formally indistinguishable from the partial order in (10b), since the first two rules are sufficient to cover all of the cells.¹² If no feature in a sequence is redundant, we call it redundancy-free.

The careful reader may have noticed that in presenting the range of sequences generable from one inventory in (12) we gave only sequences that represent total orderings among the features of the inventory. However, in our exposition above, we also included sequences that contained only a subset of the features (as in (10)). It turns out that consideration of the total sequences is sufficient for calculating the range of partitions generated by an inventory, under the assumption of completeness, which we may define as follows:

- (14) A sequence S is *complete* with respect to a paradigm P iff S generates a form (possibly zero) for every cell in P .

The partial sequence in (10) is complete, but the partial sequences in (15) are not (each leaves one cell without an exponent):

- (15) a. $\langle f_{100}, f_{010} \rangle$: AB_
 b. $\langle f_{001}, f_{011} \rangle$: _AB

¹² We can characterize redundancy abstractly as follows: In a feature sequence S the feature in position j is redundant iff the conjunction of S_j with the disjunction of the features S_1, \dots, S_{j-1} is identical to S_j . As a reviewer observes, various authors have argued that features which may be redundant in one part of an analysis may be useful in another part (see Trommer 2008 for an example). We do not exclude redundant features from inventories categorically.

In general, as a simplification in what follows we will consider only non-redundant complete sequences, since this class is sufficient to exhaustively characterize the partition set of any inventory.¹³

2.4 Restricting the hypothesis space: the minimal valid inventory

Here, we define briefly the two conditions, and note a third assumption, that will be central in the investigation that follows. We suggest these are a priori plausible conditions to restrict the class of possible inventories, and we will work through their consequences in detail in sections 4 and 5 and appendix A below.

The first, basic condition is that an inventory be valid.

- (16) An inventory I is *valid* for a paradigm P iff there exists a model M including I that generates the maximally differentiated partition of P .

The maximally distinct partition of a paradigm (space) is the partition in which each cell is distinct from every other cell. In other words, for a three cell paradigm, a valid inventory is one for which there is some set of rules that will derive ABC. Note that Validity is a property of inventories, not models (grammars). The rules of exponence in (8) demonstrate that (7) is a valid inventory, but we do not require that every grammar (model) generate ABC; syncretism by definition precludes there being such a requirement for every

¹³ Paradigms that appear to have gaps are well attested in the literature, suggesting that completeness is not a universal condition. A famous example is Russian verbs which lack a first person singular present. A reviewer likewise asks about more widespread examples such as apparently deficient (incomplete) pronoun systems in Southeast Asian languages and elsewhere. As the reviewer notes, we may make the provisional assumption here that the deficiency is a matter of lexicalization, and assume that the inventories underlying even incomplete paradigms are valid (as defined in the next subsection): the features are part of the grammar, even if they are not (always) lexicalized. Completeness for us does not constrain inventories and thus does not play a significant role in the main results discussed below. If incomplete paradigms are allowed, then there is a formal distinction to be drawn between *ABA and $\emptyset B\emptyset$; the former underivable, but the latter describable as an incomplete paradigm.

model.¹⁴ The model consisting of (7)+(9) is perfectly well formed (and examples apparently conforming to such ABB patterns are widely instantiated).

The second, and more interesting condition on inventories is Minimality:

- (17) An inventory I constitutes a Minimal Valid Feature Inventory for some paradigm P iff
- a. I is valid for P , and
 - b. there is no alternative I' s.t., I' is also a valid inventory for P and I' has fewer features than I

In Section 4 and Appendix A, we will work through these conditions for various sizes of paradigms, starting with 2-celled and then 3-celled paradigms. One finding in this paper is that the two simple assumptions on inventories just noted – that inventories use the minimal number of features to describe a paradigm space – have the curious effect that in certain paradigm spaces, notably those with three cells, certain patterns of syncretism become unstatable. In a sense to be made clear below, ABA patterns of a certain type are indescribable. More accurately, no minimal valid inventory yields a paradigm set that includes all three bifurcations of the three celled space: {AAB,ABA,ABB}. If two are included, the third is not. Since we have treated order in a paradigm as arbitrary, all of the results we describe hold only up to linear permutations in this way. This result is of interest, because it arises without the nesting/containment assumption that plays a central role in other treatments of *ABA generalizations (Bobaljik 2012; Caha 2009; Starke 2009). Another result is a curious pattern in the nature of the restrictiveness that these assumptions create.

Lastly, at least to start, we will assume following standard practice in morphology that intersection of features is available. If f_a and f_b are in an inventory, then $f_a \cap f_b \mapsto x$ is a well-formed rule of exponence. (I.e., in more standard notation, if $[F] \mapsto A$ and $[G] \mapsto B$ are well-formed rules of exponence, then so is $[F,G] \mapsto C$.) Most feature-based morphological analyses invoke this (for example, if [FEMININE] and [PLURAL] are features in the inventory, then there can be an unanalyzable exponent of [FEMININE,PLURAL], without needing a

14 Maximal differentiation is also not a requirement for every language. Famously, although there are many ways to define the case paradigms for Russian nominals, there is no paradigm that is maximally differentiated in Russian—all Russian case paradigms have some measure of syncretism (Jakobson 1936/1971; see Bobaljik 2002 for some implications of this old observation). Validity is related to, but distinct from, another condition: completeness, which we have mentioned above. Completeness (if it holds) is a property of sequences (and thus derivatively of models), not of inventories.

separate feature [FEMPL]). We assume that intersection is the only Boolean operation on features that is available (but see below for further discussion).

Adding intersection is not innocuous. Because of the way we have defined features and inventories, intersection intersects with Minimality in a non-trivial fashion. Intersection, like rule ordering, allows for exponents that do not directly conform to features that are in the inventory. If an inventory consists only of f_{110} and f_{101} , a rule can be stated referring to: $f_{110} \cap f_{101} = f_{100}$. While this generates an exponent that only expresses the first cell, it does so without the feature f_{100} being contained in the inventory. This will play an important role in the discussion of 3-cell paradigms. Of course, it is worth considering the consequences of minimality and validity without the additional assumption that intersection is available. We do so in section 5.4 below. Note that the core result holds either way, but the inclusion of intersection is a widespread assumption in morphology, so we consider that scenario first.¹⁵

Before proceeding to the discussion of our main results, we will present a number of additional concepts and assumptions which we hope will allow for more familiarity with the general notation. In particular, we offer some remarks on how various current ideas can be rendered in our notation, allowing for commensurability among analyses or frameworks. The reader interested primarily in the consequences of the assumptions just made can skip ahead to section 4.

3 Additional considerations

Our formalism allows us to selectively add or subtract conditions in order to examine the consequences of any particular set of assumptions. In principle, we can translate sets of assumptions from other feature logics into our notation, and can thus accurately investigate the algebra of different combinations. The following subsections illustrate some well-discussed conditions in the field, showing how they can be expressed and evaluated in our terms.

¹⁵ A reviewer asks how our condition of Minimality relates to the condition of Primitivity proposed in (Harbour 2016:Chapter 7). Harbour’s Primitivity condition excludes feature inventories that include features which are interdefinable. As can be seen in section 5, not all inventories excluded by Minimality would run afoul of Harbour’s primitivity condition, but the reverse should hold.

3.1 Extrinsic Order and Pāṇinian Sequences

Above we have noted that sequences (i.e., rules of exponence) must in some cases be ordered. The sequences in (18) contain the same features, but the difference in order alone yields different partitions:

- (18) a. $\langle f_{110}, f_{011} \rangle$: AAB
 b. $\langle f_{011}, f_{110} \rangle$: ABB

As is well known from early discussions of rule systems, rule ordering may be extrinsic (a stipulated language-particular order, as in (18)) or intrinsic, i.e., such that more specific rules automatically bleed more general rules. A specific formulation of the intrinsic Pāṇinian ordering principle or Elsewhere Condition is as in (19) (after Kiparsky 1973):

- (19) If two (incompatible) rules R1, R2 may apply to a given structure, and the context for application of R1 is a (proper) subset of the context for that of R2, then R1 applies and R2 does not.

We translate the operative notion into our set up as in (20), which picks out the class of sequences for which any reordering that does not introduce redundancy has no effect on the partition set. That is, a Pāṇinian sequence is not necessarily a total order, but all order with any consequence is determined by (19).

- (20) A (redundancy-free) sequence S is a *Pāṇinian* sequence if and only if any redundancy-free permutation of S yields the same partition as S .

The sequences in (18) do not satisfy (20). Since (18a) and (18b) are permutations of one another and yield different partitions, neither of them is Pāṇinian. Other than (12a), the sequences in (12) likewise cannot be Pāṇini-sequences since they are not redundancy-free. But the redundancy-free sequence in (21a) is Pāṇinian because it and its only redundancy-free permutation in (21b) yield the same partition: ABC.

- (21) a. $\langle f_{010}, f_{011}, f_{110} \rangle$: ABC
 b. $\langle f_{010}, f_{110}, f_{011} \rangle$: ABC

This raises two points. First, Pāṇinian ordering is the kind of general condition one could entertain as a restriction on rule systems. As the comparison of (18) and (21) shows, imposing a condition that sequences must be Pāṇinian may reduce the partition set for some inventory I by excising all partitions that are derived only by non-Pāṇinian sequences. Rather than build this assumption

in, we see our goal as investigating the effects of assumptions, since we have a notational apparatus that allows us to directly compare systems with and without such an assumption. As it happens (we will present this in more detail below), imposing Pāṇinian ordering will have a drastic effect on minimal feature inventories that are closed under intersection, essentially preventing analysis of syncretism in the three-cell cases.

Now consider again the observation that the sequences in (18) are not Pāṇinian, but those in (21) are. The features invoked in these sequences are not unrelated. The features f_{110} and f_{011} are common to all of these sequences, and the relation between them is that they have partial overlap, but neither is contained in the other. Generally such a relationship between two features f_a and f_b is what makes a sequence non-Pāṇinian, unless there is another feature or other features f_{c_1}, \dots, f_{c_n} that cover the intersection of f_a and f_b and are contained within both f_a and f_b . One easy case is that there is only one other feature f_c (i.e. $n = 1$), namely the intersection or conjunction of the two features f_a and f_b . This is what we see in (21): adding the feature that corresponds to the intersection of the two features in (18) renders the sequences Pāṇinian. Generally, if a redundancy-free sequence is closed under intersection, then it is Pāṇinian. This provides another reason for us to include intersection: allowing intersection makes it easier to compare intrinsically ordered and Pāṇinian analyses. It does, though, raise the question of whether any other Boolean operations on features should be countenanced.

