Morphosyntactic Patterns Follow Monotonic Mappings ⋆

Sedigheh Moradi

Stony Brook University, Stony Brook NY, USA
sedigheh.moradi@stonybrook.edu

Abstract. Apart from being a system of structures, language is a system of relations. Understanding the particular regularities underlying these relations help us predict both possibilities and gaps in linguistic organization. This paper follows Graf's work [13] in positing monotonicity as a substantial underlying restriction on possible patterns in morphosyntactic paradigms. This approach not only extends the notion of monotonicity outside semantics, but also combines this formal explanation with extralinguistic motivations. The tense hierarchy I propose for syncretism in verbal paradigms is independently motivated by Reichenbach's tense system [22]. The gender hierarchy used for gender resolution rules is directly extracted from the organization of the linguistic data. The restriction on both types of paradigms is readily explained by the fact that they only allow monotonic mappings from a base hierarchy to output forms.

Keywords: Monotonicity · Morphosyntax · Tense Syncretism · Gender Resolution Rules.

1 Introduction

It is generally accepted that language variability is not limitless and there are common restrictions on the attestibility of patterns. Out of this view grew the notion of universals in pursuit of explanations in linguistics [10]. Chomsky [8] classifies *linguistic universals* as formal and substantive. *Substantive universals* are the building blocks of grammar. These are particular regularities that the formal rules express. A *formal universal* is the property of having a grammar meeting a certain abstract condition.

The majority of linguistic work is concerned with formal universals, and this holds in particular for work grounded in mathematics or computation. For example, recent work on subregular complexity ([1, 14–16] and references therein) shows that many aspects of language — from phonology to morphology, syntax, and even semantics — are very limited in terms of their computational complexity. These limits can be used to explain why certain intuitively plausible patterns

⋆ The work reported in this article is supported by the National Science Foundation under Grant No. BCS-1845344.

do not seem to occur across languages. However, this perspective cannot explain why, say, there is a process of intervocalic voicing, but not one of intervocalic devoicing, as both processes would have exactly the same complexity. Here it is the substance of the involved elements that matters, rather than the complexity of this process. The central claim of this paper is that monotonicity can close this gap as it provides a fruitful, formally rigorous perspective on linguistic substance.

Strictly speaking, monotonicity is a semantic notion. However, it has been linked to many fundamental aspects of linguistic processing, reasoning, and grammar [17]. Monotonicity, as explained in this article, is used to provide a formal basis for certain morphosyntactic patterns. I present typological data mapping the attested variation in two morphosyntactic domains: tense syncretism and resolved gender agreement. I then show that all the attested patterns follow monotonic mappings.

Graf [13] proposes monotonicity as a formal universal of morphosyntax. The general idea is based on two criteria: I) each morphosyntactic domain comes with a base hierarchy (e.g. person: $1 < 2 < 3$), and II) the mappings from a hierarchy to surface forms must be monotonic. This dual specification already puts this approach at a major advantage because it combines substantive universals (linguistic hierarchies) and formal universals (monotonicity) to give a tighter characterization of natural language.

This paper proceeds as follows. Section 2 outlines a brief description of the notion of monotonicity. In the third section, I provide an analysis of tense syncretism based on verb paradigms. The interest in verb stem syncretism is threefold: I) it is problematic for the more restrictive *ABA generalization of Bobaljik [3], based on which in a paradigm, two forms cannot be identical to the exclusion of any forms in-between them. II) the attested patterns all follow monotonicity. III) the observed hierarchy of morphological tense is independently motivated by the logical temporal relations of Reichenbach [22].

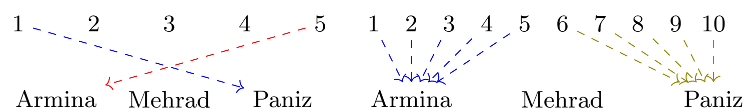
Section 3 presents the typology of gender resolution rules. Here, again, combining abstract algebra and the notion of monotonicity helps us understand the restricted set of the attested patterns. This suggests that there might be external ordering principles for gender, similar to what we see with tense. The crucial finding is that even though masculine and feminine genders should be ordered with respect to each other, the hierarchy does not favor one over another. In other words in a 3-gender system, both $m < n < f$ and $f < n < m$ can keep the system monotonic. One way to look at it is that gender is assigned along a path with two end nodes (masculine and feminine nodes). You can equally use the nodes to select objects (/assign gender). Neuter means 'neither' in Latin, so the neuter is always negatively defined as neither feminine not masculine. The objects that are not selected will be assigned to the middle part of the segment, which is the neuter.

