Monotonicity as an effective theory of morphosyntactic variation

Thomas Graf[1]
Department of Linguistics, Stony Brook University, Stony Brook, USA

ABSTRACT

One of the major goals of linguistics is to delineate the possible range of variation across languages. Recent work has identified a surprising number of typological gaps in a variety of domains. In morphology, this includes stem suppletion, person pronoun syncretism, case syncretism, and noun stem allomorphy. In morphosyntax, only a small number of all conceivable Person Case Constraints and Gender Case Constraints are found. While various proposals have been put forward for each individual domain, few attempts have been made to give a unified explanation of the limited typology across all domains. This paper presents a novel account that deliberately abstracts away from the usual details of grammatical description in order to provide a domain-agnostic explanation of the limits of typological variation. This is achieved by combining prominence hierarchies, e.g. for person and case, with mappings from those hierarchies to the relevant output forms. As the mappings are required to be monotonic, only a fraction of all conceivable patterns can be instantiated.

1 INTRODUCTION

In physics, an *effective theory* describes the behavior of a system at a higher level of abstraction that does not necessarily reflect the true causal factors that give rise to the behavior. For example, the physical laws governing the behavior of gases are effective theories of a system whose causal factors reside at the much lower level of atoms and fundamental forces. An effective theory is often easier to understand than

the actual system, and it may furnish generalizations that are harder or impossible to state at a more fine-grained level of description.

The central goal of this paper is to develop such an effective theory for certain areas of morphology and morphosyntax that have attracted a lot of attention in recent years:

(1)   a. **The *ABA generalization in morphology**
- Stem suppletion in adjectival gradation
- Syncretism in person pronoun paradigms
- Syncretism in case paradigms
- Noun stem allomorphy

   b. **Morphosyntactic constraints on clitic clusters**
- Person Case Constraint (PCC)
- Gender Case Constraint (GCC)

In each domain, the goal is to explain why not all logically conceivable patterns are attested. For example, there are 64 logically possible PCC variants, but only a handful have been reported in the literature. Such seemingly arbitrary typological gaps demand a principled explanation, and the explanation should apply across as many domains as possible.

The explanation proposed in this paper consists of two components: a base hierarchy that captures certain prominence relations between the elements in a domain, and a mapping from each base hierarchy to the relevant output forms. Crucially, the mapping must be monotonic. The shape of the base hierarchy and the monotonicity requirement conspire to greatly limit the range of possible patterns.

The approach advocated here is strongly inspired by the mathematical formalism of Graf (2014, 2017) but improves on it in important respects. A broader range of data is considered, including a wide selection of case syncretisms and a new kind of PCC reported by Tyler (2017) for Choctaw. In addition, the analysis in terms of monotonicity greatly simplifies Graf's rather byzantine machinery (I am indebted to an anonymous reviewer of Graf 2017 for pushing me to explore monotonicity as a unifying principle). In contrast to generative accounts such as Anagnostopoulou (2005), Nevins (2007), Caha (2009), Bobaljik (2012), and Zompí (2016), the monotonicity approach provides a unified solution for all the phenomena above, rather than just one or two of them. This is because as an effective theory, my proposal

can focus on describing the general behavior of the system rather than how this behavior arises from the machinery of the grammar.

My proposal is close in spirit to Bobaljik and Sauerland (2018), who seek to derive the *ABA generalization from feature combinatorics without making specific reference to the content or denotation of these features. However, my account is less radical in its quest for content-agnostic explanations as each domain may come with its own base hierarchy. This makes the approach easier to apply to specific phenomena. But as it is still an open question what specific forms the base hierarchies may take, there is also a lot of room for overgeneration. The hierarchies proposed in this paper are largely in line with current linguistic thinking, but a tight mathematical characterization of the space of possible hierarchies is still missing. The ideas pursued in Bobaljik and Sauerland (2018) might actually turn out to be equivalent to specific restrictions on base hierarchies. So even though the two approaches differ a fair bit at this point and address slightly different questions, they are at the very least fellow travelers. In particular, both largely abstract away from feature systems and the specifics of the grammar and thus are not tied to any specific grammar formalism.

The paper proceeds as follows. Section 2 defines monotonicity and explains it in intuitive terms. No other mathematical concepts are needed for this paper. Sections 3 and 4 then present the account of the *ABA generalization and the PCC, respectively. I conclude the paper with some brief thoughts on the status of monotonicity in general (Section 5.1), the psychological reality of the monotonicity account and the mechanisms it posits (Section 5.2), and the empirical robustness of the approach in light of an impoverished data sample (Section 5.3).

2 MONOTONICITY: DEFINITION AND EXPLANATION

Even though monotonicity is a well-known concept of mathematics, I include a detailed explanation here to accommodate as large an audience as possible. Readers who are already familiar with monotonic functions can skip ahead to Section 2.2, where I introduce the notion *feasibly monotonic* as a minor generalization of monotonicity. This generalization step will simplify the discussion of the *ABA generalization

in Section 3. The analysis of PCC effects in Section 4 only needs the standard notion of monotonicity.

2.1 *Monotonic functions*

Monotonicity expresses whether a mapping between two objects respects their internal structure. Suppose we are given two structures $\langle A, \leq_A \rangle$ and $\langle B, \leq_B \rangle$ such that A and B are (possibly infinite) sets with respective order relations $\leq_A$ and $\leq_B$ defined over them. For example, $\langle A, \leq_A \rangle$ may be the set of natural numbers ordered by the less-or-equal relation, and $\langle B, \leq_B \rangle$ may be the set of Latin characters in alphabetical order. Then a function f from A to B is *monotonic* iff f preserves the relative order of elements. Sticking with our example of natural numbers and alphabet letters, a monotonic function must not map 10 to H and 100 to E because $10 \leq 100$ but $f(10) = H$ occurs after $f(100) = E$ rather than before it. However, the function may map all numbers between 0 and 10 to E and all other numbers to H, as this does not invert the original order.[1]

(2) **Monotonicity**
Given two sets A and B, let $\leq_A \subseteq A \times A$ and $\leq_B \subseteq B \times B$. A function $f : A \to B$ is *monotonic* with respect to $\leq_A$ and $\leq_B$ iff it holds for all x and y in A that $x \leq_A y$ implies $f(x) \leq_B f(y)$.

A linguistic analogy for monotonicity is the ban against crossing branches in autosegmental phonology (Goldsmith 1976). In autosegmental phonology, a phonological representation is not merely a string of segments, but instead consists of multiple tiers whose elements are connected by association lines. Each tier is still linearly ordered, though, and may be regarded as a string on its own. For example, a representation may consist of a string of segments, i.e. the segmental tier, and a string of tones, i.e. the tone tier. This is illustrated below with an example of tone association in Kikuyu.

---

[1] The reader may have noticed that this description of monotonicity only applies to monotonically increasing (or *isotone*) functions. A function can also be monotonically decreasing (or *antitone*). In this case, the order must be inverted: $x \leq_A y$ implies $f(y) \leq_B f(x)$. For the purposes of this paper, the distinction is immaterial because every isotone function from A to B is antitone from A to the dual of B. For instance, an isotone function from numbers to alphabetically ordered letters would be antitone if the letters are instead ordered in reverse.

(3)

t o m a r ɔ r i r ɛ

L   H   L   H
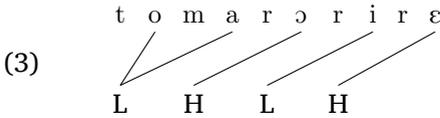
The ban against crossing branches ensures that segmental tier and tonal tier are synchronized to a certain extent. The linear order of tones must reflect the linear order of the segments they are associated to. Whenever this is not the case, some association lines illicitly cross each other as in the representation below.

(4)

t o m a r ɔ r i r ɛ

L   H   L   H
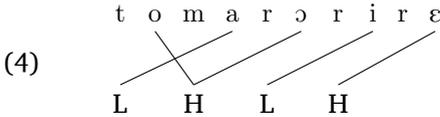
These are exactly the cases where the mapping from elements on the segmental tier to elements on the tone tier is not monotonic. In the case at hand, the segment o linearly precedes a, yet o is mapped to a high tone H that follows the low tone L that a is associated with.

Note that even though the examples above all involve linearly ordered structures, monotonicity is more general and can be evaluated for any arbitrary ordering relation. The example below depicts a mapping from a partially ordered structure **S** on the left to the algebra **2** of truth values on the right. This mapping is monotonic because there are no x and y such that $x \leq_S y$ yet $f(y) <_2 f(x)$. In particular, it is irrelevant for monotonicity that $f(2) <_2 f(3)$ because neither $2 \leq_S 3$ nor $3 \leq_S 2$ hold. If two elements x and y are unordered with respect to each other, the relative order of $f(x)$ and $f(y)$ is immaterial for monotonicity.

(5)

4

2   3

1

0

T

F
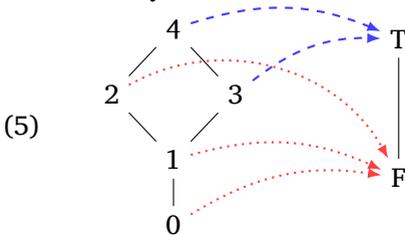
However, if the mapping is altered just a bit such that 1 is mapped to True instead of False, monotonicity is lost because then we have $1 <_S 2$ yet $f(2) = F <_2 T = f(1)$.

We will encounter both linearly and partially ordered structures in this paper. Linearly ordered structures are at the center of the *ABA-generalization for adjectival gradation and pronoun syncretism (Section 3.1). Partial orders, on the other hand, are indispensable for broadening the empirical scope to case syncretism (Section 3.3), noun stem allomorphy (Section 3.4), Person Case Constraints (Sections 4.2, 4.3), and the Gender Case Constraint (Section 4.4).

2.2 *Feasibly monotonic functions*

During the discussion of the *ABA generalization in Section 3, there will be some cases where the co-domain B does not have any natural order defined over it. Adjectival gradation, for example, involves two kinds of objects:

1. a set of adjectival degrees, i.e. {positive, comparative, superlative}, and
2. a set of surface realizations, e.g. {slow, slower, slowest}.

Whereas the former can be given a natural order in terms of semantics, the set of surface realizations lacks such an internal structure. There is no obvious ordering relation between these three phonological representations. One could put them in reverse alphabetical order, or line them up according to length or morphological complexity. For our purposes, the important thing is simply that some sufficiently strict order is defined over this domain so that monotonicity can be invoked. We make this requirement explicit via the notion of *feasible monotonicity*.

(6) **Feasible monotonicity**
Let A be a set ordered by $\leq_A \subseteq A \times A$, and B some arbitrary set. Then $f : A \to B$ is *feasibly monotonic* iff there is some linear order $\leq_B \subseteq B \times B$ such that f is monotonic with respect to $\leq_A$ and $\leq_B$.