3.2 Feature Algebra: Privativity, Defaults, Containment, and Intersection

While all Boolean operations are generally assumed to be available in semantics, most work in morphology assumes that the feature algebra is restricted and that e.g. the union operation is not available.¹⁶ As noted, we adopt the assumption that intersection, but no other algebraic operation is part of morphology (we do however consider systems without this assumption in section 5.4). We briefly mention some alternatives in the following subsections.

¹⁶ But see, for example, [Stump \(2016\)](#), who allows other Boolean operators, including union (disjunction) in the construction of complex features. We suspect that allowing union will render any valid inventory unrestrictive: any inventory that can generate ABC can generate all other partitions.

3.2.1 Binariness and Dimensions

Consider the example of binary features. Our features are, by definition, privative, rather than binary, in the sense that these terms are understood in the morphological and phonological literature. Binary features, of the sort typically written $[\pm F]$ are, in our terms, names for pairs of features: one feature that names a set of cells, and another feature that names the complement set. In our terms, feature binariness could be expressed by holding that if f_{1100} is a feature in some inventory, then f_{0011} is also a feature in that inventory, etc. Assuming binary features is tantamount to assuming privative features along with a negation operation that is restricted to atomic (non-derived) features.¹⁷ We accord no special status to pairs of features in this way: an inventory containing f_{0011} may or may not contain the complement as a second feature (see also Pullum & Tiede 2010). In at least some cases, including two- and three-cell paradigms, imposing binariness complicates the analysis (cf. Corbett 2010).

Feature binariness is connected to the notion of dimensions in paradigms, raised by a reviewer. The type of representation entertained here readily accommodates multi-dimensional syncretism, a *prima facie* challenge for theoretical approaches, such as Nanosyntax, which adopt a universal total (containment) ordering among features (see Caha & Pantcheva 2012 for ideas on how to extend the Nanosyntax model to accommodate this.) We have thus far represented paradigms as one-dimensional lists, as in (22), although one often finds four-celled paradigms presented as a 2×2 matrix, encoded as two binary features, as in (23).

(22) $\langle A, B, C, D \rangle$

(23)

	$-\alpha$	$+\alpha$
$-\beta$	A	B
$+\beta$	C	D

Translation is straightforward: the feature $-\alpha$ in (23) is encoded relative to the list in (22) in our terms as f_{1010} , $+\beta$ as f_{0011} , etc. But dimensions of paradigms, underlying horizontal and vertical syncretisms, have no a priori special status—we can just as readily define a feature f_{1001} which picks out cells A and D, a diagonal syncretism in (23). For us, this flexibility is an advantage, since it allows us to take any existing partition set and probe what the optimal underlying feature inventory might be, given any combination

¹⁷ Even proponents of binariness and intersection of features assume that $-[A \cap B]$ isn't necessarily part of an inventory containing features A and B.

of assumptions such as binarity, Pāṇinian order, Minimality etc. Rather than setting the features ahead of time, we can in this way discover whether features should be binary or not. In our terms, the two binary features that define the matrix in (23) are the inventory: $\{f_{1010}, f_{0101}, f_{1100}, f_{0011}\}$, but this is simply one of more than 32,000 inventories that could have been used in the analysis of a given four-cell paradigm.¹⁸

3.2.2 Defaults and containment

A default feature (or value) is one that is compatible in principle with all cells, i.e., $f_{11\dots 1}$, for a paradigm of any arbitrary size. Many approaches accord a special status to the default. For example, as a reviewer notes, theories that treat features as attribute:value pairs, or equivalents, such as category:feature/value etc., may allow for reference to the category as a whole as a default (Adger & Svenonius 2011 is a recent, explicit example of this, but the general approach has many antecedents). A three-celled paradigm could be described in these terms such that one element is the default, corresponding to the absence of a

¹⁸ There are $(2^4 - 1 =)$ 15 different features, and thus 2^{15} definable inventories (not all of which will be valid, of course). (i) provides an model using the inventory corresponding to binary features.

- (i) a. Inventory: $f_{1010}, f_{1100}, f_{0101}, f_{0011}$
- b. $f_{1010} \cap f_{1100} \mapsto A$
 $f_{0101} \cap f_{1100} \mapsto B$
 $f_{1010} \cap f_{0011} \mapsto C$
 $f_{0101} \cap f_{0011} \mapsto D$

An alternative analysis of the four cell paradigm without binarity is given in (ii):

- (ii) a. Inventory: $f_{1010}, f_{1100}, f_{0111}$
- b. $f_{1010} \cap f_{1100} \mapsto A$
 $f_{1100} \cap f_{0111} \mapsto B$
 $f_{1010} \mapsto C$
 $f_{0111} \mapsto D$

Note that eschewing binarity allows for a smaller inventory of features (three features instead of four), hence there is no a priori argument from simplicity in favour of binarity.

value for the attribute, as in (24), where the third line spells out the absence of a value (a category, but no feature, in Adger & Svenonius’s terms):

$$(24) \quad \begin{array}{l} f_{100} \mapsto A \\ f_{010} \mapsto B \\ f \mapsto C \end{array}$$

Such analyses are readily found in the literature. For example, analyses that treat the third person as the “absence” of person or the default person instantiate (24). But in our terms, (24) is simply a notational variant of (25). An underspecified, default exponent corresponds to an exponent that is compatible with any cell in the paradigm, and surfaces wherever it is not bled by an earlier rule.

$$(25) \quad \begin{array}{l} f_{100} \mapsto A \\ f_{010} \mapsto B \\ f_{111} \mapsto C \end{array}$$

Just as with feature binarity, we choose from the outset not to assign any privileged status to the default (or to inventories containing a default)—it is simply one feature among many to be considered. We consider full sets of inventories, including those that do and do not contain the default. Doing so allows us to compare the results of inventories that include the default with those that do not. For example, it could turn out that inventories that include the default as one of the features are more highly valued along some dimension than those that do not. But unlike Adger & Svenonius (2011), we do not build this assumption in from the start.

Related to defaults, there are also theories that build containment in as a prior assumption about feature inventories. Much of the *ABA literature relies on partial or total containment among classes of features in an inventory. The Nanosyntax framework codes this as an *fseq*, assumed to be universal and invariant across languages (Caha 2009). Other ABA literature (Bobaljik 2012; Smith et al. 2016) assumes containment in the contexts where ABA is excluded, but without a total commitment to invariant *fseqs*. As described above, feature containment relations can readily be expressed in our notation, as in (7). The *fseq* assumption would then elevate that to a general condition: for any two features f_a, f_b in an inventory, either $f_a \subset f_b$ or $f_b \subset f_a$. Once again, we do not impose a priori conditions of this sort, as our aim is to see whether these arise as plausible conditions from other considerations.

In the above paragraphs, we hope to have shown that any of a number of other conditions on inventories or sequences could be expressed in our system.¹⁹ Our primary strategy here is to limit building assumptions into our system, so that this will allow us, at least in principle, to consider the restrictiveness of various possible assumptions in the abstract, and to allow for direct formal comparison of classes of competing frameworks. We now begin the process of exploring the consequences of the assumptions we did suggest for paradigms of different sizes.

19 We have not engaged here with the proposals in Harbour (2011, 2016), and Ackema & Neeleman (2013). Harbour and Ackema and Neeleman contend that standard frameworks treat features as first order predicates, whose values serve as one-place truth functors, but that this should be replaced by a perspective in which features are operators that induce partitions of lattices (or their atoms). Harbour’s approach supports recursive composition of feature values, such as a number value like [-singular,+augmented,-augmented] = trial. This is non-contradictory, since the outer value of augmented acts on the result of having previously applied the inner value. We believe that these types of proposal can be expressed in our notation (although the composition of feature inventories becomes non-trivial) since our features fundamentally, like Harbour’s, define a partition of a set (in our case, the list indicated as a binary vector on f). Direct engagement with these proposals takes us far too far afield for the present article, though.

Another approach for which we postpone fuller engagement is that in Graf (2017), which does not use features as such, and instead operates on cells directly, positing an abstract order of cells, i.e. cell- $x < \text{cell-}y < \text{cell-}z$. Though Graf’s proposal is more complex, at this point we can offer a remark on a version of his proposal that relates to the convexity assumption (cf. Gärdenfors 2000). Assume that the set of all cells with the same exponent must be convex: if cells x and z have the same exponent and there is a cell y such that $x < y < z$ relative to the order, then y must also have the same exponent as x and z . Convexity stated in this way predicts that, if three cells are ordered linearly as $x < y < z$, the pattern *ABA is ruled out. Note that Graf’s proposal relies on three assumptions that should be discussed further: that there is always an order of cells in a paradigm, and that the convexity constraint must apply directly to cell-exponent relations, and cannot be satisfied at an intermediate featural level. Graf discusses the former assumption explicitly – consider, for example, that an ABA-pattern does not violate convexity if the cells only stand in a partial order where $x > y$ and $x > z$ but where y and z are not ordered relative to one another (cf. Smith et al. 2016). The latter assumption is not discussed by Graf in detail, but is clearly necessary: Consider both a feature f_{xyz} shared by all three cells xyz and a feature f_y singling out just y are convex assuming the linear order $x < y < z$. Then the exponence mappings $f_y \mapsto B$ and $f_{xyz} \mapsto A$ derive the ABA-pattern. Specifying these assumptions helps us locate our discussion relative to Graf’s proposal. Namely, the proposal for deriving *ABA we explore in the following takes the opposite direction of Graf’s: we assume no inherent order of cells, but rely strongly on features. It is interesting that these two at least superficially quite different routes arrive at similar results and we hope this will encourage a more detailed comparison of the two approaches in the future.

4 Two-Cell Paradigms

Consider first the case of paradigms with two cells. Analysis of a two-cell paradigm space is relatively trivial, but serves as a warm up for the more interesting cases, and offers an opportunity to become more familiar with the notation for presenting the analysis.

For the analysis of a two-celled paradigm space, there are three logically possible features: f_{10} , f_{01} , and f_{11} – this corresponds to the general formula that for n -cells there are $2^n - 1$ possible features. From three features, eight distinct inventories of features may be defined, i.e., the power set of the features. Of these, we may discard the empty set - if there are no features, nothing can be described.

Of the seven remaining inventories, any inventory consisting of just a single feature will fail our criterion of Validity: The maximally differentiated partition of a two-celled paradigm space is AB, i.e., the two cells are distinct. Since our features are privative, a single feature is not sufficient to analyze the AB paradigm: If the single feature is f_{11} it isn't possible to make the required distinction between the A and the B cell – the only partition that can be generated is AA. And if the single feature was either f_{10} or f_{01} , no analysis of the two cell paradigm is possible at all. Only one cell could receive an exponent. Recall that we made the decision not to assign the 'default' f_{11} some special status but to include it as just one possible feature among many. Therefore if only a rule of exponence $f_{10} \mapsto A$ is specified, the second cell wouldn't be filled at all. Therefore this analysis fails to be valid under (16). This shows that at least 2 features are required to analyze the AB paradigm.