Section 4 concludes the paper.

2 Monotonicity

Monotonicity is a mathematical property that corresponds roughly to the intuitive notion of order preservation. Suppose an ordering relation $\leq$ over a set $\{p, q, r, s, \ldots\}$ such that $p \leq r \leq s$. Then in a monotonic function, one cannot map both p and s to some A without also mapping r to A.

Let us consider a very intuitive example. Suppose A is a list of ordered numbers and B is a list of names in alphabetical order, then a function f from A to B is monotonic *iff* it preserves the relative order of elements. If f maps 1 to *Paniz* while 5 is mapped to *Armina*, f is not monotonic (this can be seen in crossing branches). However, mapping all the numbers between 1 and 5 to *Armina* and all the numbers from 5 to 10 to *Paniz* still preserves the original order and the function is monotonic.

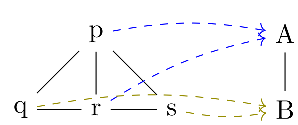


Now consider the *ABA generalization, which was first proposed by Bobaljik in [3] as an explanation for the absence of certain patterns in morphology and morphosyntax. Suppose an order of positive-comparative-superlative in an adjectival paradigm. For the paradigm of the English adjective *bad*, the first stem is A (*bad*), the second stem is B (*wors(e)*), and the third stem is again B (*wors(t)*). Using this notion of suppletion, one can abstract away from linguistic forms to see the underlying structure, where the positive and superlative cannot share a root distinct from the comparative, hence *ABA. If $\leq$ is a linear order, monotonicity corresponds exactly to the *ABA generalization.

Another linguistically familiar example of linear monotonicity is the ban against crossing branches in autosegmental phonology [12]. Autosegmental structures are usually presented in tiers, and within each tier segments are linearly ordered. The ban on crossing branches assures that all mappings from tones to segments follow the linear order of the two tiers.

But monotonicity is more general because it is also defined for partial orders. Suppose that $p \leq r \leq s$ as before, and $q \leq p$, but q is unordered with respect to r and s. Then a monotonic mapping could map p and r to A but q and s to B.

(1) *Monotonic mappings in a partially ordered structure*



Monotonicity has already been used as an abstract condition on morphological paradigms to explain typological gaps in adjectival gradation, case syncretism, pronoun syncretism, Person Case Constraint and Gender Case Constraint [13, 19]. I this paper, as we will see, the tense hierarchy is a partially ordered structure that is the same across all languages. But gender resolution rules can form both linear and partial structures depending on the number of gender values that are involved in resolution processes.

3 Tense Stem Syncretism

The *ABA generalization, introduced by Bobaljik [3], states that, given a fixed order of cells in a morphological paradigm, two cells cannot be syncretic to the exclusion of any cells between them. Bobaljik uses a specific notion of suppletion defined based on the form of the stems in a paradigm. Previously in 2, I briefly explained this with an example from adjectival gradation: the positive and superlative cannot share a root distinct from the comparative. English uses the ABB pattern for the adjective *bad* (*bad* - *wors(e) wors(t)*), but neither English nor any other language can ever use an ABA pattern here.[1]

Bobaljik accounts for *ABA in terms of feature containment. Within the Containment Hypothesis [3], this gap is due to the fact that a superlative morpheme does not directly attach to the adjectival root because the superlative always embeds a comparative. This means that if the positive form has the feature [ADJECTIVE], then the comparative form will be [[ADJECTIVE] COMPARATIVE]] which is itself a subset of the superlative form [[[ADJECTIVE] COMPARATIVE] SUPERLATIVE].

While Bobaljik is mostly concerned with the absence of ABA patterns in adjectival gradation, he also briefly discusses tense syncretism in verb stems. He draws on Wiese's analysis [24] of ablaut in German verbs to explain German stem alternations within the same framework. Bobaljik [3] extends this presentation of verb stem alternations to English verbs. He notes that no verbs in English and German display ABA patterns if one assumes an order of present-participle-past.

Table 1. Verb suppletion patterns in German & English

	PRS	PARTICIPLE	PAST	PATTERN
German	sprech-e	ge-sproch-en	sprach	ABC
	gieβ-e	ge**goss**en	**goβ**	ABB
	geb-e	ge-**geb**-en	gab	AAB
English	sing	sung	sang	ABC
	shine	**shone**	**shone**	ABB
	come	**come**	came	AAB
	walk	**walk**ed	**walk**ed	AAA

[1] In adjectival paradigms AAB pattern, where positive and comparative share a root distinct from superlative, is also missing cross-linguistically [3]. The absence of this pattern does not concern us here.