(7) **Linear order**
A relation $\leq_B \subseteq B \times B$ is a linear order iff all of the following hold for all $x, y, z \in B$:

- *reflexivity*: $x \leq_B x$
- *antisymmetry*: $x \leq_B y$ and $y \leq_B x$ jointly imply $x = y$
- *transitivity*: $x \leq_B y$ and $y \leq_B z$ jointly imply $x \leq_B z$
- *totality*: $x \leq_B y$ or $y \leq_B x$

The figure below shows three mappings. The leftmost one is monotonic. The one in the middle is not monotonic, but it is feasibly monotonic because we can switch the order of B and C and thus obtain a monotonic mapping. The function on the right, on the other hand, is neither monotonic nor feasibly monotonic: no matter which order one picks for A, B, and C, two branches will always cross, indicating that the mapping is not monotonic.

(8)

1	2	3		1	2	3		1	2	3
A	B	C		A	B	C		A	B	C
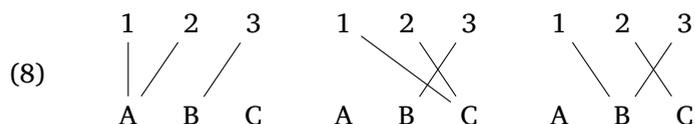
To sum up, monotonic mappings are order-preserving in the sense that they do not invert existing orderings: $x \leq_A y$ entails $f(x) \leq_B f(y)$. The notion of feasibly monotonic mappings extends this to cases where the co-domain lacks internal structure. It does so by considering all possible ways to order the co-domain such that feasible monotonicity holds iff monotonicity holds for at least one of those orders. The next section discusses the *ABA generalization as the first application of (feasible) monotonicity, with the PCC following in Section 4.

3   *ABA GENERALIZATION

The first part of the empirical analysis is devoted to the *ABA generalization. I initially limit myself to suppletion in adjectival gradation and pronoun systems (Section 3.1). Each domain involves only 3 cells, which simplifies the discussion. Both of them will be explained in terms of a linearly ordered based hierarchy – one for adjectival degrees, another one for person. Crucially, the mappings from these hierarchies to surface forms must be feasibly monotonic. This requirement severely restricts the range of possible suppletion patterns, providing a close fit for the attested typology and reducing the *ABA generalization to monotonicity.

I then expand this approach to larger, partially ordered hierarchies to account for case syncretism (Section 3.3) and noun stem allomorphy (Section 3.4). The general idea remains the same, though: once a suitable, linguistically motivated hierarchy has been fixed, the

range of cross-linguistic variation falls out from the limitation to (feasibly) monotonic functions.

As I explain in Section 3.2, the monotonicity program is still in its early stages, and as a result the construction of these hierarchies is primarily guided by empirical concerns. Conceptual or mathematical restrictions on the shape of hierarchies must be postponed until a larger class of hierarchies has been identified. That is not to say, though, that the hierarchies presented in this paper are completely arbitrary. They display many regularities, and are very natural from a linguistic perspective.

3.1 *3-cell paradigms: Adjectival gradation and pronoun allomorphy*

The *ABA generalization was formulated in Bobaljik (2012) and refers to a particular typological gap in numerous morphological paradigms. Given a morphological subsystem where one may posit an underlying hierarchy $x > y > z$, z cannot pattern with x to the exclusion of y.

The best-known example of the *ABA generalization is suppletion in adjectival gradation, which was analyzed at great depth in Bobaljik (2012). Bobaljik points out that if a language allows for stem suppletion in either comparatives or superlatives, it must allow for both. Data illustrating this generalization is given in Table 1. If one follows the convention to list the three forms in the order *positive*, *comparative*, *superlative* and uses letters to indicate which forms use the same stem, one can decompose the typological gaps into two constraints: *AAB and *ABA. The central puzzle is why these specific constraints should hold for adjectival gradation but not, say, *ABB or *ABC.

In Bobaljik (2012), the ban against ABA patterns is explained via structural mechanisms. Adjectival forms are decomposed into a tree
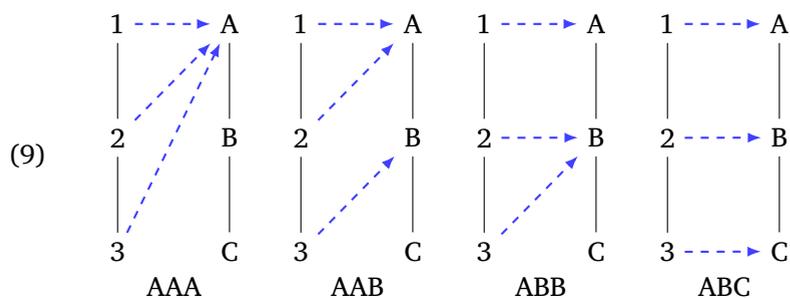
Table 1:
Examples of attested
suppletion patterns from
Smith *et al.* (2018)

Language	Positive	Comparative	Superlative	Pattern
English	smart	smart-er	smart-est	AAA
English	good	bett-er	be-st	ABB
Finnish	hyvä	pare-mpi	parha-in	ABB
Latin	bon-us	mel-ior	opt-imus	ABC
Welsh	da	gwell	gor-au	ABC
unattested	good	bett-er	good-est	*ABA
unattested	good	good-er	be-st	*AAB

template such that comparatives contain the positive base form as a subtree and are in turn themselves subtrees of the corresponding superlative forms. Then *ABA follows from specific assumptions about the rewrite rules that convert these tree structures into morphological surface forms. Bobaljik and Sauerland (2018) provide an alternative explanation grounded in the combinatorics of feature systems. Under both approaches *ABA falls out from the fact that it is impossible for a rewrite rule to target positive and superlative forms to the exclusion of the comparative. Both works also agree that *ABA is the more important constraint of the two – whereas *ABA holds for many morphological paradigms, *AAB seems to be specific to adjectival suppletion and requires additional stipulations.

The increased importance of *ABA relative to *AAB is noteworthy because the former can be explained in terms of monotonicity, but not the latter. Suppose that there is a universal underlying hierarchy of the form *positive > comparative > superlative*. For the sake of succinctness, I abbreviate this hierarchy as $1 > 2 > 3$. Now let $\{A, B, C\}$ be the set of possible surface forms. Irrespective of how this set is ordered, there can be no monotonic function f with $f(1) = f(3) \neq f(2)$. We already saw this in (8) at the end of Section 2.2. Hence no feasibly monotonic function over $1 > 2 > 3$ can produce a pattern of the form ABA, and consequently the *ABA generalization reduces to a ban against functions that are not feasibly monotonic.

But the other patterns AAA, AAB, ABB, and ABC can be produced by feasibly monotonic functions, as is shown below with an assumed ordering of $A > B > C$.

(9)

Some trivial variations are not depicted here, such as a function that maps 1, 2, and 3 to C instead of A. Keep in mind that no further assumptions are made about the exponents of A, B, and C, so it is not the

case that A is the fixed counterpart of a positive form or C the fixed counterpart of a superlative form. Instead, A, B, and C are just abstract variables or bins, and any two forms that are put in the same bin must have the same exponent. Consequently, a function mapping all three forms to A is empirically equivalent to one mapping all three forms to C. In this system, there are only five ways to map three different forms to exponents. But one of them is the illicit (and not monotonic) ABA pattern, so that (9) already exhausts the full range of options.

We see, then, that the typology of adjectival gradation is partially explained by the assumption that

1. there is a universal hierarchy *positive > comparative > superlative*, and
2. the mapping from this hierarchy to surface forms must be feasibly monotonic.

These two assumptions explain the absence of ABA patterns, but they still allow for AAB patterns, which are unattested cross-linguistically. Just like the previous analyses in Bobaljik (2012) and Bobaljik and Sauerland (2018), monotonicity cannot give a unified explanation of the absence of both ABA and AAB patterns.

However, this is actually a welcome state of affairs because AAB patterns do show up in other empirical domains. Harbour (2015) conducts an extensive survey of pronoun systems. His findings are summarized in Table 2. Putting aside number and the inclusive-exclusive distinction to focus exclusively on person specification, we can infer that all of the languages surveyed by Harbour adopt one of four person systems:

- all persons are the same (AAA),
- first and second person are the same (AAB),

Table 2:
Typology of pronoun systems
according to Harbour (2015, p. 137)

Pronominal contrasts	Language
none	Wichita
1|23	Damin, Elseng/Morwap[?]
12|3	Winnebago
1|2|3	Jarawa, Kiowa
1EX|1IN|2|3	Imonda, Matses, Waris
1|23 × SG|PL	Sanapaná

- second and third person are the same (ABB),
- all persons are different (ABC).

Again the ABA pattern is missing, and this fact is expected if all languages use an underlying person hierarchy of $1 > 2 > 3$ and the mapping to surface forms must be feasibly monotonic. At the same time, the four attested patterns AAA, AAB, ABB, and ABC are completely expected from this perspective. I thus conclude that monotonicity can successfully limit the possible range of variation in these morphological domains, although certain options such as AAB may still be unattested due to unrelated factors.[2]

3.2 *Motivating the hierarchies*

The advantage of the monotonicity approach lies in its ability to account for the *ABA generalization across seemingly unrelated domains. Person and adjectival gradation have nothing in common semantically, and it also seems unlikely that they are syntactically related. More specifically, I am not aware of any proposals where first person is structurally contained by second person which in turn is contained by third person, mirroring Bobaljik's treatment of positive, comparative, and superlative. But from the abstract high-level perspective advocated here, person and adjectival degrees are exactly parallel because their respective hierarchies are isomorphic. Each one is of the form $1 > 2 > 3$, and the only difference is what each element of the hierarchy denotes.

---

[2] An anonymous reviewer wonders how the person hierarchy can be extended to the exclusive-inclusive distinction without losing its explanatory force. The reviewer notes that it would be very natural to analyze 1IN as the combination of 1 and 2, such that one gets the ordering relations $1\text{IN} > 1 > 3$ and $1\text{IN} > 2 > 3$. But with this hierarchy, 1 and 3 can now be syncretic to the exclusion of 2, undermining the results of this section.

This case shows that the choice of hierarchies must be carefully guided by data. An empirically more adequate hierarchy might preserve $1 > 2 > 3$ while also adding $1\text{IN} > 1$ and $1\text{IN} > 2$. This would incorporate the natural idea that 1IN combines first and second person without loosening the relative order of 1 and 2. At this early stage of the monotonicity enterprise, though, the choice of hierarchy is primarily driven by empirical data, and without a careful analysis of this data all claims about the shape of hierarchies are highly speculative.