That two features are sufficient is shown by looking at Table 1. This table displays the four inventories with two or three features. For each inventory, the set of possible rules of exponence are given (redundant features are in parentheses), and in the rightmost column, the corresponding partitions that can be generated. As the table shows, any selection of two features from the three possible features will allow an analysis of the AB-paradigm. Thus these three subsets represent possibilities for a restrictive Universal Grammar satisfying Minimality — the three-feature inventory (#4) is excluded by this criterion.

The same information can be represented graphically as patterns of squares, here aligned vertically using colour to define features, exponents and partitions. In Table 2, we display the three valid systems that contain the minimal number of features, i.e., two features, in the two-cell case in this way. As we discuss immediately below, while there are three such distinct feature inventories, inventory 1 and inventory 2 predict the same sets of possible partitions.

#	inventory	sequence	partition
1	f_{01}, f_{11}	f_{11}	AA
		$f_{11}, (f_{01})$	AA
		f_{01}, f_{11}	AB
		f_{01}	**
2	f_{10}, f_{11}	f_{11}	AA
		$f_{11}, (f_{10})$	AA
		f_{10}, f_{11}	AB
		f_{10}	**
3	f_{10}, f_{01}	f_{10}, f_{01}	AB
		f_{01}, f_{10}	AB
		f_{01}	**
		f_{10}	**
4	f_{10}, f_{01}, f_{11}	$f_{10}, f_{01}, (f_{11})$	AB
		$f_{01}, f_{10}, (f_{11})$	AB
		$f_{10}, f_{11}, (f_{01})$	AB
		$f_{11}, (f_{01}, f_{10})$	AA
		f_{01}	**
		f_{10}	**

Table 1 Table of valid feature inventories and corresponding partition sets for two-cells. ** = incomplete sequence


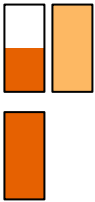
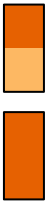

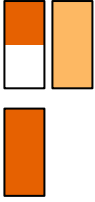
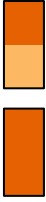



#	inventory	sequences	partitions
1			
2			
3			
count	3		2

Table 2 Graphical display of table of minimal valid two-cell inventories and derivation of the predicted partitions. Compared to table 1, incomplete sequences and redundant features are omitted here, and sequences that are order-invariant shown only once. Inventory 4 of table 1 is excluded here by minimality.

But inventory 3 predicts a smaller set of possible partitions, namely only the AB partition.

Inventory #1 is valid, since there is a sequences of features from this inventory, which generates the maximally differentiated partition AB. This sequence is in the first line: there are two, ordered rules of exponence ($f_{01} \mapsto B$, and $f_{11} \mapsto A$). As the table shows, the AA partition may also be generated from the same inventory. The first sequence provides a rule of exponence only for the feature f_{11} . This generates the fully syncretic paradigm: AA. Continuing through the table, we see in this way that the first and the second possible universal inventory each allow two classes of languages corresponding to the partitions AA and AB. The third feature inventory, although Valid and Minimal, only predicts the AB partition as possibility. On this analysis, if AA were to surface in any language it would need to be the result of accidental homophony.

The fourth inventory contains all three features and therefore allows 6 sequences with rules of exponence for 3 features, 6 sequences with 2 features, and 3 with single features, which we show in a condensed form in Table 1. However, as noted, this inventory fails the Minimality condition.

Recall from above that we defined the partition set of an inventory as the set of paradigms that can be derived from it. For the three minimal complete inventories of the two cell case, the partition sets can be read off the partitions column of Table Table 1.

Typological evidence or learning experiments ultimately can inform us which partitions are attested. If a typological survey shows that both AA and AB patterns exist, then the the third inventory, though valid and minimal, is not the actual inventory made available by UG. We note in passing that it is the only minimal valid analysis that uses a binary feature, rather than the equivalent of a default and “marked” combination.

However, the typological evidence cannot alone decide between different inventories that both predict the same possible partitions like inventories #1 and #2 above.

Despite the relatively trivial nature of the exercise with the two-cell paradigm space, the preceding discussion demonstrates that assumptions have consequences, and the the assumption that UG inventories be both minimal and valid has reduced the space of possible inventories from 7 (or 8 with the empty set) to 3. We have shown how typological evidence can be brought to bear on the choice. Finally, we note that the two minimal valid inventories that are capable of generating both AA and AB patterns are in fact *linear permutations* of one another. Since we have taken the linear order of the paradigm cells to be arbitrary, there is no way to resolve this choice further on our assumptions. For example, to describe a two-way number contrast, one could invoke two features: singular and plural, corresponding to inventory #3, or posit a single marked number value (say, f_{01} = plural) and leave the other unmarked (f_{11} = number), or combine these to use two marked features and a default. Minimality prefers one of the first three inventories; if there is syncretism in some paradigms, then one of the first two is to be preferred. But considerations of Minimality and Validity alone do not resolve the venerable debate about which value of number is marked (Sauerland et al. 2005 and others).

Our first result is that in a paradigm space that constitutes only a binary opposition, the only minimal valid analyses that also permit syncretism are the ones that takes UG to have a single feature that names one member of the opposition, and which is contrasted with a default feature, compatible with both members. In this way, there would be an empirically-grounded argument to be made that if Minimality is assumed, then Binarity should be rejected as

a general condition on feature inventories. In the manner just noted, the two assumptions make contrasting predictions about the state of the world. But we have no way on these considerations alone of saying which member of the opposition is ‘marked.’

5 Three-Cell Paradigms

Turning to the three-cell paradigm space, we begin to see the growth in the space of analytical possibilities, and we also see how various assumptions such as Minimality and intrinsic, i.e., Pāṇinian ordering restrict that space. For a three-cell paradigm space, there are $2^3 - 1 = 7$ possible features, listed in (26):

$$(26) \quad f_{100}, f_{010}, f_{001}, f_{110}, f_{101}, f_{011}, f_{111}$$

If features could be freely chosen to form inventories, then 128 distinct feature inventories could in principle be constructed from these features (including the empty set), i.e., $(2^{2^n - 1})$, a function with double exponential growth). Each inventory is in turn relatable to $n!$ (total) sequences.

5.1 Restrictions

We now consider the degree to which the assumptions mentioned above restrict the space of possible grammars (analyses).

5.1.1 Validity

The first restriction we impose is Validity, as in (16). For example, the inventory $f_{100}, f_{010}, f_{001}$ is valid, in that it describes a three-way contrast, while the inventory $f_{100}, f_{110}, f_{010}$ is invalid - it is not complete, as it provides no means to describe the third cell. It turns out that 96 of the 128 possible inventories of active features are valid in this sense in the three cell case (see Table 3 below). With four cells, the ratio is 31 962 out of 32 768 (see Table 6 below). Validity thus restricts the number of feature sets, but the restriction is not particularly strong.

5.1.2 Minimal Feature Inventory

The more interesting (and less obviously empirically motivated) requirement is Minimality. As defined above, a minimally valid feature inventory is an inventory that contains the minimal number of features needed to describe the maximally differentiated partition. For the two-cell space, the minimality

requirement does not restrict the possibilities in any interesting way (it excludes only one inventory out of the 4 valid ones), but for the three-cell space, the minimal number of features that is needed to describe the maximally differentiated partition is two, as we show presently. Validity plus Minimality together thus restrict the choice from among 128 different logically possible feature inventories to the following three:

- (27) a. f_{110}, f_{101}
 b. f_{101}, f_{011}
 c. f_{110}, f_{011}

To see that (27) are the minimal valid inventories, consider first that they are indeed each valid: (28) gives the rules of exponence that generate the maximally differentiated partition from the inventory in (27c). For (27a) and (27b) analogous sequences can be specified.

- (28) a. $f_{110} \cap f_{011} = f_{010} \mapsto B$
 b. $f_{110} \mapsto A$
 c. $f_{011} \mapsto C$

Now consider minimality: Obviously, no inventory with fewer than two features can be valid, hence we only need to show that the inventories in (27) are the only valid two cell inventories. Assume that there was another valid inventory I with only two features. Because (27) lists all combinations of the features f_{110}, f_{101} , and f_{011} , I would need to contain one of f_{100}, f_{010} and f_{001} , or f_{111} . But it is easy to see that for any these features, it is impossible to satisfy validity by only adding one further feature to I . Hence, (27) are the three minimal inventories for three cells.

5.1.3 Order and the Effect of Pāṇini

Note that in order to describe the ABC pattern, the rules of exponence must be (partially) ordered, such that the exponent of the conjoined features takes preference over the rules in (28b)-(28c) (this holds for any of the three inventories in (27)). The property of Order was not relevant in the two-cell paradigm, but the Pāṇinian order condition has a strong effect with the valid three cell inventories.

Because each of the inventories in (27) has two basic features that may be conjoined to define a third feature, the number of possible sequences for each inventory is 16, although many of these sequences will be redundant or incomplete. The order (28) is in addition to deriving the complete partition

also Pāṇinian as defined above. To see this consider first that if the order of (28b) and (28c) is changed as in (29), the resulting sequence still derives the complete partition ABC.

- (29) a. $f_{110} \cap f_{011} = f_{010} \mapsto B$
 b. $f_{011} \mapsto C$
 c. $f_{110} \mapsto A$

Other orders of the rules in (28) render feature f_{010} redundant. For example, if the order of the first two rules in (29) is exchanged, then the rule $f_{110} \mapsto A$ will assign the exponent A to the first two cells, *bleeding* rule (30b) (i.e., rendering rule (30b) redundant).

- (30) a. $f_{110} \mapsto A$
 b. $f_{110} \cap f_{011} = f_{010} \mapsto B$
 c. $f_{011} \mapsto C$

As a general property (well understood from studies of Rule Ordering), ordering f_a before $f_a \cap f_b$ will render the conjunction redundant, and is thus equivalent to not selecting (or having no rule referencing) the conjoined feature. This corresponds to an *intrinsic* order: if the conjoined rule is active, it must be ordered before its individual conjuncts.

If the conjoined rule is omitted or not ordered first, it can be omitted and only the order between the two rules referring to the basic features matters. The resulting sequences are not Pāṇinian, but require extrinsic order. (31) yields AAC while (32) yields ACC.