Wiese and bobaljik explain the gap in the data, i.e., the unattested identicality of the present and the past to the exclusion of the participle, using the Containment Hypothesis. Given the hierarchy present < Perfect participle < past, the present tense is the default with no featural specifications ([]), the participle is contained in the past sharing the [past] feature with it; and [[] PAST] and the preterite, the highest in the hierarchy, contains the [finite] feature in addition to its [past] feature [[[] PAST] FINITE] [3].

Based on Bobaljik's approach present and past are never syncretic to the exclusion of participle and more generally all tenses can be linearly ordered across languages so that no ABA patterns ever arise. The first assumption is compatible with the fact that there are Germanic languages which lack the past tense (preterite), which Bobaljik argues follows from its marked status. Furthermore, the participle can be used in constructions that are semantically related to the present tense. Leading us to conclude that it may share present features with the present, and past features with the preterite. "Such an intuition is particularly amenable to an analysis with overlapping decomposition [6], which could be represented schematically as [PRESENT], [PRESENT, PAST], [PAST]" [2]. In what follows, however, I show that Bobaljik's second prediction is only partially borne out once one considers a wider range of data: ABA patterns do arise if one also considers the future. This is problematic for Bobaljik's system, but can be readily explained via a partial order of morphological tenses in the monotonicity framework of Graf [13]. Crucially, this partial order is induced by the tense system of Reichenbach [22] and thus arises from third factor principles [7].

3.1 Corpus of Tense Syncretism

In order to extract the following data, I have used an opportunity sample of tense syncretism in the verbal paradigms of more than 20 languages. The languages under scrutiny represent a typologically-diverse sample belonging to the following families: Altaic, Germanic, Indo-Iranian, Romance, and Slavic, among others. For simplicity, I assume that two tenses have distinct stems if the stems differ for at least one person/number cell. This may result in multiple patterns in a single language. Also remember that the criterion for stem change is the specific notion of suppletion used by Bobaljik and introduced in the previous section.

The variety of ways verbal stems are paradigmatically related vary a lot, even within a language. The language sample I studied, rendered the following 10 patterns of verb stem syncretism (Table 2) with an ordering of past-participle-present-future.

In order to better understand the nature of the attested patterns and anticipate the kind of hierarchy we need, let's take a look at the unattested patterns. The total number of possible patterns for a paradigm with 4 cells is 15 (Bell number of 4), from which we already have 10. The remaining 5 unattested patterns are given in Table 3.

Out of all logically possible patterns, only 5 are unattested: ABAX (where future is A, B, or C), ABBA, and ABCA. The absence of ABAX patterns shows

Table 2. Attested patterns of tense syncretism

	Pattern	Example	past	participle	present	future
(1)	AAAA	Turkish	**gel**di	**gel**miʃ	**gel**ijor	**gel**eçek
(2)	AABB	Japanese	**ʃit**a	**ʃit**eita	**suru**	**suru**
(3)	AABA	Serbo-Croat	**hte**o sam	**hte**o	**ho**ću	**hte**ću
(4)	ABCD	German	warf	geworfen	wirf	werfen
(5)	AABC	Sindhi	**wa**yo	**wa**yo ho	wanje t^ho	wiindo
(6)	AAAB	French	**all**	**all**	**all**	ir
(7)	ABCC	Kurdish	xwɑrd	xoriɑ	**xweid**	**xweid**
(8)	ABCB	Spanish	fu	∅i	v	∅ir
(9)	ABBB	English	went	**go**ne	**go**	will **go**
(10)	ABBC	French	vin	**ven**	**ven**	viend-r

Table 3. Description of Unattested Patterns with Pst-Prf-Prs-Fut order

	Pattern Description	Linear Order
(1)	past = present; participle = future	ABAB
(2)	past = future; participle = present	ABBA
(3)	past = present = future; Separate root for participle	ABAA
(4)	past = present; Distinct roots for participle and future	ABAC
(5)	past = future; Distinct roots for participle and present	ABCA

that syncretism of present and past to the exclusion of participle is not attested. The behavior of future is problematic, though. While future is never syncretic with past to the exclusion of either present or participle, AABA and ABCB violate the *ABA generalization. But if one allows for partial orders, ABA patterns with future can be accounted for in terms of the monotonicity constraint [13].