This raises the question, though, whether there is any independent motivation for these hierarchies beyond their central role in accounting for the data. There certainly is, but before discussing this in detail I would like to point out that every account of the *ABA generalization has to assume some base hierarchy for the domain in question, at least at a descriptive level. Otherwise, the *ABA generalization in its current form cannot be stated. Suppose that adjectival gradation patterns were by default listed in the order *comparative-positive-superlative*. An attested pattern like *good-better-best* would then be regarded as *better-good-best*, which is an ABA pattern. The discussion of the *ABA generalization thus presupposes an agreed-upon base order for every domain under investigation.[3] The monotonicity account simply takes this base order at face value and describes how the expected typology is narrowed down by restricting our attention to feasibly monotonic mappings from base hierarchies to surface forms.

Crucially, though, the hierarchies posited so far are highly plausible from a cognitive perspective. The hierarchy *positive > comparative > superlative* directly reflects the semantics of each form. The person hierarchy $1 > 2 > 3$, on the other hand, has already been argued for by Zwicky (1977) for entirely different reasons. This hierarchy is also implicit in feature-based systems such as that of Nevins (2007), where first person is [+author, +participant], second person is [−author, +participant], and third person is [−author, −participant]. Suppose we represent these specifications in privative terms as {author, participant}, {participant}, and {}, respectively. If one orders these sets by the superset relation, the very same ordering emerges as with Zwicky's person hierarchy. Admittedly, there are other well-known feature systems that give rise to a different ordering, e.g. Harley and Ritter (2002). The posited person hierarchy is also missing the crucial contrast between inclusive and exclusive. The current person hierarchy thus might present an overly simplified picture. But future refinements would only make the hierarchy an even closer

---

[3] This holds even for the account of Bobaljik and Sauerland (2018). While their mathematical analysis is order-independent and eliminates ABA patterns based on the shape of the feature system, the application to empirical data requires picking a suitable feature system. This is tantamount to positing a base hierarchy for the empirical domain under investigation.

match for current linguistic thinking. Overall, then, the hierarchies for person and adjectival gradation are both on linguistically solid ground.

Ultimately, the monotonicity approach must furnish a restrictive theory of linguistic hierarchies lest it devolve to a purely descriptive enterprise where hierarchies are tuned and tweaked until monotonicity yields the desired result. But these restrictions cannot be put in place *a priori*. They must be inferred by defining empirically adequate hierarchies for a wide range of phenomena and by isolating properties that separate these hierarchies from conceivable alternatives that produce undesirable data patterns (cf. fn. 2). This bottom-up strategy is a major methodological difference to Bobaljik and Sauerland (2018), who start out with abstract yet linguistically natural restrictions on feature systems and use those to derive the absence of ABA patterns. Such a top-down strategy could also be applied to the monotonicity approach, but I believe that a largely data-driven approach will prove more fruitful for a nascent enterprise like this. Without a rich body of well-established hierarchies, the best option is to craft restrictive hierarchies to fit the data and evaluate their linguistic plausibility. As the number of hierarchies grows, their shared properties will become more apparent and serve to constrain the shape of newly posited hierarchies.

It is also of interest in this connection how the hierarchies relate to linguistic assumptions about feature systems or structural projections. A clearer understanding of this link would make it easier to convert assumptions about linguistic feature systems into constraints on hierarchies. I have already hinted at such a connection between feature systems and hierarchies in my brief discussion of Nevins (2007) and its strong correspondence to the person hierarchy of Zwicky (1977). In this particular case, the relation is easy to discern thanks to the simple nature of both the person hierarchy and the feature system. But as we will see in the remainder of this paper, other empirical domains require much more elaborate hierarchies whose connections to features or projections from the linguistic literature is much less clear. The very next phenomenon, case syncretism, is already a striking example of the complexity of hierarchies.

3.3   *Moving beyond 3 cells: Case syncretism*

Even though adjectival gradation and pronoun allomorphy pertain to vastly different morphological domains, they are both similar in that their paradigms distinguish only three cells: positive-comparative-superlative for the former, first person-second person-third person for the latter. Many paradigms, however, involve more than three cells. Case is a prime example of this, with many languages distinguishing at least four different cases. It will be interesting to see if monotonicity still holds in these larger paradigms, and if so, what shape the relevant hierarchies have.

Caha (2009, 2013) provides a detailed study of case syncretism, i.e. which cases in a noun inflection paradigm may systematically display the same surface form. Such syncretisms are common across languages with multiple morphologically realized cases, e.g. Russian (Caha 2009, p.12).

(10)

	window (sg)	teacher (pl)	100
Nom	okn-o	učitel-ja	st-o
Acc	okn-o	učitel-ej	st-o
Gen	okn-a	učitel-ej	st-a
Loc	okn-e	učitel-jax	st-a
Dat	okn-u	učitel-am	st-a
Inst	okn-om	učitel-am-i	st-a

The first column shows syncretism of nominative and accusative. In the second column, accusative and genitive are syncretic. The third column displays two syncretisms, nominative-accusative on the one hand and genitive-locative-dative-instrumental on the other. With six cases, there are 203 logically conceivable patterns, but only a fraction of those are attested. As it it unlikely that all these typological gaps are purely accidental, a more principled explanation of this limited typology is needed.

Even though Caha's primary concern is to accommodate the typological facts in the framework of nano-syntax, he first formulates a purely descriptive universal. His *strong case contiguity hypothesis* limits case syncretism to contiguous areas of Blake's case hierarchy (Blake 2001):

(11) **Blake's case hierarchy (strict version)**

Nom > Acc > Gen > Dat > Inst > others

This means that a language may mark, say, accusative, genitive, dative and instrumental the same, but not accusative and instrumental to the exclusion of dative and genitive. In other words, the strong case contiguity hypothesis extends the *ABA generalization beyond systems with three-way contrasts.

The strong case contiguity hypothesis is yet another instance of monotonicity. Any feasibly monotonic function can only map continuous parts of Blake's case hierarchy to the same exponent. If such a function mapped, say, accusative and dative to A but genitive to B, then we would have both $f(\text{acc}) \leq f(\text{gen}) \leq f(\text{dat})$ and $f(\text{acc}) = f(\text{dat}) \neq f(\text{gen})$, which is impossible. Hence (feasible) monotonicity over Blake's case hierarchy rules out case syncretisms of the ABA-variety.

Curiously, though, such ABA-style case syncretisms do exist. Harðarson (2016) points out that accusative and dative are frequently syncretic to the exclusion of the genitive in Germanic languages. For the monotonicity account, the only way of incorporating this fact is to change the case hierarchy. By relaxing Blake's hierarchy such that genitive and dative are unordered with respect to each other, we can keep all the syncretisms of the original hierarchy while also allowing for accusative-dative syncretism.

(12) **Blake's case hierarchy (relaxed version)**

```
        Nom
         |
        Acc
       /    \
   Gen        Dat
       \    /
        Inst
         |
       Others
```
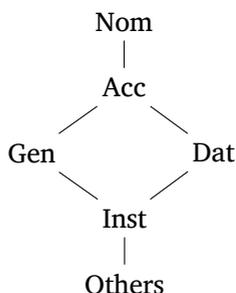
This is the first instance where I have to posit a hierarchy that is only partially ordered. But since monotonicity is not limited to linear orders (Section 2.1), the step to partial orders for base hierarchies is a natural one. With the hierarchy in (12), dative and genitive can still

be syncretic because they are unordered with respect to each other. Recall from Section 2.1 that monotonicity only limits the possible values for ordered elements, which entails that syncretism of these two unordered cases cannot violate monotonicity. At the same time, accusative still cannot be syncretic with instrumental to the exclusion of dative or genitive as this would violate the ordering relations established by the hierarchy.

However, such accusative-instrumental syncretism are actually attested, which means that the current hierarchy is still too restrictive. In fact, the range of attested case syncretisms goes far beyond what our relaxed hierarchy allows for. In a painstaking literature survey, Zompí (2016) has compiled an extensive list of case syncretisms across numerous typologically diverse languages, including languages with an ergative-absolutive system and even nominative-ergative-absolutive systems. His findings are summarized in Table 3 and include many syncretisms that are unexpected even with the relaxed version of Blake's hierarchy in (12).[4]

Zompí (2016) argues for a radically simplified case hierarchy to account for the permissive typology. Cases come in three types: unmarked core case (Nom, Abs), marked core case (Acc, Erg), and oblique case (Gen, Dat, Loc, Inst, Prep, possibly others). The case hierarchy is then simplified to *unmarked < marked < oblique*, and every syncretism must be continuous over this ordering of classes. For example, nominative-accusative syncretism is licensed because it covers two adjacent classes, unmarked and marked. Similarly, nominative-accusative-dative-instrumental syncretism involves only adjacent classes and thus is permitted. An unattested syncretism of absolutive and genitive, on the other hand, is correctly ruled out because it would involve an unmarked case and an oblique case to the exclusion of all marked cases.

---

[4] Zompí (2016) classifies some cases like ablative and allative as instances of other cases. The table faithfully reproduces his terminology to the best of my abilities. Note that Prep is short for *prepositional case*. Zompí also discusses what he calls a *prepositional locative*, and it is unclear whether this should be subsumed under Loc like a standard locative or under Prep. I decided to list it as Loc, but treating it as an instance of Prep would work equally well for the purposes of this paper.

Syncretism	Page(s)
Nom-Acc	6, 18, 35, 36
Nom-Acc + Dat-Inst	6, 28
Nom-Acc + Gen-Loc	28
Nom-Acc + Gen-Loc + Dat-Inst	18
Nom-Acc + Erg-Inst	37
Nom-Acc + Erg-Prep	39
Nom-Acc-Prep + Erg-Gen	39
Nom-Acc-Dat	19, 22
Nom-Acc-Dat-Inst	23, 24
Nom-Gen	83
Nom-Gen + Acc-Dat	87
Nom-Erg	33, 35, 36, 37
Acc-Gen	18
Acc-Gen + Loc-Dat	28
Acc-Dat	19, 21, 22
Acc-Inst	24
Acc-Prep	39
Acc-Loc-Dat	27
Acc-Loc + Gen-Dat-Inst	26
Gen-Dat	23
Gen-Loc-Dat	26
Loc-Dat	27, 28
Dat-Inst	23
Abs-Erg	32
Erg-Gen	32
Erg-Inst	33

Table 3:
Case syncretism patterns
from Zompí (2016)

While Zompí's approach represents a marked improvement, it still falls short as it is both too permissive and too restrictive. For one thing, the hierarchy allows syncretisms such as nominative-ergative-dative, which do not seem to occur. When a syncretism involves nominative and an oblique case, the marked case is always accusative and never ergative. Admittedly this might just be a statistical confound: languages with a three way contrast between nominative, ergative, and absolutive are exceedingly rare, and so are syncretisms that involve all three of Zompí's case types. Hence, one is very unlikely to

come across a language that could conceivably display syncretism of nominative, ergative, and an oblique. Still, a more principled explanation that curbs overgeneration and provides a tighter fit for this typological gap would be welcome.