- (31) a. $f_{110} \mapsto A$
 b. $f_{011} \mapsto C$
 (32) a. $f_{011} \mapsto C$
 b. $f_{110} \mapsto A$

The following table shows, for one inventory, the six possible sequences (six distinct orders of three rules) and the three corresponding partitions that are derived. (As before, redundant elements in the sequences are in parentheses). The analogous table for the other two choices can be readily constructed. As an expository device, we use green text to indicate a feature that is derived as the intersection of the two basic features. As explained in section 2.4, the green features are not part of the feature inventory, but are a convenient abbreviation for rules of exponence that make reference to the intersection of two features in their structural description.

(33)	inventory	sequences	partition
	f_{110}, f_{011}	$f_{010}, f_{110}, f_{011}$	ABC
		$f_{010}, f_{011}, f_{110}$	ABC
		$f_{110}, (f_{010}), f_{011}$	AAC
		$f_{110}, f_{011}, (f_{010})$	AAC
		$f_{011}, (f_{010}), f_{110}$	ACC
		$f_{011}, f_{110}, (f_{010})$	ACC

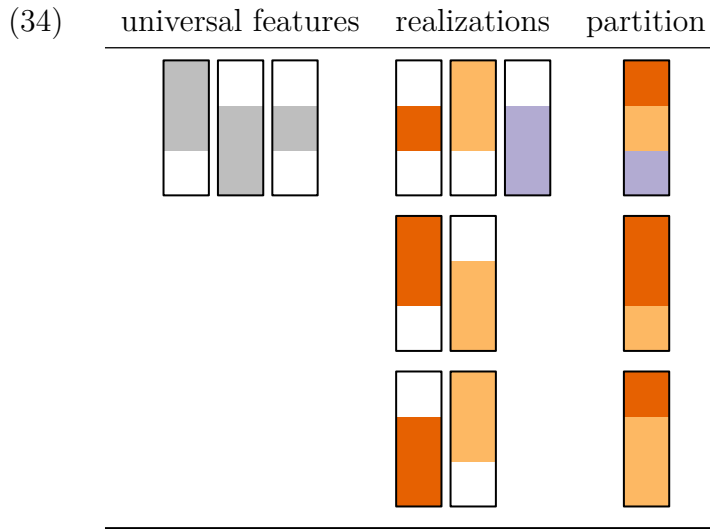
3

What (33) shows is the following: There are (only) three minimal valid feature inventories that can generate a maximally differentiated three-celled paradigm space. One such inventory is $\{f_{110}, f_{011}\}$. From that inventory, 6 ($= 3!$) sequences may be formulated, where each sequence is a distinct, total ordering of rules of exponence for the two features and their intersection.²⁰ While there are six rule orderings possible, only three distinct partitions are generated. The first two lines in (33) derive the same surface patterns (partitions), since the ordering of the last two rules is irrelevant.

As the reader may verify, the other two minimal valid inventories (in (27a)) have the same properties as (33). The three inventories amount to permutations in the order of the cells, but are otherwise identical in their formal properties. Each inventory generates a partition set that contains only two of the three logically possible bifurcations of the paradigm. Since we have not stipulated a meaningful order of the paradigm cells, the three are equivalent, up to linear order.

The information in (33) is represented graphically in (34):

²⁰ If non-total sequences are included, there are 15 possibilities, but the additional sequences are either incomplete, or indistinct from the sequences in (33) which have redundant rules.



5.2 Result: *ABA

At this point, we note two properties we believe to be of theoretical interest. For a three-celled paradigm space, there are $B_3 = 5$ distinct partitions. However, imposing the conditions of Validity and Minimality on the UG feature inventories restricts the expressive power of the system, such that each inventory generates only 3 of the 5 possible partitions. The three inventories that are permitted are moreover linear permutations of one another. We believe this is of interest since it appears to be true at least in some domains that the number of attested partitions is a small subset of the logically possible ones. The example we noted above was that in the 8 cell division of the person/number space, only 60-some-odd distinct partitions, out of $B_8 = 4,140$ possibilities, are attested in Cysouw's 250+ language sample. Being able to predict restrictions on the space of possibilities is thus of potential theoretical interest, if the restrictions indeed line up with the data. In the case at hand, the following restrictions obtain:

Of the five possible partitions of a three-cell space, four show some differentiation among the cells. However, each of the inventories in (27a) generates only three of those partitions. As in the case of inventory #3 in the two-celled paradigms, we are now able to connect our formal results to potential empirical evidence. If there is, as we have hypothesized, a fact of the matter for some domain, such that UG contains only one of the inventories in (27a), then this should show up as the following empirical generalization: across the relevant domain, only three of the four possible patterns of differentiated partition should be attested. In (33), we show that the inventory f_{110}, f_{011} generates

the partition set $\{ABC, AAB, ABB\}$; that inventory does not generate ABA. No sequence from that inventory will generate a pattern in which the first and last cell share an exponent, to the exclusion of the middle cell. The same holds for the other two inventories, up to linear order: each inventory will fail to generate exactly one of the possible partly syncretic partitions.

This result is noteworthy in the current context, since it provides a means of characterizing the absence of *ABA patterns without assuming featural containment. Existing accounts of *ABA patterns invoking containment are all built on what, in our terms, is a non-minimal feature structure, with strict nesting of features – some version of: $f_{100}, f_{110}, f_{111}$.

In other words, what we have just shown has two parts. The easy part is a demonstration that it is possible to derive a *ABA generalization for some domain without invoking containment. We have just done so. The slightly harder part was the demonstration that the type of feature inventory that derives *ABA without containment is not only possible, but is in fact preferred (over containment), if UG makes use of Minimally Valid feature inventories. We postpone until the next section some speculative remarks on whether this result constitutes a plausible alternative scenario for the account of *ABA generalization examples in the literature.

Before that discussion, we note one further point about these inventories. No valid, minimal inventory for a 3-cell paradigm space generates the maximally undifferentiated partition AAA. Curiously, it is not a general property of our assumptions that such undifferentiated partitions are universally excluded in the minimally valid inventories, and we show below that it does not hold for four cells. We can say that at this point that the undifferentiated partitions are excluded when the number of cells is from the sequence $2^n - 1$ for $n \geq 2$, i.e. 3, 7, 15, ... We note this, but leave it as an unexplored aspect of the system. Total syncretism appears to exist, of course, and we do not exclude it across the board. We return to this issue again in section 5.4. Giving up the assumption that intersection is always permitted will preserve the generalized *ABA result considered here, but will admit AAA patterns. The upshot of that section will be that the (equivalent of the) 3 inventories considered to be minimally valid inventories with intersection become three among a larger class of minimally valid inventories (including the containment patterns). Some inventories from among the larger class permit AAA, but the general result holds: no member of that larger class admits all three bifurcations of the paradigm space: any minimal valid inventory whose partition set contains ABB and AAB will necessarily exclude ABA.

5.3 More on the 3-cell space: Non-Minimal sets

Thus far, we have examined only the three minimally valid inventories that generate a three-cell paradigm. To evaluate the effect of minimality, we now look also at non-minimal inventories. In the two-cell case, we were able to present a complete discussion of all the possible inventories and of the partition sets described by each inventory. There were only 8 possible inventories for the features definable over a two-cell paradigm, and 4 inventories were invalid. But for a three cell space, there are 128 inventories, and numerous sequences to consider.

Table 3 provides a summary of important aspects of the grammar of three-celled paradigms and the models that generate them. In the next paragraphs, we walk through this table in some detail, identifying various properties that are of potential interest. Among these, we note that imposing Pāṇinian ordering — limiting all models to intrinsic rule ordering — turns out to have rather drastic consequences. Possibly of more interest, we note that there are some partition sets that do not arise under any constellation of the assumptions considered here. Even without Minimality, for example, feature inventories turn out to be somewhat restrictive.

Table 3 is divided horizontally into two halves. Each half tabulates all the valid feature inventories, and counts inventories grouped by the number of features they contain (y-axis) \times the partition sets that may be generated from them (x-axis). The two halves of the table differ as follows: In the top half, it is assumed that extrinsic order of rules of exponence is permitted, while in the bottom half, we add the additional assumption that only intrinsic (Pāṇinian) ordering is permitted. We discuss the differences below.

The columns in Table 3 represent possible partition sets of a three-cell paradigm space, using colour instead of letters, as in (34) above: the same colour in two cells indicates the same exponent (syncretism). There are $B_3 = 5$ distinct partitions (the rightmost column) and 16 different subsets of partition that contain the maximally differentiated partition (ABC = dark orange, light orange, light purple).

The header of each column represents a distinct partition set, and the number in a given column represents the number of formally distinct (valid) inventories that can in principle generate that set. The three minimal valid inventories that we have discussed above are in the top row of the top half of the table (columns 6-8). These are the only three valid, two-feature inventories. But the table provides a range of information about what happens if we do not include the minimality requirement.

2 features, order	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3 features, order	1	2 2 2	-	1 1 1	3 3 3	-	-	-	-	-	-	-	-	-	-	-
4 features, order	-	1 1 1	-	3 3 3	-	-	-	-	-	-	-	-	-	-	-	-
5 features, order	-	-	-	1 1 1	-	-	-	-	-	-	-	-	-	-	-	-
6 features, order	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
7 features, order	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
total with order	1	3 3 3	0	8 8 8	0	0 0 0	0	0 0 0	8	5 5 5	5	5 5 5	5	5 5 5	5	39
2 features, Pāṇini	3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3 features, Pāṇini	4	4 4 4	-	-	-	-	-	-	-	1	3 3 3	3	3 3 3	3	-	-
4 features, Pāṇini	-	3 3 3	-	1 1 1	-	-	-	-	-	3	2 2 2	14	2 2 2	2	-	-
5 features, Pāṇini	-	-	-	1 1 1	-	-	-	-	-	3	-	15	-	-	-	-
6 features, Pāṇini	-	-	-	-	-	-	-	-	-	1	-	6	-	-	-	-
7 features, Pāṇini	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-
total Pāṇini	7	7 7 7	0	2 2 2	0	0 0 0	0	0 0 0	8	5 5 5	5	5 5 5	5	5 5 5	5	39

Table 3 Table showing which 12 of the 16 logically possible three cell partition sets can be generated by feature inventories (applying intersective closure) and by how many inventories. Partition sets related by cell permutation are grouped together.

In the leftmost column of the line “3 features, order”, one finds the number 1. Assuming extrinsic rule ordering is allowed, there is exactly one choice of an inventory with three features, from among the 7 possible features, which yields only an ABC partition. We have seen that already; it was the inventory in (5). If that inventory is chosen, from among the 128 possible inventories, then the only partition that can be generated is ABC.

On the same line, the number in the rightmost column is 3. There are (exactly) three distinct choices of feature inventories from each of which all five logically possible inventories can be derived. One such inventory is $f_{110}, f_{101}, f_{111}$, i.e. it is derived from a valid two-feature inventory by adding the default f_{111} . The other two inventories are also of this type; i.e., the two linear permutations of this inventory.²¹

This line also shows that there are 3-feature inventories that generate a partition set which excludes ABA. For example, the third column from the right notes that there are three inventories whose partition sets contain ABC, AAB, ABB, and AAA, but not ABA. One of the three inventories which generate this partition set is $f_{001}, f_{011}, f_{111}$ as we saw above already (the containment inventory). A second possibility is $f_{100}, f_{110}, f_{111}$ (a linear permutation of the previous one). Finally, the inventory $f_{100}, f_{001}, f_{111}$ also generates this partition set, but without containment. Furthermore, all three inventories exclude *ABA from their corresponding partition set regardless of whether extrinsic ordering is allowed or not.

Bear in mind that the numbers in this table do not count models or sequences, but count inventories. Other than those in the leftmost column, each valid inventory in the table may be contained in multiple models, thus yielding sets of generable partitions. For example (33) (= (34)) is here coded by the number 1 in the top line, column 7. This is a two-feature inventory that generates the partition set at the top of column 7; moreover, this is the only choice of (two) features which generates that exact partition set (and requires extrinsic rule ordering to do so).

²¹ The following partial list of (redundancy-free) sequences demonstrates that this inventory is unrestricted:

- (i) a. $\langle f_{110} \cap f_{101}, f_{110}, f_{111} \rangle$: ABC
- b. $\langle f_{110} \cap f_{101}, f_{111} \rangle$: ABB
- c. $\langle f_{101}, f_{111} \rangle$: ABA
- d. $\langle f_{110}, f_{111} \rangle$: AAB
- e. $\langle f_{111}, \dots \rangle$: AAA

5.3.1 Results: Impossible Partition Sets

One point of interest is that there are four partition sets that are underivable no matter the size of the inventory: four columns total to zero (in fact the same four with or without a limitation to Pāṇinian ordering). As the fifth column shows, there is, for example, no valid inventory (minimal or otherwise) that has the partition set {ABC, AAA}. In other words, no combination of features will admit all and only the maximally and minimally differentiated partitions. Also excluded are patterns that allow ABC, AAA and exactly one syncretic grouping (columns 9-11).

This latter fact is particularly interesting, since the last of these (column 11) is what Bobaljik (2012) finds empirically for suppletion in adjective gradation: ABA and AAB are unattested, but the other patterns are allowed. Our result means that the suppletion pattern of gradation isn't predicted by any variation of the morphological assumptions we consider here – i.e. whether Pāṇini, Minimality or other similar conditions are assumed. However, Bobaljik also proposes to separate the component accounts of *ABA from *AAB in adjectival gradation, arguing that only *ABA is excluded by the logic of features and syncretism, and proposes an additional, syntactic locality condition to exclude *AAB (see also Bobaljik & Wurmbrand 2013).

5.3.2 Pāṇini

Before leaving the domain of three-cell paradigm spaces, we will consider the effect of one additional restriction, namely the idea that there is no extrinsic ordering of rules, and only Pāṇinian ordering. Each of the three valid, minimal feature inventories makes use of two basic overlapping features, and derives a third by using the intersection of those two. We showed above that reordering the rules has the effect of deriving syncretic patterns, in effect, by rendering the intersective feature redundant. The order in (30a) is equivalent to a system that uses only the two basic features, but not their conjunction.

We may consider imposing Pāṇinian-order-only as a restriction on valid sequences, corresponding to the hypothesis that grammars make use of only intrinsic, but not extrinsic, ordering of rules. Comparing the top and bottom halves of Table 3 allows us to evaluate the effects of this assumption, for three-celled paradigms.

One result which we find interesting is that for 3-cell paradigms, imposing Pāṇinian ordering has no effect on the total number of valid inventories. (This turns out to be different for 4-celled paradigms). We simply note this here, without further comment.

However, comparing the first line of each half of the table shows that imposing Pāṇinian ordering in addition to Minimality is a severe restriction. This constellation of assumptions has the effect that only the maximally distinct partition is describable (the leftmost column in Table 3). All three valid minimal inventories will derive that order and no other. Technically, intrinsic ordering does not restrict the relative order of f_{110} and f_{011} , but since the conjunction will identify the middle cell, the remaining ordering is free (the two are non-distinct).

Since syncretism is abundant in paradigms of all sizes, imposing a Pāṇinian ordering, along with the other assumptions considered above, seems, in its combination with intersection, pathologically over-restrictive. Somewhat different results obtain if we do not assume that intersection is freely available, so we turn to that now.

5.4 Deriving *ABA in Intersection Free Systems

As we mentioned in section 3.2 above, assuming that intersection of features is available to rules of exponence accords with standard practice in morphological theory. Systems where the feature set is closed under intersection, combined with the assumptions of minimality and validity, yield tight restrictions on paradigm sets, including one that seems to be of special interest in current morphology and therefore we have focussed so far on systems with intersection. In this section we discuss what happens if we drop this requirement. In particular, we show that these allow a different route to derive a generalized *ABA constraint. Table 4 shows an overview of the possibilities for deriving the 16 valid partition sets for the three cell case when feature intersection isn't available.

One consequence of assuming that intersection is not available for morphological feature algebras is that a valid inventory must contain least three features.²² This can be seen as follows: the minimally valid inventories with two features when intersection is available contain two features such as f_{110} and f_{011} which each contain two cells. But if intersection is not available, such a system cannot describe the maximally differentiated paradigm and thus is not valid. In this case, $\{f_{110}, f_{011}, f_{010}\}$ is available only as a three feature inventory – the feature f_{010} corresponding to the intersection of the other two features must be included explicitly.

Generally, any partition set that can be generated in a morphology that allows intersection can also be generated in a non-intersective one by explic-

²² More generally, only an inventory that contains at least n different features can describe the maximally differentiated n -cell paradigm if intersection is not available.

itly adding the features intersection would derive. In the three-cell case, this relationship holds also in the reverse direction as the comparison of Table 4 with 3 shows: any partition set that can be generated non-intersectively can also be generated in a system with intersective closure.²³

Returning to the *ABA constraint, and related considerations of restrictiveness, there are results that we believe should be of interest here. First, we note that admitting or banning intersection has no role on the overall impossibility of the four partition sets identified in section 5.3.1 as ungenerable. But rejecting intersection does make a difference in the definition of minimal valid inventories. In (the top half of) Table 3, where intersection is admitted, there are exactly three minimally valid inventories, which are linear permutations of one another, all of which derive a partition set with exactly three members. When intersection is not admitted, the number of minimally valid inventories increases to 25. Even so, the system is restrictive: without the Pāṇinian restriction, only 10 of 16 conceivable paradigm sets are generable.²⁴ None of the minimal valid inventories generates the unrestricted paradigm set (the rightmost column in Table 4) and none generates the partition set that excludes only AAA (column 12). Note that that pattern is derivable from non-minimal feature inventories, as indicated in the table. Hence it is Minimality that is playing a key role in excluding that partition set.

In other words, we include the following among our results. Regardless of whether intersection is admitted, and regardless of whether Pāṇinian ordering is enforced, the assumption of Minimal Validity as a condition on feature inventories ensures the following:

- (35) No minimal valid feature inventory for a 3-cell paradigm space includes all three bifurcations of the paradigm in its partition set.

These results amount to a generalization of the *ABA generalization up to linear permutation. A special case of (35) is the rightmost column of Tables 3 and 4: no minimal valid feature inventory generates an unrestricted partition set. The assumption of minimality always entails a restriction. Another special case is the implication that if any two bifurcations of the three-celled space are in the partition set of (minimal valid) inventory, then the third is not. Up

²³ For $n > 3$, there can be partition sets that can be generated only in non-intersective systems. One such system for $n = 6$ is derived from the features $f_{111100}, f_{001111}, f_{100000}, f_{010000}, f_{001000}, f_{000100}, f_{000010}, f_{000001}$. This set of features cannot generate the pattern AABCC, but its intersective closure can generate AABCC.

²⁴ The Pāṇinian restriction without intersection restricts this further to 7 possible paradigm sets, as seen in the line “3 features, Pāṇini” in Table 4.

to linear permutations of the cell orders, this is the *ABA generalization: if AAB and ABB are admissible paradigms, then ABA is not.

6 *ABA - empirical considerations

Coming up out of the heady sea of numbers for air, we are now at a point to step back and ask whether the results of our investigation of the formal combinatorics of features has any bearing on the actual *ABA generalizations discussed in the literature. Our tentative conclusion is that some domains where a *ABA generalization is observed do not seem to conform to the profile of the minimal valid inventory (with intersection), while for others, the situation is less clear, and the minimal valid inventory, with overlapping features, rather than containment, seems to us to be a direction worth pursuing.

We opened this article with reference to the *ABA generalization in adjectival gradation, investigated extensively in Bobaljik (2012). We see no reason from the discussion here to think that it would be profitable to reanalyze that as arising from a minimally valid 2-feature inventory. Doing so would invoke two privative features, one shared by the positive and comparative grade (but not the superlative), and another shared by the comparative and superlative, but not the positive. There is, however, fairly extensive evidence independent of patterns of suppletion for a containment relation in adjectival gradation: the superlative transparently contains the comparative in many languages.²⁵ Some examples are given here (from Bobaljik 2012:31):

(36)

	POS	CMPR	SPRL	
a. Persian:	kam	kam -tar	kam -tar-in	‘little’
b. Cimbrian:	šüa	šüan -ar	šüan -ar-ste	‘pretty’
c. Czech:	mlad-ý	mlad-ší	nej- mlad-ší	‘young’
d. Hungarian:	nagy	nagy -obb	leg- nagy -obb	‘big’
e. Latvian:	zil -ais	zil -âk-ais	vis- zil -âk-ais	‘blue’
f. Ubykh:	nüs [◦] ə	ç’a- nüs [◦] ə	a-ç’a- nüs [◦] ə	‘pretty’

In addition, it is not at all obvious that it makes sense to consider adjectival degrees as grammatical features, in the way that, for example, classificatory elements such as gender are.

²⁵ Recall that the inventory that respects containment is a minimal valid inventory if feature intersection is not permitted, as discussed in 5.4. As noted there, no inventory at all, not even the containment one, yields the actual profile seen in adjectival suppletion, where the partition set is exclusively {AAA, ABB, ABC} excluding both ABA and AAB.