The more severe problem, as Zompí (2016, p. 88) readily admits, is that his solution still undergenerates because it cannot account for the robustly attested syncretism of nominative and genitive. This is an instance of an unmarked case being syncretic with an oblique case to the exclusion of all marked cases, directly contradicting Zompí's central claim that syncretism must be contiguous across case classes. Zompí (2016, p. 87f) concludes that genitive must enjoy some special status, but does not offer a detailed account of how genitive is supposed to work in his system.

The monotonicity approach can remedy these shortcomings by building on Zompí's idea of three distinct case classes and combining them with the insight that hierarchies need not be linearly ordered. The result is a hierarchy that roughly breaks down into three "case layers", but also grants special status to genitive and locative. (To minimize clutter, the hierarchy below omits Zompí's case Prep; like other oblique cases such as allative, ablative, and so on, it would reside in the third layer.)

(13) **Case layer hierarchy**
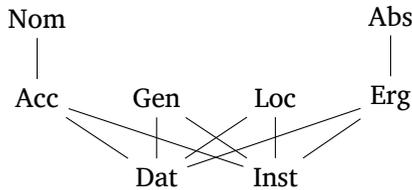
If one were to consolidate the individual cases into classes, this hierarchy would directly follow Zompí in positing *unmarked < marked < oblique*, except that one also has *special < oblique* for genitive and locative.

Let us look at several syncretism patterns from Table 3 and why they can be regarded as feasibly monotonic maps over this hierarchy. We start with nominative-accusative-dative as the only syncretism in the paradigm. Suppose that all three cases are mapped to some exponent C. Then genitive, locative, and ergative must be mapped to some

$B_g, B_l, B_e < C$, respectively; absolutive is mapped to some $A \leq B_e$; and instrumental has some $D > C$ as its exponent. Any set that furnishes a sufficient number of case exponents can be ordered in this way, so the mapping is feasibly monotonic.

Next, consider a system with three distinct syncretisms: nominative-accusative, genitive-locative, and dative-instrumental. Suppose nominative and accusative are mapped to some A. Then genitive and locative must be realized as some B, but it does not matter whether $A < B$ or $B < A$ since nominative and accusative are both unordered with respect to genitive and locative. Dative and instrumental must have an exponent C with both $A < C$ and $B < C$, and the remaining cases can be handled as before. Again, it is possible to produce such an ordering of exponents, and consequently we are dealing with yet another feasibly monotonic mapping over the case hierarchy.

Three more examples will prove instructive. First, note that syncretism of nominative and genitive to the exclusion of accusative is readily available in this system because genitive is unordered with respect to nominative and accusative. Hence, the explanation for nominative-genitive syncretism is exactly parallel to our previous account for genitive-dative syncretism in the relaxed version of Blake's hierarchy.

Second, the hierarchy captures the fact that nominative-accusative-dative syncretism is attested, but not nominative-ergative-dative. Since accusative occurs between nominative and all oblique cases (except genitive and locative), any syncretism involving nominative and one of these cases must also include accusative due to monotonicity. Yet it is also possible for nominative and ergative to be syncretic to the exclusion of accusative, which is also an attested pattern.

Third, locative must not be in the same case layer as other oblique cases because of paradigms where accusative and locative are syncretic while genitive, dative, and instrumental display a different syncretism. If we had *genitive < locative*, with A and B as the respective exponents, then monotonicity would require $A \leq B$. Since locative and genitive are not syncretic in this specific case, $A < B$ must hold. As accusative and dative are not syncretic, either, and we have *accusative < dative*, we also have $C < D$. But accusative and locative are syncretic, so $B = C$. Similarly, syncretism of genitive and dative

implies $A = D$. Put together, these equations entail both $A < B$ and $B < A$. It is clearly impossible for a linear order to obey both $A < B$ and $B < A$. Consequently, the described syncretism pattern cannot be feasibly monotonic unless locative and genitive are unordered, which requires locative to assume a special place in the hierarchy alongside genitive.

The last point highlights an important generalization of the monotonicity account.

(14) **Ban against multiple cross-level case syncretisms**

No case paradigm may display two distinct syncretism patterns A-X and B-Y such that A and B belong to the second case layer and X and Y to the third.

Hence the hierarchy in (13), albeit permissive, still puts a fair number of principled restrictions on case syncretism.[5]

Overall, then, the monotonicity approach does a decent job at accommodating the wide range of attested case syncretisms while still enforcing some testable restrictions on the typology. Its main advantage over competing approaches such as the nano-syntax analysis of

---

[5] The empirical impact of the generalization as presented here hinges on how one interprets language-specific data. No language instantiates all the cases listed in the hierarchy, which makes the status of unrealized cases an important issue. One option is to treat unrealized cases as if they were not part of the hierarchy at all. This is equivalent to positing a language-specific hiearchy that omits all unrealized cases. But instead one may treat case absence as case syncretism. For example, if a language lacks a distinct genitive but can use a dative for this purpose, one might analyze this as dative and genitive being syncretic across all paradigms. These two approaches are not empirically equivalent.

Consider a language that has both genitive and dative, but the two are sometimes syncretic. Suppose furthermore that the language also has an instrumental, which can also serve the role of a locative. However, instrumental and locative never have distinct forms. If unrealized cases are ignored, such a language is expected to exist since a monotonic function can map both genitive and dative to some exponent A while mapping instrumental to some other exponent B. If, on the other hand, missing cases are analyzed as an instance of complete syncretism, then both locative and instrumental have to be mapped to B. But then the ban against multiple cross-level syncretisms makes it impossible for both genitive and dative to be mapped to A.

I conjecture that the second approach is not empirically feasible, wherefore unrealized cases must be excised from the hierarchy.

Caha (2009) or the case type hierarchy of Zompí (2016) is the flexibility that comes with partially ordered hierarchies. This is difficult to achieve in a syntactic approach, where case hierarchies are replicated in terms of constituency containment (similar to how Bobaljik 2012 analyzes the superlative as containing the comparative, which in turn contains the base form). At the same time, though, the flexibility of the monotonicity approach also risks depriving it of any explanatory power – if just about any hierarchy will do, one can always fit the data as needed.

This is indeed a problem, but I maintain that all the hierarchies so far reflect common linguistic intuitions. As already explained in Section 3.2, the hierarchy for adjectival gradation is directly grounded in semantics, and the person hierarchy is both compatible with contemporary assumptions about person features and can be traced all the way back to Zwicky (1977). The case hierarchy looks more bewildering, but it, too, is built on linguistic ideas. The core of the hierarchy is the stratification into three types of cases, as argued for in Zompí (2016). The distinction between core cases and oblique cases is well-established, and within the core cases it is also standard to single out accusative and ergative as dependent cases that have, in a certain sense, lesser status than nominative and absolutive. The split between nominative and accusative on the one hand, and ergative and absolutive, is typologically well-supported, with most languages adopting one of the two but not both. This only leaves the special position of genitive and locative in need of an explanation.

However, our approach is not an exception in granting these cases privileged status. Caha (2009, p. 130) posits multiple distinct locatives, some of which occur very high in his linear hierarchy:

(15) **Refined case hierarchy of Caha (2009)**

Nom $>$ Acc $>$ Loc_1 $>$ Gen $>$ Loc_2 $>$ Dat $>$ Loc_3 $>$ Inst

Zompí (2016, p. 87f), on the other hand, argues that genitive exhibits special behavior because it can be an unmarked ($=$ default) case or an inherent case in syntax. Depending on its syntactic status, it may occur higher or lower in the case hierarchy. Instead of distinguishing multiple types of genitive and locative, the hierarchy in (13) instead assigns them a position that makes them more prominent than obliques but not directly comparable to the core cases.

At this point, then, we have a unified explanation of syncretism across three vastly different domains: adjectival gradation, pronouns, and case. In all three areas, typological gaps are accounted for by limiting the range of possible systems to those that can be produced by feasibly monotonic maps from some base hierarchy. Each base hierarchy is motivated by linguistic considerations. In the case of pronouns, the account provides a perfect fit for the typology, whereas it only carves out a superset of the attested patterns for adjectival gradation and case syncretism. For adjectival gradation, only the absence of the pattern AAB remains unexplained and requires a different account, e.g. in terms of syntactic containment. For case syncretisms, the amount of overgeneration is hard to assess because the number of possible combinations is so large that any gaps may just be due to insufficient data rather than principled exceptions. I put off discussion of this methodological issue until Section 5 – for now, I take the hierarchy to present a reasonable approximation of the typology of case syncretism.

I conclude this section on morphological syncretism with a short observation on noun stem allomorphy before moving on to a completely different phenomeon, the Person Case Constraint in morphosyntax.

3.4 *Case in noun stem allomorphy*

So far, I have only considered strict case syncretism, i.e. whether two forms that differ in case may have exactly the same surface realization. Hence the focus is on total identity for case. Instead, one can also look at partial identity, in particular whether two distinct cases attach to identical noun stems. This does not always occur. In Latin, for example, the nominative of 'man' is *hom-o*, whereas the accusative is *homin-em*. Nominative and accusative thus are formed with different stems of the same noun. This is known as *noun stem allomorphy*. As it turns out, the kinds of allomorphy observed with singular stems can also be captured in terms of monotonicity (I ignore plural stem allomorphy here because I am unaware of any detailed studies in this area).

McFadden (2018) proposes that all languages obey a strict condition on stem allomorphy.

(16) **Nominative stem-allomorphy generalization**

If noun stem allomorphy is conditioned by case, it distinguishes the nominative from all other cases.

In other words, noun stem allomorphy always displays an AB^n pattern, where A and B may be identical. Hence only the nominative may pick out a different stem. For a language with three cases, McFadden's generalization permits only AAA and ABB while excluding AAB, ABA, and ABC.

We already know from our discussion in Section 3.1 that AAA and ABB can be produced from a linear hierarchy by feasibly monotonic maps. Generalizing from this, it does not take much effort to verify that both A^{n+1} and AB^n are licit patterns given the case layer hierarchy in (13). The puzzle of noun stem allomorphy, hence, is not why the attested patterns are possible. Once the case layer hierarchy is in place for case syncretism, the attested noun stem allomorphy patterns are also readily available. Instead, the question is why the majority of conceivable allomorphy patterns are unattested.

Although I cannot envision a convincing reason as to why noun stem allomorphy is much more restricted than case syncretism, it is worth noting that the observed restrictions can be given a natural account in terms of the case layer hierarchy in (13). In this hierarchy, all cases in the second layer are more prominent than the cases in the third layer. But suppose that the hierarchy contains a cycle such that all third-layer cases are also more prominent than the second-layer cases. This is illustrated below, with arrows reflecting prominence.

(17) **Conflated case layer hierarchy**
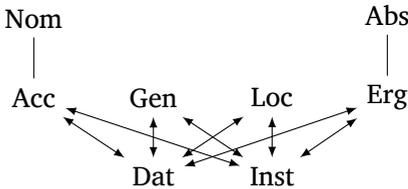
Such cycles naturally enforce identity of exponents. Let f be a feasibly monotonic map and suppose that $x \le y$ and $y \le x$. By monotonicity, this implies $f(x) \le f(y)$ and $f(y) \le f(x)$. But since f is feasibly monotonic, the exponents are linearly ordered. Therefore $f(x) \le f(y)$ and $f(y) \le f(x)$ both hold if and only if $f(x) = f(y)$. So if all second-layer and third-layer cases are part of one large cycle, any two cases x and y that are part of this cycle stand in the relation $x \le y$ and $y \le x$, and consequently they must be mapped to the same exponent.

From this it also follows that if nominative is syncretic with any case, it is syncretic with all of them. The one exception to this is absolutive, but since the data in McFadden (2018) does not include any languages with both a nominative and an absolutive, it remains to be seen whether this exception is undesirable.[6] For McFadden's sample, the solution in (17) works as desired. It allows for A^{n+1} and AB^n, and nothing else.

As I admitted earlier on, it is unclear at this point why noun stem allomorphy should use the conflated hierarchy in (17) instead of the more stratified one in (13). Like the missing AAB pattern in adjectival gradation, this issue is beyond the scope of this research project. It is one of the cases where abstracting away from the concrete mechanisms of morphology and syntax comes at the cost of leaving some aspects entirely unexplained. But as we will see next, this is exactly what makes it possible to relate morphological paradigms for adjectival gradation, person, and case to morphosyntactic well-formedness constraints.

4 PERSON CASE CONSTRAINT

Our exploration of monotonicity in language now transitions from morphology proper to a widely studied phenomenon of morphosyntax, the Person Case Constraint (PCC). The PCC is a restriction on clitic clusters. A direct object (DO) and indirect object (IO) clitic may co-occur in a cluster only if their person specifications are compatible. The "case" in PCCs thus refers to the DO-IO distinction rather than morphological case.

(18) **PCC (Spanish)** (Ormazabal and Romero 2007, p. 316f)

    a.   Pedro {me,   te}   lo   envía.
        Pedro {1SG.DAT, 2SG.DAT} 3SG.M.ACC send.PRES.3SG
        'Pedro sends it to {me, you}.'

    b.   * Pedro le   {me,   te}   envía.
        Pedro 3SG.M.DAT {1SG.ACC, 2SG.ACC} send.PRES.3SG
        'Pedro sends {me, you} to him.'

---

[6] An anonymous reviewer, citing observations in Bobaljik (2008), points out that every known split-ergative language with nominative and absolutive case has identical stem forms for the two. If this generalization is indeed exceptionless, then (17) must also contain a loop between those two cases.

The PCC is very different from the phenomena we have considered so far. It is morphosyntactic in nature, not morphological. It does not regulate the range of available exponents, but the well-formedness of clitic combinations. But just like the phenomena from Section 3 surrounding the *ABA generalization, the PCC presents an interesting puzzle where only a fraction of all conceivable options are typologically attested. And just like for the previous phenomena, the typological gaps can be readily explained in terms of monotonic mappings from a base hierarchy to some domain of values.

In the following, I first describe the PCC typology in detail (Section 4.1) and explain how PCCs can be conceptualized as monotonic maps (Section 4.2). Section 4.3 then uses this perspective to explain why only a few PCC variants seem to occur in natural languages. In contrast to previous approaches (Anagnostopoulou 2005; Adger and Harbour 2007; Nevins 2007, a.o.), the monotonicity approach correctly predicts the existence of a recently described PCC variant in Choctaw. As shown in Section 4.4, it also generalizes straightforwardly to the Gender Case Constraint (Foley *et al.* 2017). Monotonicity thus manages to provide a unified perspective on a wide range of seemingly unrelated phenomena.

4.1   *The PCC typology*

The PCC was first observed in Perlmutter (1971), but it is only in recent years that it has attracted significant attention (see e.g. Bonet 1994; Anagnostopoulou 2005; Adger and Harbour 2007; Nevins 2007, 2011; Rezac 2007; Walkow 2012; Graf 2014, 2017). The current literature distinguishes four different types of PCC.

(19) a. **S(trong)-PCC**
          DO must be 3. (Bonet 1994)

      b. **U(ltrastrong)-PCC**
          DO is less prominent than IO, where 3 is less prominent than 2, and 2 is less prominent than 1. (Nevins 2007)

      c. **W(eak)-PCC**
          3IO combines only with 3DO. (Bonet 1994)

      d. **M(e first)-PCC**
          If IO is 2 or 3, then DO is not 1. (Nevins 2007)

Very recent work argues for the existence of additional PCCs (Stegovec 2016; Tyler 2017), but for the sake of exposition I will limit the discussion to these four PCCs for now and return to the others later on.

Even with just four PCCs under discussion, it is hard to deny that the class of PCCs as defined above seems rather bewildering. However, a clearer picture emerges if one simply represents the licit and illicit combinations in tabular form as in Table 4. The tables make it readily apparent that each PCC represents a specific way for grammaticality to spread from the top right corner – the 1-3 combination – towards the bottom left, where 3-1 resides.

Table 4: Attested variants of the PCC

IO↓/DO→	1	2	3
1	NA	*	✓
2	*	NA	✓
3	*	*	NA

(a) S-PCC

IO↓/DO→	1	2	3
1	NA	✓	✓
2	*	NA	✓
3	*	*	NA

(b) U-PCC

IO↓/DO→	1	2	3
1	NA	✓	✓
2	✓	NA	✓
3	*	*	NA

(c) W-PCC

IO↓/DO→	1	2	3
1	NA	✓	✓
2	*	NA	✓
3	*	✓	NA

(d) M-PCC

Before we proceed, an important disclaimer is in order regarding the diagonal of each table. The cells where IO and DO have the same person specification are marked as NA, which is short for "not applicable". The reason for this value is not a lack of available data, but rather how this data should be analyzed. These cases are known to exhibit special behavior, such as *3-3 effects* (Perlmutter 1971, p. 132).

(20)　 * Le　　　 lo　　　　 recommendé.
　　　 3SG.DAT 3SG.M.ACC recommend.PAST.1SG
　　　 'I recommended it to him.'

Although (20) looks like a PCC-effect, it is generally assumed that the source of 3-3 effects is morphological in nature (see Foley *et al.* 2017 for a full discussion). Since PCCs are taken to be syntactic in nature, morphologically conditioned phenomena do not fall under their purview and should be treated differently. I follow this common line of reasoning and exclude all cells along the diagonal from the PCCs.

Even with the diagonal excluded, there are $2^6 = 64$ conceivable PCC variants. Only a small fraction of those 64 are actually attested. Every account of the PCC thus has to explain why the range of variation is severely limited. As we will see soon, the monotonicity approach does so with little effort.

4.2   *PCCs as monotonic maps*

We now turn our attention to the four patterns in Table 4 and why each one of them can be regarded as the result of a monotonic map. First, we have to define an appropriate structure to represent IO/DO-combinations. We do this by constructing a specific crossproduct based on the person hierarchy $P := 1 > 2 > 3$ of Zwicky (1977), which we already encountered during the discussion of person syncretisms in Section 3.1. Given two hierarchies A and X, their crossproduct is the structure $\langle A \times X, \leq \rangle$ such that $\langle a, x \rangle \leq \langle b, y \rangle$ iff $a \leq_A b$ and $x \leq_X y$. An example is shown below.

(21)

$$A \qquad\qquad X \qquad\qquad A \times X$$
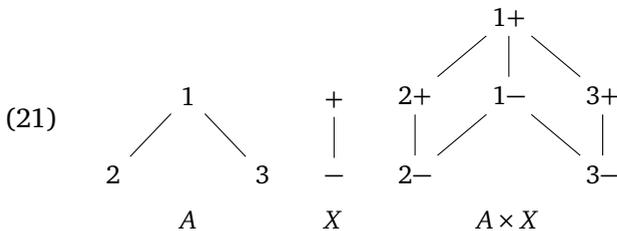
With the right crossproduct, all the PCCs in Table 4 will turn out to be captured by monotonic maps from this crossproduct to the algebra **2** of truth values.

It is very tempting to go with the intuitively pleasing option to construct the crossproduct $P \times P$. This hierarchy, which is shown below, combines the person hierarchy with itself so that every node of the hierarchy represents a specific IO-DO combination.

(22) **Full person-person hierarchy** $P \times P$
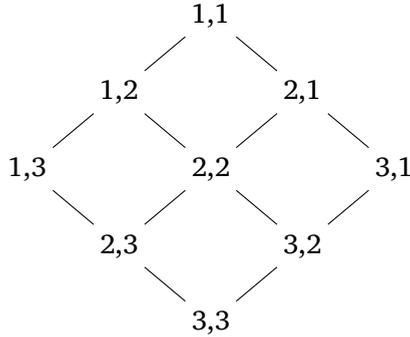
Since we ignore the diagonal, we remove all nodes of the form x, x, for $x \in \{1, 2, 3\}$. This leaves us with the reduced hierarchy $\mathscr{R}$ below.

(23) **Reduced person-person hierarchy** $\mathscr{R}$
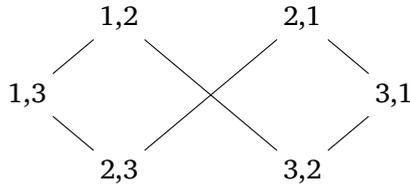
However, this is not the correct hierarchy for our purposes. The W-PCC does indeed correspond to a monotonic mapping from this hierarchy to the algebra **2**, where T indicates a well-formed combination and F an ill-formed one. But the same does not hold for the U-PCC. Let us look at this in detail, starting with the W-PCC:

(24) **W-PCC as a monomorphemic map**
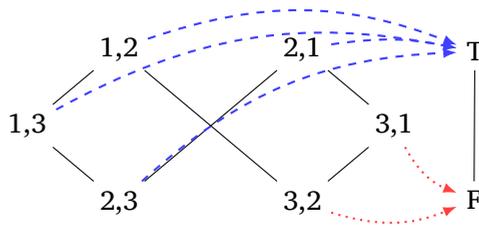
The relation between a map like in (24) and the PCC tables is as follows: if a node of the form x, y is mapped to T, then the cell in row x and column y has a checkmark. In other words, a combination of an IO with person x and a DO with person y is well-formed. The reader is

invited to verify for themselves that the mapping in (24) does indeed define the pattern of the W-PCC.

Note that (24) contains no x and y such that $x \leq_{\mathscr{R}} y$ yet $f(y) \leq_2$ $f(x)$, which establishes that the W-PCC mapping is monotonic. But the mapping for the U-PCC is not monotonic.