On the other hand, there are other domains in which *ABA generalizations have been observed, where there is less independent reason to think that the constituent elements are arranged in a containment relation.

One such domain, perhaps, is person. Vanden Wyngaerd (2016) sees a *ABA generalization in (plural) independent pronouns. Building on prior cross-linguistic investigations (Cysouw 2003; Baerman et al. 2005), he observes that there are languages where first and second (plural) pronouns are syncretic, contrasting with the third person (such as Slave, in (37), from Cysouw 2003:124), and there are languages where second and third (plural) are syncretic, contrasting to the first person (as in the Nez Perce ‘unmarked’ pronouns in (38), Cysouw 2003), but virtually no good examples of syncretism of first and third person, contrasted with second.²⁶

		SG	PL
(37)	1	sɨ	naxɨ
	2	nɨ	naxɨ
	3	ʔedɨ	ʔegedɨ

		SG	PL
(38)	1	'iin	núun
	2	'im	'imé
	3	'ipí	'imé

Vanden Wyngaerd (2016) argues for a containment relation among the features that define person, as in the following:²⁷

- (39) a. 1st: [[[PERSON] PARTICIPANT] AUTHOR]
 b. 2nd: [[PERSON] PARTICIPANT]
 c. 3rd: [PERSON]

In our terms, this is (a linear permutation of) the inventory in (7): $f_{100}, f_{110}, f_{111}$ and its properties are well understood. However, there are few, if any, languages in which such a decomposition of pronouns is transparently manifest in surface forms. As we have seen above, this inventory is valid, but non-minimal.

²⁶ In bound person marking (agreement) more patterns are attested, though of varying frequency (Cysouw 2003, 2010; Baerman et al. 2005). The asymmetry whereby 1-3 syncretism is rarer than the other two combinations is generally supported in these studies: see Ackema & Neeleman (2013) for discussion, but see Chapter 1 of Harbour (2016) for important reservations.

²⁷ This is one of the current prominent views about the decomposition of person features in the literature; see for example: Sauerland (2008); Zeijlstra (2015). For contrasting views, see Bobaljik (2008); Ackema & Neeleman (2013); Harbour (2016).

A Minimal Valid inventory would be one that composes the three persons out of two privative features: f_{110} corresponding to the feature ‘participant’, and f_{011} , which is in essence the privative feature ‘non-author’. On this alternative analysis, first and third person pronouns cannot be syncretic, excluding the second person, since they share no feature. Hence *ABA. Ackema & Neeleman (2013:905) offer an analysis in essentially these terms, motivated in large part by the patterns of syncretism. As noted above, Ackema and Neeleman’s approach to features treats them as functions that operate on a set of discourse referents, but the key point is the proposal that first and second person share a feature, as do second and third person, but first and third do not.²⁸

In work in progress (see Sauerland & Bobaljik 2013) we are exploring the typology of syncretism in person feature inventories more broadly, drawing on the extensive data in Cysouw (2003), to determine what feature inventory assigns a high likelihood to a pattern like the observed partition sets, not just in plural pronouns, but in the full range of person marking systems, including clusivity distinctions. We may wager that if we are right to suspect a minimal valid inventory at work in the patterns of syncretism in the free-standing pronouns, then we will see that emerge as well in the larger study.

Before closing, we note as well that *ABA generalizations have also been noted in verbal inflection (Wiese 2008; Starke 2009), case (Caha 2009; Smith et al. 2016), and number (Smith et al. 2016). Of these, case is another domain in which there is little independent morphological evidence for containment relations, at least among ‘core’ cases.²⁹

Pavel Caha (personal communication and 2017) calls our attention to at least one sub-part of the case hierarchy which appears to reflect the kind of feature structure we would expect on the approach taken here. Blansitt (1988) surveys the marking of the following four functions across the world’s languages: direct object, dative (recipient), allative (goal of motion), and location. Blansitt notes a generalization, exceptionless in transitive clauses, whereby no two functions are marked identically unless all intervening func-

²⁸ Ackema & Neeleman (2013:925) note that there is a sense in this perspective in which the second person is the most “marked” person – the first and third are each defined by a single feature, while the second person is the intersection of two features. They suggest that evidence from the acquisition of agreement supports this view, inasmuch as second person agreement forms are often acquired last among the persons.

²⁹ Caha (2009); Smith et al. (2016) report some examples of, e.g., dative built on accusative, etc., but these are surprisingly rare, in contrast to, e.g., what we find with adjectival gradation. For spatial/locative cases, there is a much richer amount of transparent embedding; see Comrie & Polinsky (1998); Radkevich (2010); Pantcheva (2011). See Zoppi (2017) however for other arguments that the dependent case hierarchy can profitably be understood in terms of containment.

tions in the order just given are also marked identically. In other words, a *AB(B)A generalization. One way to approach this, following Caha (2009) (but see also Caha & Pantcheva 2012), would be to assume that there is a monotonic containment relationship among the features (we consider the last three for ease of exposition):

- (40) a. f_{111} = dative
 b. f_{011} = allative
 c. f_{001} = locative

An alternative, following the approach laid out here, would be the minimal valid inventory in (41):

- (41) a. f_{110} = dative
 b. f_{011} = locative

From this inventory, the allative can be described as the intersection of the other two cases. As Caha notes, Blansitt offers at least one language that seems to transparently reflect (41) rather than (40). Tigrinya prepositions include *ne* which marks dative (and some objects, presumably an instance of differential object marking, which quite commonly uses the dative, Bossong 1985) and locative *ab*. The allative is marked by the conjunction of the two: *nab* < *ne ab*. This is also broadly consistent with the results of Radkevich (2010) who found no evidence of a simple, monotonic transparent relationship among local cases as (40) might predict (although her survey also finds cases of portmanteaus and internally complex case morphology that are equally hard to reconcile with (41)).

7 Conclusion

In this paper, we introduced a notation for approaching feature logic from an algebraic perspective, abstracting away both from any empirical consideration and from any assignment of particular meanings to the features. Features are merely names for addresses (cells or groups of cells) in a list. In this way, we provided a calculus by which one can derive the paradigm set corresponding to any inventory of features, under varying sets of assumptions. This has two benefits. In the first place, we can investigate the formal properties of adding or subtracting individual assumptions, translating competing approaches into a common notation and working through the consequences at a formal level. The size of the partition set derivable from any inventory serves as a measure

of restrictiveness—combinations of assumptions that decrease the number of partition sets are more restrictive.

We have shown a number of results that are, we hope, of potential interest regarding three-celled paradigms. One of these is that certain partition sets are indescribable—no inventory of features yields exactly these partition sets without further assumptions. This group includes the set that has only the maximally differentiated and undifferentiated partitions: AAA, ABC (column 5 in Tables 3 and 4), as well as the three that allow only one of the three bifurcations in addition.

Another result arises from the assumption that feature sets must be minimal. With that assumption, a variety of generalized *ABA-like constraints are derived, among which the actual *ABA generalizations appear to be a special case. From this basic result, further restrictions are obtained by adding in the assumption that intersection is permitted (reducing the space of possibilities from 25 inventories and 10 paradigm sets to 3 inventories deriving 3 paradigm sets).

The effect of imposing Pāṇinian ordering as a condition on grammars (models) was also considered. With feature intersection, it proved overly drastic, excluding syncretism from the minimal valid inventories, but without feature intersection, the Pāṇinian restriction was weaker, excluding 3 of 10 paradigm sets admitted by intersection-free minimal valid inventories that incorporate the possibility of extrinsic order.

One specific result of interest to the study of *ABA generalizations is that the containment relationship among features, which is standardly invoked in accounts of *ABA generalizations in the literature turns out to be not only not the only type of inventory that can explain the generalization, but in fact, under the assumption that intersection is permitted, also not one of the minimal valid inventories.

In work in progress, we investigate additional extensions of the considerations presented here. In Appendix A, we begin the process of looking at larger paradigm spaces. As paradigms grow, the considerations become more intricate, but there may still be ways in which the minimal valid inventory stands as a contender for imposing restrictions that map to observed typological generalizations. In Sauerland & Bobaljik (2013), we note that for the four-cell paradigm space corresponding to the first person (inclusive vs. exclusive \times singular vs. plural), 9 of the 15 possibilities are indeed attested (Cysouw 2003). The four cell space can be described as two intersecting binary features, but that inventory is not minimal. Rather, using intersection, a more minimal inventory is the one containing the three features in (42) (and thus allowing the intersection of the first two in the rules of exponence):

- (42) a. f_{0101}
 b. f_{0011}
 c. f_{1111}
 d. $(f_{0101} \cap f_{0011} = f_{0001})$

While eschewing binary features, this yields a partition set that contains 9 of the $B_4 = 15$ logically possible partitions of the four cell space (this corresponds to the first line of the third block in Table 7). If we map the lists in the partition set to a binary table as in (23), we may observe that this partition set contains partitions corresponding to horizontal and vertical syncretisms, but no diagonal syncretisms. In Sauerland & Bobaljik (2013), we reached the conclusion on independent grounds that this was the optimal analysis of the first person paradigm space, i.e., the inventory that yields the best fit to the observed distribution of paradigms as documented in Cysouw (2003), while minimizing the incidence of accidental homophony. We turn to more discussion of larger paradigm spaces in Appendix A, below.

Without probing deeper, we hope to have shown that the derivation of *ABA generalizations entertained here may indeed get off the ground in some domains, leaving for future work the fuller empirical investigation of this approach.

Finally, returning to the question we raised at the outset, we may step back even further and ask why UG might have the types of constraints it does. We are obviously far from an answer, but can add a few, very tentative remarks here.

To this point, we have assumed that it is reasonable to think that UG feature inventories respect a condition of Minimality, and have shown how this assumption restricts the hypothesis space to be considered in determining the actual feature inventory corresponding to paradigms of a given size. Minimality has a somewhat different flavour than some of the other restrictive assumptions we have entertained. In principle, one could think of this from a different perspective. Rather than imposing a condition of Minimality on inventories, one could imagine instead that the features are whatever they are, but that UG shows maximal use of the features it has. For a domain with two features, UG generates in principle a three-celled space: each feature on its own, plus their intersection. This builds in the assumption of minimality - and thus means that all true three-celled paradigms are those projected from the two-feature inventories, yielding the *ABA prediction (up to linear permutation).

This alternative (maximal use of minimal resources), implies that there should be no four, or five-celled paradigms. If there are two features (in a given

domain) then the maximal paradigm in that domain will have three cells. If there are three features, then the paradigms generated will have 7 cells. The appearance of a four-celled paradigm in some domain then necessarily involves syncretism.