(25) **U-PCC mapping**
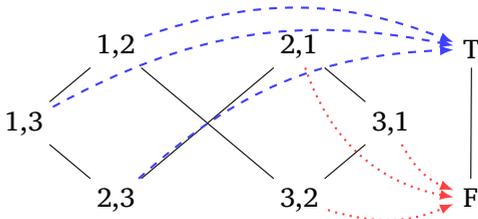
Here, we have $2,3 \leq_{\mathscr{R}} 2,1$ yet $f(2,1) = F \leq_2 T = f(2,3)$. Switching the order of T and F does not help, for then the problematic pair would be $1,2$ and $3,2$. So the U-PCC mapping isn't even feasibly monotonic for $\mathscr{R}$, which entails that it cannot be monotonic. If the typology of PCCs is to be analyzed as yet another instance of monotonicity, a different hierarchy is needed.

Recall from the initial discussion of the patterns in Table 4 that well-formedness seems to be growing out from the top-right corner, which corresponds to the combination $1,3$. This suggests that we should construct a hierarchy where the top element is $1,3$ rather than $1,1$. One candidate is the crossproduct of the person hierarchy $P := 1 > 2 > 3$ with its dual $P^{-1} := 3 > 2 > 1$. The result is shown below.

(26) **Dual Person Hierarchy** $P \times P^{-1}$
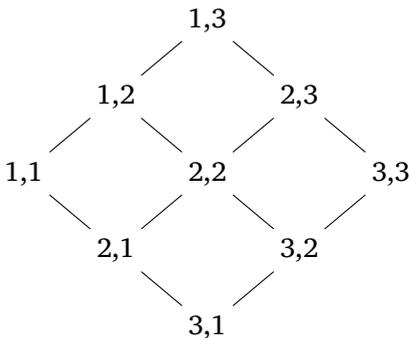
Once again all elements of the form x, x are removed, which results in a new person hierarchy $\mathscr{P}$.

(27) **Reduced dual person hierarchy $\mathscr{P}$ (final)**
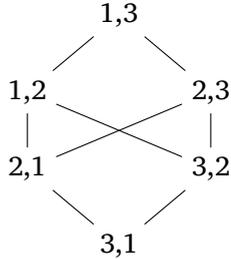
Over $\mathscr{P}$, each one of the four attested PCCs corresponds to a monotonic map to **2**. These are depicted in Table 5. For the sake of simplicity, I omit **2** and instead put a box around a node iff it is mapped to T.

In conclusion, the attested PCC variants can all be regarded as monotonic functions from the hierarchy $\mathscr{P}$ to the algebra **2** of truth values.

Table 5:
The four PCC variants as monotonic maps from the person hierarchy to the algebra of truth values; boxed nodes are mapped to true, all others to false

4.3   *Explaining the typology*
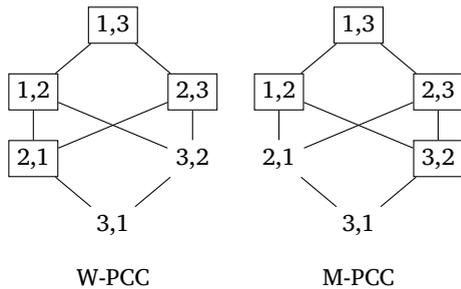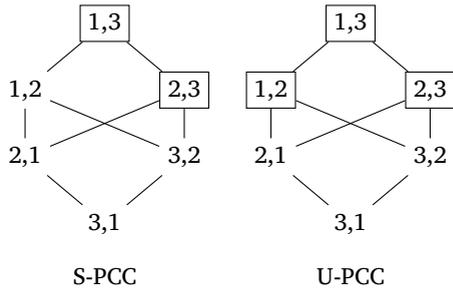
The four mappings depicted in Table 5 do not exhaust the range of available monotonic functions. So we have succeeded in correlating each attested PCC with monotonicity, but we have not given a monotonicity-based characterization of the PCC typology. But as I argue next, this too is fairly simple.

Only five other maps from $\mathscr{P}$ to **2** are monotonic. The first two map everything to T or everything to F.

(28)
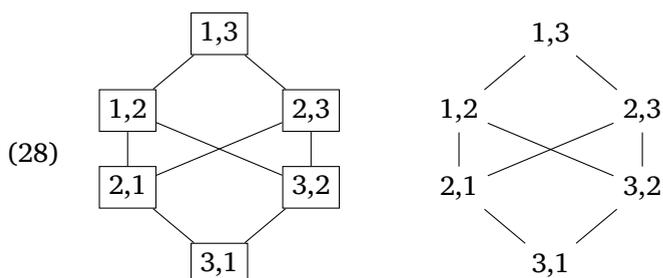
This corresponds to PCCs where either all combinations are allowed or all options are forbidden. Even though such patterns are usually not considered PCCs in the literature, they do exist. Many languages allow clitics to be freely combined, e.g. German. And at least in Cairene Arabic, clitics may never be combined, irrespective of their person specification (Shlonsky 1997; Walkow p.c.). We may consider these to be instances of a F[ree]-PCC and an I[ndiscriminate]-PCC, respectively. So these two monotonic maps do have attested counterparts in the typology.

The next mapping is a mirror image of the S-PCC, where the only licit combinations are $1, 3$ and $1, 2$ instead of $1, 3$ and $2, 3$.

(29)
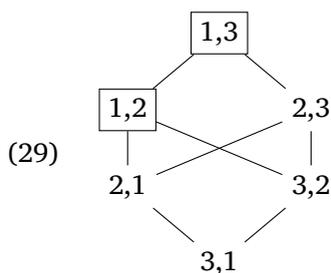
For the longest time, this pattern has been believed not to exist. But Tyler (2017, p. 10) reports that Choctaw has a rather unusual PCC of the following form:

(30) **PCC in Choctaw** (as reported)

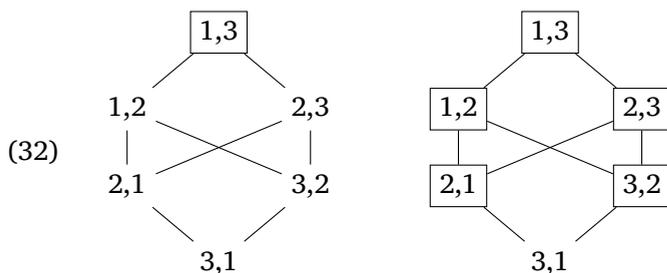| IO↓/DO→	1	2	3
1	NA	✓	✓
2	*	NA	*
3	✓	✓	NA

This pattern is clearly not monotonic over any of the hierarchies entertained so far: (22), (23), (26), and (27). However, Tyler (p.c.) states that the combinations 3, 1 and 3, 2 never surface in the data. He argues that they are blocked for reasons that are unrelated to the PCC – similar to how the diagonal often exhibits special behavior. The value ✓ in the corresponding cells thus does not indicate an empirically attested combination, but rather the theoretical claim that these patterns would be well-formed if it were not for these independent factors. We might also entertain the scenario, then, that these combinations are also illicit with respect to the PCC, which yields a very different table.

(31) **PCC in Choctaw** (reanalyzed)

IO↓/DO→	1	2	3
1	NA	✓	✓
2	*	NA	*
3	*	*	NA

Let us call this pattern the C[hoctaw]-PCC. The C-PCC corresponds exactly to the monotonic mapping in (29) where only 1, 3 and 1, 2 are mapped to T.

So far then, three monotonic maps beyond the initial four have been successfully related to some attested data pattern. This leaves only two more monotonic maps to consider. In one 1, 3 is the only licit combination, in the other one 3, 1 is the only illicit combination.

(32)

$$
\begin{array}{c}
\boxed{1,3} \\
1,2 \qquad 2,3 \\
2,1 \qquad 3,2 \\
3,1
\end{array}
\qquad
\begin{array}{c}
\boxed{1,3} \\
\boxed{1,2} \qquad \boxed{2,3} \\
\boxed{2,1} \qquad \boxed{3,2} \\
3,1
\end{array}
$$

To the best of my knowledge, no language exhibits a pattern of this kind. So just as in the case of adjectival gradation, monotonicity is a bit too loose as a characterization of the typology.

That said, it is easy to distinguish these two unattested patterns from the rest. They are the only ones that define a class with only a single member. Either the set of well-formed combinations is a singleton, or the class of ill-formed combinations is. There may be independent reasons that drive languages towards patterns that do not put a single combination in opposition to the rest of the paradigm. Or perhaps this is yet another case where the answer can only be found at a less abstracted level of description.

Be that as it may, monotonicity in combination with the hierarchy $\mathscr{P}$ from (27) provides a very tight fit for the attested PCC typology, with only minimal overgeneration. This establishes a direct connection to the morphological syncretism phenomena discussed in Section 3. But as we will see next, it can also be extended to other aspects of morphosyntax, such as the recently reported Gender Case Constraint.

4.4   *Hierarchical reversal in the Gender Case Constraint*

Besides furnishing two unattested patterns, the monotonicity account also seems stipulative in that it fails for the intuitively most pleasing hierarchy $P \times P$ and instead has to use (a reduced version of) $P \times P^{-1}$, where the order in the second component is the reverse of that in the first component. Why should natural language operate with such a peculiar hierarchy? I have no insightful answer to this puzzle, but I would like to point out that the puzzle is not limited to the PCC.

Foley *et al.* (2017) report that Zapotec displays a restriction on subject-object clitic clusters that is driven by gender rather than person, a *Gender Case Constraint* (GCC). This constraint only acts on third

person clitics. Zapotec distinguishes four genders: elder human, non-elder human, animal, and inanimate. For the sake of simplicity, I refer to these as 1, 2, 3, and 4. The Zapotec GCC then produces the following pattern:

(33) **Zapotec GCC** (Foley *et al.* 2017, p. 6)

S↓/O→	1	2	3	4
1	NA	✓	✓	✓
2	*	NA	✓	✓
3	*	*	NA	✓
4	*	*	*	NA

Mirroring standard practice for the PCC, Foley *et al.* (2017) argue that the diagonal can be subject to a separate constraint against identical combinations. For this reason, I have given those cells the value NA here, but nothing in the subsequent discussion hinges on that.

After our extensive discussion of monotonicity in the PCC, the reader should be able to see immediately that the pattern in (33) is monotonic if one starts with a hierarchy $1 > 2 > 3 > 4$ and combines it with its inverse $4 > 3 > 2 > 1$.

(34) **Gender hierarchy with the second component reversed**
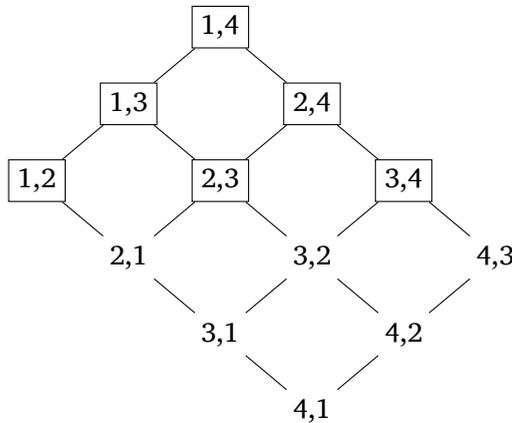
If, on the other hand, the hierarchy were simply built by combining $1 > 2 > 3 > 4$ with itself, the Zapotec GCC would not be monotonic by virtue of not being feasibly monotonic.

(35) **Gender hierarchy with identical components**
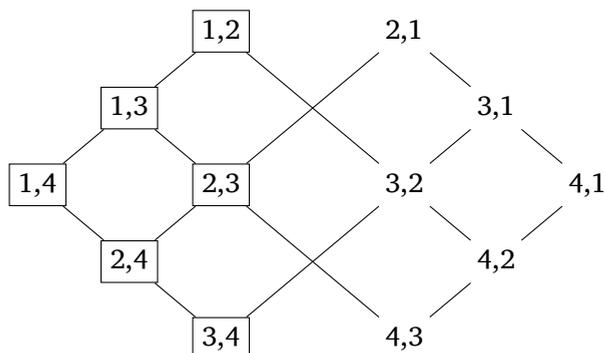
Here we have $3, 4 \leq 3, 2 \leq 1, 2$ and thus $f(3,4) = T \leq f(3,2) = F \leq f(1,2) = T$ – but $T \leq F \leq T$ cannot be satisfied irrespective of how one orders T and F. It seems, then, that it is a general fact of language that clitic combination patterns "grow" from the top-right corner of the paradigm, which is formally captured by reversing the second hierarchy. I doubt that the reason for this can be discerned at the level of abstraction at which the monotonicity approach operates. This question requires the more fine-grained and detailed assumptions of syntactic and/or morphological formalisms.

4.5   *Other PCCs*

There is one more point regarding the typology of PCCs that merits discussion. Stegovec (2016) argues based on data from Slovenian that the established PCCs also have counterparts where the relevant contrast is not that between IO and DO, but rather which clitic occurs linearly first. For example, Slovenian has a counterpart of the S-PCC that looks exactly the same as the one in Table 4 except that the x and y-axis do not correspond to DO and IO, but rather to the linearly second and the linearly first clitic.

Stegovec's findings present a major challenge for syntactic approaches such as Anagnostopoulou (2005), which derive the limited PCC typology from structural asymmetries between IO and DO. Without this structural asymmetry, syntactic accounts lose all their force. In particular, it becomes mysterious why the range of possible PCC patterns remains the same when the conditioning factor is linear order instead of the IO-DO asymmetry.

The monotonicity approach, on the other hand, can easily accommodate these findings. All the work is done by the hierarchy in (27), which is agnostic about what each component of a node encodes. The standard interpretation of, say, 3,1 is that IO is third person and DO is first person. But we might just as well interpret it as saying that the linearly first clitic is third person and the second one is first person. The monotonicity approach is not concerned with identifying potential triggers or the exact linking from syntactic configurations to the hierarchies it operates over. What it does is provide an abstract characterization of the range of variation once the appropriate triggers have been identified. As we have seen throughout this paper, this high-level approach has great unifying power, but it comes at the expense of leaving certain issues entirely unaddressed. Stegovec's findings show that this kind of agnosticism, albeit occasionally unsatisfying, increases the robustness of the approach.

5 METHODOLOGICAL REMARKS

Before I turn to the summary of this paper's key findings, there still are several methodological issues that deserve a careful exploration. These concern the status of monotonicity as a desirable property of mappings (Section 5.1), the cognitive status of the analysis advanced in this paper (Section 5.2), and the risks of building a formal model on a limited range of data (Section 5.3).

5.1 *Monotonicity in language*

Monotonicity plays a central role in this paper. It is the key ingredient that narrows down the range of variation once a suitable base hierarchy has been defined. One may wonder, then, why monotonicity should be a desirable property for language. This is a very deep question that cannot be answered in a few lines. That said, the role of monotonicity seems to extend beyond the phenomena discussed in the preceding sections.

In Section 2.1, I used the *No Crossing Branches* constraint from autosegmental phonology to illustrate the concept of monotonicity. But monotonicity in phonology goes beyond this constraint. To give but one example, there seems to be no phonological process that

targets high and low vowels to the exclusion of mid vowels – a kind of phonological *ABA constraint, and arguably an instance of monotonicity.

Monotonicity can also be found in syntax. The Accessibility hierarchy of Keenan and Comrie (1977) classifies different kinds of NP by their relative prominence as SU > DO > IO > OBL > GEN > OCOMP and states that if a language allows for NPs of type x to be relativized, it also allows this for every NP type $y > x$. This implicational universal amounts to the requirement that mappings from the linear hierarchy of NP-types to the algebra **2** must be monotonic (being mapped to T means that relativization is allowed).

Another example comes from the ordering of operations in syntax. A phrase can undergo three types of operations: selection, A-move, and A′-move. Once a phrase has undergone A-movement, it can no longer be selected or select arguments of its own. Furthermore, A-movement is impossible once the phrase has undergone A′-movement. This, too, can be viewed as an instance of monotonicity. For any given phrase, consider the linear sequence of operations it undergoes during the syntactic derivation. For example, an arbitrary DP's record may read *Select-A-A′-A′*. Now suppose furthermore that operation types are linearly ordered such that *Select < A-move < A′-move*. Then the inability to select after A-moving or to A-move after A'-moving follows from the requirement that the mapping from a phrase's derivational record to the hierarchy of operation types must be monotonic.

A more complex example from syntax is the analysis of adjunct island effects in Graf (2013). There, the fact that extraction from an adjunct is ungrammatical is reinterpreted as a monotonicity requirement over a specific kind of algebra. This monotonicity entailment can produce situations where a syntactic structure is illicit even though it does not violate any syntactic constraints. Hence, there is no such thing as an Adjunct Island Constraint, the observed effects are a direct consequence of monotonicity.

Monotonicity also surfaces in semantics. The denotations of determiners, for example, are always monotonic (Keenan and Westerståhl 1996; Peters and Westerståhl 2006). Even in the realm of lexical semantics, it has been argued that word meanings tend to be convex (Gärdenfors 2000; Jäger 2007, 2010), a notion that is closely related to monotonicity.

Overall, then, there is plenty of evidence for monotonicity in language, although the motivations for that are still largely unclear. This paper just adds a few more entries to a long list of phenomena that involve monotonicity.

5.2 *Cognitive commitment*

The previous section suggests that monotonicity plays a fundamental role in language. This seems to be at odds with the initially stated aim for an effective theory, i.e. an account that correctly characterizes the system under investigation but does not necessarily encompass the causal factors that give rise to this system. Upon reflection, though, these two positions are perfectly compatible.

It is true that the monotonicity account deliberately abstracts away from those factors that linguists consider the nuts and bolts of mental grammars: features bundles, feature checking, structure-building operations, and so on. As a consequence, the concepts I rely on may not have direct counterparts in the grammar. For example, the posited person hierarchy of $1 > 2 > 3$ is entirely agnostic about the long-standing issue whether third person is a feature or the absence of person features. At the same time, this does not entail that the person hierarchy is merely a descriptive device without any cognitive reality. Rather, the claim is that the mechanisms of the grammar, whatever they may be, are such that they give rise to this kind of hierarchy for the surface forms we observe in the data. In a sense, this is no different from syntacticians asserting the cognitive reality of their grammar formalism while leaving the neural substrate of the grammar unspecified. The monotonicity approach employs the same strategy, characterizing the high-level behavior of language while abstracting away from the grammar substrate.

As the reader has seen throughout the paper, this has several advantages. One can now generalize across domains that arguably do not look very similar and behave very differently at the usual level of grammatical description. By abstracting away from technical details, the account remains remarkably simple on a formal level. It also is largely framework agnostic and is directly compatible with Minimalism (Chomsky 1995), Distributed Morphology (Halle and Marantz 1993; Embick and Marantz 2008), GPSG (Gazdar *et al.* 1985), HPSG (Pollard and Sag 1994), LFG (Bresnan 1982), and TAG (Joshi 1985),

among others. By not directly linking into the tools and concepts of existing grammar formalisms, the account also enjoys a greater amount of freedom and can easily be adjusted to fit new data – the discovery of new PCC types in Stegovec (2016), for example, poses a major problem for syntactic accounts, but not for the monotonicity approach.

The obvious downside of abstraction is that some typological gaps cannot be adequately explained. Without the connection to the grammar substrate, domain-specific limitations such as the absence of AAB patterns in adjectival gradation remain mysterious. There is a risk that this abstractness, combined with the malleability of the approach, will ultimately lead to blind descriptivism, with the hierarchies constantly being tweaked and refined until they fit the data. As I argue next, though, the goals of the enterprise make this an unappealing option and hence an unlikely outcome.

5.3	*The risk of overfitting*

Typological accounts always face the danger of overfitting their theory to an unrepresentative data sample. Even large-scale studies rarely contain data from more than 150 languages. Since many phenomena such as the PCC are exceedingly rare, the sample of languages that display the phenomenon in question is even smaller. At the same time, combinatorial explosion leads to large numbers of logically possible systems in certain domains. For instance, there are 203 distinct case syncretism patterns for a system with 6 cases. In the face of such numbers, it is doubtful that our current data exhausts the full range of variation. In addition, the analysis of existing data is fraught with difficulties. Syncretism, for example, has to be distinguished from merely accidental homophony, which leaves plenty of wiggle room in how the data is interpreted (but see Sauerland and Bobaljik 2013 for a more rigorous approach to accidental homophony). Even in cases where sufficient data is available, then, it might have been misanalyzed.

While all these points are certainly correct, they are unavoidable given the realities of doing empirical work in linguistics. All competence data is heavily theory-laden, and since we do not know the full extent of universal grammar, even 6000 languages may only provide a small sample of the full space of grammars. Instead of stopping dead in their tracks, linguists accept that the empirical landscape may change

from one day to the next and try to formulate the most insightful theories given the currently available data.

Things are no different for the monotonicity approach, but it has several properties that mitigate the data problem. First and foremost, the account is about identifying principles that hold across many different domains. Hence, a data shortage in one domain is less of an issue because it can be offset by insights from another domain. Person syncretism and the PCC, for instance, mutually support each other as they both operate over person hierarchies. In addition, the monotonicity approach is robust in the sense that it does not try to perfectly fit the existing data but rather identifies upper bounds on the range of variation. For example, the ABA pattern is ruled out on systematic grounds, whereas the absence of AAB patterns in adjectival gradation has to be stipulated. By focusing on what holds across many domains, the approach avoids overfitting the data for any given domain.