This concludes the discussion of 2 and 3-cell paradigm spaces, and the connection with the *ABA patterns. As an appendix, we turn to a rather less in-depth investigation of the effects of the assumptions here on larger paradigm spaces, notably 4-cell paradigms.

A Appendix: Beyond three cells

As we move beyond a three-cell space, the system grows and changes in various ways. Unlike the two- and three-cell spaces, we will not walk through all examples in as much detail, but will call attention to various points and provide some more general discussion in the abstract.

A three-cell paradigm space allows for $2^3 - 1 = 7$ features, and thus $2^7 = 128$ logically possible inventories. We showed in the previous section that Validity reduces that set to 96 and Validity plus Minimality further reduces the number of contenders to 3 feature inventories. By countenancing arbitrary ordering of the rules of exponence (sequences), each of the three minimally valid inventories generates 3 of the $B_3 = 5$ partitions of a 3-member set: none derives the undifferentiated partition. Adding Pāṇini as a restriction reduces that further: from the three minimal valid feature inventories, only one partition can be generated—the maximally differentiated partition.

The four-cell case allows $B_4 = 15$ partitions (paradigms) including the maximally differentiated partition. Logically, $2^{B_4-1} = 16384$ different valid partition sets exist. If there were no constraints on feature inventories or rules of exponence, this is the number of possible states of the world that we could investigate typologically, to ask which is the actual state. This number is so large that the answer to the question which of these subsets can be derived from how many valid feature inventories cannot be visualized in the same way as in Table 3. To be able to still compare the three and the four cell results, we report therefore comparable summary statistics for both the three cell case in Table 5 and the four cell case in Table 6. We report here only the total count of different ordered (O) and Pāṇinian (P) partition sets using the abbreviations *OPS* and *PPS*: we define the $OPS(I)$ for an inventory of features I as the set of partitions such that the valid, ordered feature sequences drawn from the intersective closure of F can generate all and only the partitions in $OPS(I)$. Similarly, the $PPS(I)$ is the set of partitions such that the valid, Pāṇinian fea-

ture sequences drawn from the intersective closure of F can generate all and only the partitions in $\text{PPS}(I)$.

In the three cell case, the values shown in Table 5 can be mostly read off from the table in 3. The first column shows the size of the sets of features under consideration, ranging from the minimum number such that there is a valid feature set of that size to the size of the set of all features, i.e. $2^c - 1$ where c is the number of cells. The second column shows the number of subsets of that size whether valid or not, i.e. the binomial coefficient $\binom{2^c - 1}{\# \text{features}}$. The next column lists how many of these features sets are valid inventories. This corresponds to the sum over the values of the corresponding row in Table 3 – recall that valid feature sets are valid regardless of whether order is extrinsic or intrinsic. The penultimate column indicates how many different OPSs can be derived from the valid inventories. This corresponds to the number of non-zero entries in the corresponding row of Table 3. The last column shows the number of PPSs derivable from the valid inventories. The last row of Table 3 gives the values for feature sets of any size. The number of inventories and that of valid inventories in the last row are the sum of the entries for feature sets of a specific cardinality because there can't be any overlap. But the OPSs and PPSs can overlap – a paradigm set that is derivable from a three feature inventory can in some cases also be derived from a four feature inventory. Therefore the values in the last row of the #PPS and #OPS columns aren't the column sums.

#features	#Inventories	#validInventories	#OPS	#PPS
0	1	0	–	–
1	7	0	–	–
2	21	3	3	1
3	35	29	12	9
4	35	35	11	11
5	21	21	5	5
6	7	7	2	2
7	1	1	1	1
any	128	96	12	12

Table 5 Summary statistics of logically possibly three cell partitions

For the four-cell paradigm space, the picture changes drastically. This is largely because describing a four cell space requires at least three features (as

we show below). The table corresponding to Table 5 for four cells is shown as Table 6. We computed the table using an Apple Macbook Pro Laptop computer (2012 model) running a program in the R programming language (R Core Team 2012) that tests all the feature sets for validity (computer code available, Sauerland 2017). The total computation time amounted to several hours.

#features	#Inventories	#validInventories	#OPS	#PPS
0	1	0	–	–
1	15	0	–	–
2	105	0	–	–
3	455	140	116	47
4	1 365	1 015	317	239
5	3 003	2 793	347	402
6	5 005	4 935	310	420
7	6 435	6 425	240	369
8	6 435	6 435	160	279
9	5 005	5 005	99	193
10	3 003	3 003	54	112
11	1 365	1 365	25	61
12	455	455	12	24
13	105	105	6	6
14	15	15	2	2
15	1	1	1	1
any	32 768	31 692	361	463

Table 6 Summary statistics of logically possibly four cell partitions

The results for the four-cell systems differ noticeably from the three-cell case in terms of the effect of the Pāṇini assumption and the overall restrictiveness. The biggest difference is that in the three cell case the minimal inventories contain 2 features, but for four cells a least 3 features are required. Validity and Minimality are about equally restrictive, but in relation to the 256-times greater set of feature sets. This leaves 140 minimal valid feature inventories. Unlike the three-cell space, where adding Pāṇini (assuming closure under intersection) eliminated the possibility for variation, in the four-cell space, Pāṇinian sequences allow 47 distinct possible partition sets, each of which generates between 1 and 7 of the $B_4 = 15$ distinct partitions of

a 4-celled space. None generate more than 7 partitions. The occurrence of the number 47 – a prime number that is not a Mersenne prime – in the table is, we think, noteworthy: It might indicate that even though a Pāṇinian sequence is a straightforward mathematical concept, the concept leads to some irreducible formal complexity.³⁰ We include a display of these 47 paradigm sets as Appendix B.

A second main difference concerns the relation between the order-sensitive partition sets (OPSs) and Pāṇinian partition sets (PPSs). In the three cell case, the OPS (top half of Table 3) and PPS (bottom half of the table) were overall the same. But with four cells, sequences with extrinsic order and Pāṇinian sequences predict different restrictions on the paradigm sets. As shown by the last line of Table 6, the Pāṇini assumption allows slightly more variation – 463 paradigm sets are generable with Pāṇinian sequences, while only 361 are compatible with extrinsically ordered systems.³¹ We furthermore computed the amount of overlap between the OPSs and PPSs. We found that the union of OPSs and PPS contains 557 partition sets, and therefore 267 are both OPS and PPS, 94 are only OPSs, and 196 are only PPSs. Some examples are shown in Table 7: the first three rows show three of the 267 partition sets that are both OPSs and PPSs. The second block of three rows shows three that are only PPSs, and the third block shows two that are only OPSs. The last block illustrates with two examples the 15 827 partition sets that cannot be generated by a feature-based morphological system unless additional restrictions are introduced.

Overall, feature-based analyses are surprisingly restrictive in the four-cell case. Of the 16 384 logically possible subsets of paradigms, only fewer than 3.4% (557 : 16 384) are compatible with a feature-based approach, 15 827 logically conceivable typological states are ruled out.³² Pāṇini's assumption is

³⁰ Work in progress of ours in collaboration with Tom Green supports this conclusion.

³¹ It may seem counter-intuitive that systems that allow any order of rules generate less variation than systems which are limited only to Pāṇinian rule orders. In terms of individual partitions, since Pāṇinian orders are a (proper) subset of the orders that are storable as extrinsic orders, Pāṇinian systems will not generate fewer partitions than those with unrestricted ordering. But what we count here is not the partitions that are generated, but the partition sets. In general, allowing extrinsic order will mean that the partition set generable from a given inventory will be a superset of the set generable from the same inventory with only Pāṇinian order. The numbers above reflect the effect that there is more convergence among the larger sets: with unrestricted rule ordering, more distinct inventories converge on the same partition sets than in the case of restricted rule ordering, overall.

³² One reviewer points out that permutations among the cells of a paradigm lead to multiple counting of equivalent patterns as was shown by the column grouping in Table 3 above. Since taking equivalence under permutation into account would reduce both the number of logical possibilities as well as that of OPSs and PPSs, we expect that it would not affect the




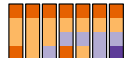

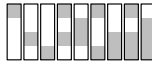




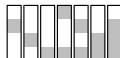
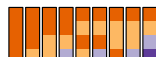



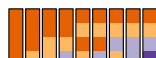

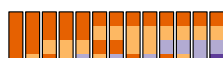
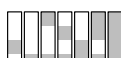

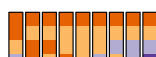
partition set	features (with order)	features (Pāṇini)
	1: 	57, e.g. 
	88, e.g. 	90, e.g. 
	1024, e.g. 	128, e.g. 
	0	32, e.g. 
	0	2, e.g. 
	0	128, e.g. 
	2, e.g. 	0
	12, e.g. 	0
	0	0
	0	0

Table 7 Examples of four-cell partition sets with the number of ordered and Panini analyses and examples of feature inventories (assuming intersective closure)

even slightly more restrictive, allowing fewer than 2.9% (463 : 16384) of all possible sets of partitions. This is surprising since the feature assumption and also the Pāṇini assumption only ruled out 4 of 16 paradigm sets in the 3 cell case, i.e. 75% of the paradigm sets are compatible. Even assuming minimal Pāṇini-systems turns out to be more restrictive in the four cell than in the three cell case, allowing fewer than 0.3% (47 : 16384) of the paradigm sets with four cells compared to 1 of 16 with three cells. The numerical comparisons need to be taken into account when evaluating feature-based morphological theories against typological data since the more restrictive a constraint is the less likely it is that is satisfied by chance.

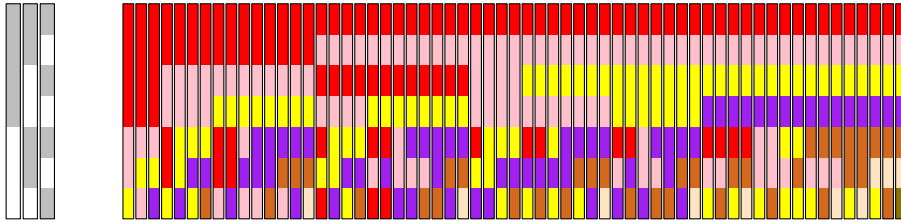
In many ways, the constraints imposed by feature inventories on universal grammar are at this point difficult to grasp concretely, though future formal work may find a way to do so. We explore therefore a four cell version of the *ABA constraint, *ABBA (see Caha 2009; Smith et al. 2016 for empirically-based proposals that exclude this pattern). There are two other permutations of the ABBA partition, ABAB and AABB. We investigated therefore which partition sets containing all three of ABBA, ABAB and AABB, as well as ABCD, can be generated from ordered and Pāṇini inventories. There are 2^{11} , i.e. 2048 such partition sets, but only 41 of them can be generated by feature inventories. Of those, 21 can be generated with either ordered or Pāṇinian systems, and a further 10 each with either only an ordered or only a Pāṇinian system. The examples in the third and fifth row of Table 7 illustrates two such systems. The system in row three predicts 13 partitions and can be generated from only three features if order is allowed. But if Pāṇini's restriction is obeyed, at least six features are necessary to generate this system. Row 5 shows an ABBA-system with the smallest number of features required to be generated in a Pāṇinian way, namely 4. So in the Pāṇini case, ABBA-systems require more than the minimal number of features as we also found for ABA-systems.