How well this enterprise will work out in practice remains to be seen. But at this point, there is no reason to dismiss it on conceptual or methodological grounds.

6   CONCLUSION

The account proposed in this paper derives typological gaps from two components: a fixed underlying hierarchy shared across all languages (a person hierarchy, case hierarchy, and so on), and the requirement that the mappings from these hierarchies to output forms must be monotonic. This simple principle produces close approximations of the range of variation for each domain, with only some requiring further stipulations (e.g. the ban against AAB pattern in adjectival gradation, or the absence of PCC patterns with only one well-formed or ill-formed clitic combination). The relevant hierarchies are summarized in Table 6.

While it must remain an open question for now why monotonicity should play such an important role, it cannot be denied that it surfaces in many other areas of language, including phonology, morphology, syntax, and semantics. Hopefully future work will be able to shed light on this fundamental question.

A lot of analytic work also remains to be done. The literature on typological generalizations is enormous, and only a few could be

Phenomenon	Hierarchy
Adjectival gradation	positive | comparative | superlative
Pronoun syncretism	1 | 2 | 3
Case syncretism	Nom — Acc · Gen · Loc · Erg — Abs · Dat · Inst
Noun stem allomorphy	Nom — Acc · Gen · Loc · Erg — Abs · Dat · Inst
PCC	1,3 · 1,2 · 2,3 · 2,1 · 3,2 · 3,1
GCC	1,4 · 1,3 · 2,4 · 1,2 · 2,3 · 3,4 · 2,1 · 3,2 · 4,3 · 3,1 · 4,2 · 4,1

Table 6:
Overview of posited
hierarchies for each
phenomenon

explored here. Number was completely ignored, and preliminary work on syncretism patterns in verbal inflection suggests that its behavior is more complicated than that of person in the PCC. Other topics for future research are inverse marking and resolved agreement.

Expanding the range of domains is of vital importance as it may provide additional support for hierarchies posited here. For example, the person hierarchy $1 > 2 > 3$ should surface in every domain that involves person. In addition, hierarchies for entirely new domains will increase our understanding of what hierarchies can (and cannot) look like. This is vital for keeping the approach from devolving into pure stipulation of suitable hierarchies.

ACKNOWLEDGMENTS

REFERENCES

David ADGER and Daniel HARBOUR (2007), Syntax and Syncretisms of the Person Case Constraint, *Syntax*, 10:2–37, doi:10.1111/j.1467-9612.2007.00095.x.

Elena ANAGNOSTOPOULOU (2005), Strong and Weak Person Restrictions: A Feature Checking Analysis, in Lorie HEGGIE and Francisco ORDOÑEZ, editors, *Clitics and Affix Combinations: Theoretical Perspectives*, pp. 199–235, John Benjamins, Amsterdam.

Barry J. BLAKE (2001), *Case*, Cambridge University Press, Cambridge.

Jonathan D. BOBALJIK (2008), Where's ϕ? Agreement as a Post-Syntactic Operation, in David ADGER, Daniel HARBOUR, and Susana BEJAR, editors, *Phi Theory: Phi-Features Across Modules and Interfaces*, pp. 295–328, Oxford University Press, Oxford.

Jonathan D. BOBALJIK (2012), *Universals in Comparative Morphology: Suppletion, Superlatives, and the Structure of Words*, MIT Press, Cambridge, MA.

Jonathan D. Bobaljik and Uli Sauerland (2018), ABA and the Combinatorics of Morphological Features, *Glossa: A Journal of General Linguistics*, 3, doi:10.5334/gjgl.345, article 15.

Eulàlia Bonet (1994), The Person-Case Constraint: A Morphological Approach, in *The Morphology-Syntax Connection*, number 22 in MIT Working Papers in Linguistics, pp. 33–52.

Joan Bresnan (1982), *The Mental Representation of Grammatical Relations*, MIT Press, Cambridge, MA.

Pavel Caha (2009), *The Nanosyntax of Case*, Ph.D. thesis, University of Tromsø.

Pavel Caha (2013), Explaining the Structure of Case Paradigms by the Mechanisms of Nanosyntax: The Classical Armenian Nominal Declension, *Natural Language and Linguistic Theory*, 31:1015–1066, doi:10.1007/s11049-013-9206-8.

Noam Chomsky (1995), *The Minimalist Program*, MIT Press, Cambridge, MA, doi:10.7551/mitpress/9780262527347.003.0003.

David Embick and Alec Marantz (2008), Architecture and Blocking, *Linguistic Inquiry*, 39:1–53.

Steven Foley, Nick Kalivoda, and Maziar Toosarvandani (2017), Forbidden Clitic Clusters in Zapotec: Implications for the Person-Case Constraint, handout of a talk presented at CLS, May 25 2017, University of Chicago.

Gerald Gazdar, Ewan Klein, Geoffrey K. Pullum, and Ivan A. Sag (1985), *Generalized Phrase Structure Grammar*, Blackwell, Oxford.

John Goldsmith (1976), *Autosegmental Phonology*, Ph.D. thesis, MIT.

Thomas Graf (2013), The Syntactic Algebra of Adjuncts, in *Proceedings of CLS 49*, `http://thomasgraf.net/doc/papers/Graf13CLS.pdf`, to appear.

Thomas Graf (2014), Feature Geometry and the Person Case Constraint: An Algebraic Link, in *Proceedings of CLS 50*, `http://thomasgraf.net/doc/papers/Graf14CLS.pdf`, to appear.

Thomas Graf (2017), Graph Transductions and Typological Gaps in Morphological Paradigms, in *Proceedings of the 15th Meeting on the Mathematics of Language (MOL 2017)*, pp. 114–126, `http://www.aclweb.org/anthology/W17-3411`.

Peter Gärdenfors (2000), *Conceptual Spaces: The Geometry of Thought*, MIT Press, Cambridge, MA.

Morris Halle and Alec Marantz (1993), Distributed Morphology and the Pieces of Inflection, in Ken Hale and Samuel J. Keyser, editors, *The view from building 20*, pp. 111–176, MIT Press, Cambridge, MA.

Daniel HARBOUR (2015), Poor Pronoun Systems and What They Teach us, *Nordlyd*, 41(1), doi:10.7557/12.3314.

Heidi HARLEY and Elizabeth RITTER (2002), Person and Number in Pronouns: A Feature-Geometric Analysis, *Language*, 78:482–526, doi:10.1353/lan.2002.0158.

Gísli Rúnar HARÐARSON (2016), A Case for a Weak Case Contiguity Hypothesis — A Reply to Caha, *Natural Language and Linguistic Theory*, 34:1329–1343, doi:10.1007/s11049-016-9328-x.

Aravind JOSHI (1985), Tree-Adjoining Grammars: How Much Context Sensitivity is Required to Provide Reasonable Structural Descriptions?, in David DOWTY, Lauri KARTTUNEN, and Arnold ZWICKY, editors, *Natural Language Parsing*, pp. 206–250, Cambridge University Press, Cambridge.

Gerhard JÄGER (2007), The Evolution of Convex Categories, *Linguistics and Philosophy*, 30:551–564.

Gerhard JÄGER (2010), Natural Color Categories Are Convex Sets, in Maria ALONI, Harald BASTIAANSE, Tikitu DE JAGER, and Katrin SCHULZ, editors, *Logic, Language and Meaning*, pp. 11–20, Springer, Berlin, Heidelberg, doi:10.1007/978-3-642-14287-1_2.

Edward KEENAN and Dag WESTERSTÅHL (1996), Generalized Quantifiers in Linguistics and Logic, in Johan VAN BENTHEM and Alice TER MEULEN, editors, *Handbook of Logic and Language*, pp. 837–894, North-Holland, Amsterdam.

Edward L. KEENAN and Bernard COMRIE (1977), Noun Phrase Accessiblity and Universal Grammar, *Linguistic Inquiry*, 8:63–99.

Thomas MCFADDEN (2018), *ABA in Stem-Allomorphy and the Emptiness of the Nominative, *Glossa: A Journal of General Linguistics*, 3, article 8.

Andrew NEVINS (2007), The Representation of Third Person and Its Consequences for Person-Case Effects, *Natural Language and Linguistic Theory*, 25:273–313, doi:10.1007/s11049-006-9017-2.

Andrew NEVINS (2011), Multiple Agree with Clitics: Person Complementarity vs. Omnivorous Number, *Natural Language and Linguistic Theory*, 28:939–971, doi:10.1007/s11049-006-9017-2.

Javier ORMAZABAL and Juan ROMERO (2007), The Object Agreement Constraint, *Natural Language and Linguistic Theory*, 25:315–347.

David M. PERLMUTTER (1971), *Deep and Surface Structure Constraints in Syntax*, Holt, Rinehart and Winston, New York.

Stanley PETERS and Dag WESTERSTÅHL (2006), *Quantifiers in Language and Logic*, Oxford University Press.

Carl POLLARD and Ivan SAG (1994), *Head-Driven Phrase Structure Grammar*, CSLI and The University of Chicago Press, Stanford and Chicago.

Milan Rezac (2007), Escaping the Person Case Constraint: Reference-Set Computation in the ϕ-system, *Linguistic Variation Yearbook*, 6:97–138.

Uli Sauerland and Jonathan D. Bobaljik (2013), Syncretism Distribution Modeling: Accidental Homophony as a Random Event, in Nobo Goto, Koichi Otaki, Atsusi Sato, and Kensuke Takita, editors, *Proceedings of GLOW in Asia 9*, pp. 31–53.

Ur Shlonsky (1997), *Clause Structure and Word Order in Hebrew and Arabic*, Oxford University Press, Oxford.

Peter W. Smith, Beata Moskal, Ting Xu, Jungmin Kang, and Jonathan D. Bobaljik (2018), Case and Number Suppletion in Pronouns, *Natural Language & Linguistic Theory*, doi:10.1007/s11049-018-9425-0.

Adrian Stegovec (2016), Personality Disorders and Missing Persons: Deriving the Person-Case Constraint without Case, `http://ling.auf.net/lingbuzz/002632`, ms., University of Connecticut.

Matthew Tyler (2017), Choctaw PCC Repair: Basque-Style PCC Repair in a Language with no Dative, handout of a talk presented at CLS, May 25 2017, University of Chicago.

Martin Walkow (2012), *Goals, Big and Small*, Ph.D. thesis, University of Massachusetts Amherst.

Stanislao Zompí (2016), *Case Decomposition Meets Dependent-Case Theories. A Study of the Syntax-Morphology Interface*, Master's thesis, University of Pisa.

Arnold Zwicky (1977), Hierarchies of Person, in *Chicago Linguistic Society*, volume 13, pp. 714–733.