To conclude this section, consider briefly the seven cell paradigm. Because both 3 and 7 have $2^n - 1$ cells for some n , the structures of the two paradigms are similar. The minimal number of features of a valid inventory for a seven cell paradigms is three. Any set of three features f_a , f_b , and f_c where each basic feature contains 4 cells, each conjunction of two features contains 2, and the conjunction of all three features contains 1 cell is valid. There are $7! = 5040$ such systems, and all have the same structure except for permutations of the cells. One system is shown in (43) along with its order-sensitive partition set (OPS) – like in the 3-celled paradigm space, only the maximally differentiated

overall picture very much. Since the permutation-dependent calculations that we report are easier to verify, we decided to present those here.

paradigm is possible with Pāṇini sequences. It is noticeable that of the 61 generable partitions in the OPS none contains only A and B – only paradigms with at least a three-way partition are possible. Since all OPSs are identical except for cell-permutations, we can derive in this way a general *AB constraint for seven cell paradigm spaces. I.e., if there are 7-celled spaces, then they cannot be limited only to bipartitions.

(43)



Zentrum Allgemeine Sprachwissenschaft in Berlin. We are particularly grateful to Tom Green for his contributions to our understanding of the combinatorics more details of which still remains to be published. This work has been financially supported in part by the NSF (grant BCS-0616339), the Alexander von Humboldt-Stiftung (Bessel-award Bobaljik), and by the Bundesministerium für Bildung und Forschung (BMBF) (Grant 01UG1411).

Competing Interests

In appendix A, the authors assert that the occurrence of the value 47 as the number of complete paradigm sets derivable from minimal Pāṇinian inventories is surprising (see also appendix B). The authors may have been biased because at the same time of writing both were of that age. Other than this, the authors declare no competing interests.

References

- Ackema, Peter & Ad Neeleman. 2013. Person features and syncretism. *Natural Language and Linguistic Theory* 31. 901–950.
- Adger, David & Peter Svenonius. 2011. Features in Minimalist syntax. In Cedric Boeckx (ed.), *The Oxford handbook of linguistic minimalism*. Oxford: Oxford University Press.
- Albright, Adam & Eric Fuß. 2012. Syncretism. In Jochen Trommer (ed.), *The morphology and phonology of exponence*, 326–287. Oxford: Oxford University Press.
- Baerman, Matthew, Dunstan Brown & Greville G. Corbett. 2005. *The syntax-morphology interface: a study of syncretism*. Cambridge: Cambridge University Press.
- Benveniste, Émile. 1956. La nature des pronoms. In Morris Halle, Horace G. Lunt, Hugh MacLean & Cornelis H. van Schooneveld (eds.), *For Roman Jakobson: Essays on the occasion of his sixtieth birthday*, 34–37. The Hague, Netherlands: Mouton.
- Blansitt, Edward L. 1988. Datives and allatives. In Michael Hammond, Edith A Moravcsik & Jessica Wirth (eds.), *Studies in syntactic typology*, 173–191. Amsterdam: Benjamins.
- Bobaljik, Jonathan David. 2002. Syncretism without paradigms: Remarks on Williams 1981, 1994. In Geert Booij & Jaap van Marle (eds.), *Yearbook of morphology 2001*, 53–85. Dordrecht: Kluwer.

- Bobaljik, Jonathan David. 2008. Missing persons: A case study in morphological universals. *The Linguistic Review* 25(1-2). 203–230.
- Bobaljik, Jonathan David. 2012. *Universals in comparative morphology: Suppletion, superlatives, and the structure of words*. Cambridge: MIT Press.
- Bobaljik, Jonathan David & Susi Wurmbrand. 2013. Suspension across domains. In Ora Matushansky & Alec Marantz (eds.), *Distributed Morphology today: Morphemes for Morris Halle*, 185–198. Cambridge: MIT Press.
- Bossong, Georg. 1985. *Empirische Universalienforschung. Differentielle Objektmarkierung in den neuiranischen Sprachen*. Tübingen: Narr.
- Caha, Pavel. 2009. *The nanosyntax of case*. Ph.D. thesis, CASTL Tromsø.
- Caha, Pavel. 2017. Deriving Blansitt's generalization by an overlapping decomposition: a case against the subset principle. *Glossa: a journal of general linguistics* (this volume, to appear).
- Caha, Pavel & Marina Pantcheva. 2012. Datives crosslinguistically. Unpublished handout CASTL.
- Chomsky, Noam. 1956. Three models for the description of language. *IRE Transactions on Information Theory* 2. 113–124.
- Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English*. New York: Harper and Row.
- Comrie, Bernard & Maria Polinsky. 1998. The great Dagestanian case hoax. In Anna Siewierska & Jae Jung Song (eds.), *Case, typology, and grammar*, 95–114. Amsterdam: Benjamins.
- Corbett, Greville G. 2008. Determining morphosyntactic feature values: the case of case. In Greville G. Corbett & Michael Noonan (eds.), *Case and grammatical relations: Papers in honor of Bernard Comrie*, 1–34. Amsterdam: Benjamins.
- Corbett, Greville G. 2010. Features: essential notions. In Kibort & Corbett (2010), 17–36.
- Cysouw, Michael. 2003. *The paradigmatic structure of person marking*. Oxford, UK: Oxford University Press.
- Cysouw, Michael. 2010. On the probability distribution of typological frequencies. In *The mathematics of language*, 29–35. Springer.
- Cysouw, Michael. 2011. The expression of person and number: a typologists perspective. *Morphology: Special Issue on the Morphosemantics of Agreement* 41(2).
- Gärdenfors, P. 2000. *Conceptual spaces. the geometry of thought*. Cambridge, MA: MIT Press.
- Graf, Thomas. 2017. Graph transductions and typological gaps in morphological paradigms. 15th Meeting on the Mathematics of Language (MoL 2017) (in print).

- Harbour, Daniel. 2008. On homophony and methodology in morphology. *Morphology* 18(1). 75–92.
- Harbour, Daniel. 2011. Descriptive and explanatory markedness. *Morphology* 21. 223–240.
- Harbour, Daniel. 2016. *Impossible persons*. Cambridge: MIT Press.
- Jakobson, Roman. 1936/1971. Beitrag zur allgemeinen Kasuslehre. Gesamtbedeutungen der russischen Kasus. In *Selected writings*, vol. 2, 23–71. The Hague: Mouton.
- Kibort, Anna & Greville G. Corbett (eds.). 2010. *Features: Perspectives on a key notion in linguistics*. Oxford: Oxford University Press.
- Kiparsky, Paul. 1973. “elsewhere” in phonology. In *A festschrift for Morris Halle*, 93–106. New York: Holt, Rinehart and Winston.
- Kiparsky, Paul. 1979. *Pāṇini as a variationist*. MIT Press.
- Kracht, Markus. 1997. Inessential features. In Christian Retoré (ed.), *Logical aspects of computational linguistics: First international conference, lacl '96, selected papers*, 43–62. Berlin: Springer.
- McCawley, James. 1974. The sound pattern of English [review]. *International Journal of American Linguistics* 40(1). 50–88.
- Pantcheva, Marina. 2011. *Decomposing path: The nanosyntax of spatial expressions*. Ph.D. thesis, Universitetet i Tromsø.
- Pertsova, Katya. 2011. Grounding systematic syncretism in learning. *Linguistic Inquiry* 42(2). 225–266.
- Pullum, Geoffrey K. & Hans-Jörg Tiede. 2010. Inessential features and expressive power of descriptive metalanguages. In Kibort & Corbett (2010), 272–292.
- R Core Team. 2012. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Radkevich, Nina. 2010. *On location: The structure of case and adpositions*. Ph.D. thesis, University of Connecticut.
- Sauerland, Uli. 2008. On the semantic markedness of ϕ -features. In David Adger, Susana Béjar & Daniel Harbour (eds.), *Phi theory: Phi features across interfaces and modules*. Oxford: Oxford University Press.
- Sauerland, Uli. 2017. Syncretism distribution modeling: R-code. URL <http://github.com/ulisauerland/sdm>.
- Sauerland, Uli, Jan Anderssen & Kazuko Yatsushiro. 2005. The plural is semantically unmarked. In Stephan Kepser & Marga Reis (eds.), *Linguistic evidence – empirical, theoretical, and computational perspectives*, 413–434. Berlin, Germany: Mouton de Gruyter. <http://dx.doi.org/10.1515/9783110197549.413>.

- Sauerland, Uli & Jonathan D. Bobaljik. 2013. Proceedings of GLOW in Asia 9. In Nobu Goto, Koichi Otaki, Atsushi Sato & Kensuke Takita (eds.), *Proceedings of GLOW in Asia*. Tsu, Japan: University of Mie.
- Smith, Peter, Beata Moskal, Ting Xu, Jungmin Kang & Jonathan David Bobaljik. 2016. Case and number suppletion in pronouns. Manuscript Goethe Universität Frankfurt, University of Connecticut, and Syracuse University.
- Starke, Michal. 2009. Nanosyntax: A short primer to a new approach to language. *Nordlyd* 36(1). 1–6.
- Stump, Gregory. 2016. *Inflectional paradigms: Content and form at the syntax-morphology interface*. Cambridge: Cambridge University Press.
- Trommer, Jochen. 2008. Third-person marking in Menominee. In Daniel Harbour, David Adger & Susana Béjar (eds.), *Phi theory: phi features across interfaces and modules*, 221–250. Oxford: Oxford University Press.
- Vanden Wyngaerd, Guido. 2016. The feature structure of pronouns: a probe into multidimensional paradigms. Unpublished manuscript CRISSP, lingbuzz/003166.
- Wiese, Bernd. 2008. Form and function of verb ablaut in contemporary standard German. In Robin Sackmann (ed.), *Explorations in integrational linguistics*. Amsterdam: John Benjamins.
- Zeijlstra, Hedde. 2015. Let's talk about you and me. *Journal of Linguistics* 51. 465–500.
- Zompì, Stanislao. 2017. *Case decomposition meets dependent-case theories*. Master's thesis, Università de Pisa.