Note that there is no way of totally ordering all four tenses such that there are no ABA configurations. Consider the attested pattern where past, participle and future are syncretic to the exclusion of present, as is the case in Persian and Serbo-Croatian. This pattern will be AABA with a Pst-Prf-Prs-Fut ordering and ABAA with a Pst-Prs-Prf-Fut ordering, both of which violate the ABA generalization. Our linear order won't violate *ABA only if it posits present at either end of the order. But any such order will be problematic for other attested patterns leading to the violation of *ABA. Once a specific connection between semantic and morphological tenses is made, the availability of some ABA patterns is due to the fact that the semantic relations between morphological tenses only induce a partial ordering.

Suppose that present $\leq$ participle $\leq$ past, and present $\leq$ future, but future is unordered with respect to participle and past. Then future can be syncretic with any one of the three tenses to the exclusion of others, allowing for a limited range of what appear to be ABA patterns. This is illustrated in Figure 1 for the

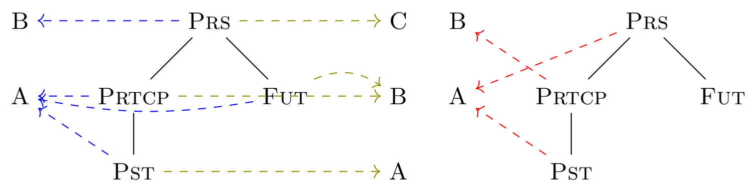


Fig. 1. Monotonic (left) and non-monotonic (right) mappings in tense syncretism

attested *ABA violations AABA and ABCB. The unattested ABAX patterns do not obey monotonicity (crossing branches).

This partial hierarchy might seem obvious given that the future and the participle both are intuitively associated with the present. However, Reichenbach's tense relations [22] provides a logical framework to motivate this morphological order.

3.2 Semantic Motivation: Reichenbach's System

In Reichenbach's system [22], tense denotes a three-way relation between speech time (S), event time (E) and reference time (R). The introduction of the notion of the reference time is considered Reichenbach's greatest contribution to the study of temporal relations. The position of R relative to S distinguishes the 3 tenses: 'past', 'present', and 'future'. The present time is the default setting in which S=E=R. Gradual shifts from this default point builds a partial hierarchy of temporal relations. R is located before S in the past (R $<$ S) and after S in the future (R $>$ S).

The position of E with respect to R distinguishes 3 further possibilities: 'posterior' (R $<$ E, viewing the situation E from an earlier point; looking forward), 'simple' (R $=$ E, used for the coincidence of R and E) and 'anterior' (R $>$ E, viewing the situation E from a later point; looking backward).

In the case of perfect, E is located before R. In past perfect, R precedes S (E $<$ R and R $<$ S). In present perfect, R overlaps S (E $<$ R and R $=$ S). Likewise, in future perfect, both S and E precede R (S $<$ R and E $<$ R).

All possible combinations involving a single time of speech (S) include three simple tenses (where R $=$ E), five anterior tenses (where E $<$ R), and five posterior tenses (where R $<$ E). Thus, the temporal system of a language could include up to 13 tenses. The actual number of tense realizations in each language depends on the number of grammaticalized combinations [22].

Here I argue that in addition to absolute tenses (present, past, future), perfect should also be part of the hierarchy of morphological tenses. Reichenbach and Comrie agree that perfect cannot be viewed as a canonical aspect since it tells us nothing about the internal temporal organization of the situation [9]. Perfect is like tense in that it locates an eventuality relative to some reference point. In the sentence *Paniz has eaten the cake*, there is an eventuality to the act of eating: it is done in the past. This makes the present perfect very similar to the simple past. In Reichenbach's terms, the simple past expresses a temporal precedence between the Speech time and the Reference time, while the perfect expresses a temporal precedence between the Event time and the Reference time. Another point of difference between the present perfect and the simple past will be apparent once we add a past-oriented adverb to our example: **Paniz has eaten the cake yesterday*. It is unexpected for an anterior temporal relation to be incompatible with a past-oriented adverb (Klein [18] refers to this situation as "the present participle puzzle").

More in support of positioning perfect among tense relations is the fact that perfect refers to a bundle of meanings that is maintained no matter what absolute tense it is associated with. Generally, three main readings are associated with perfects: The *universal reading* asserts that an eventuality holds for an interval of time; in the *experiential reading*, the eventuality holds for a proper subset of an interval; and finally, in a *resultative reading* the result of the eventuality holds at the speech time [21]. These readings make different claims about the location of the underlying eventuality, although in some languages only a subset of them is allowed. For example, in Greek perfect participles are marked as perfective and as a result the universal reading is not possible [21].

With these facts in order, I include perfects as part of the tense system (though this should not deny their aspectual properties in some languages).[2] Once one considers only those tenses that are morphologically realized across languages, the partial hierarchy of tenses emerges clearly.

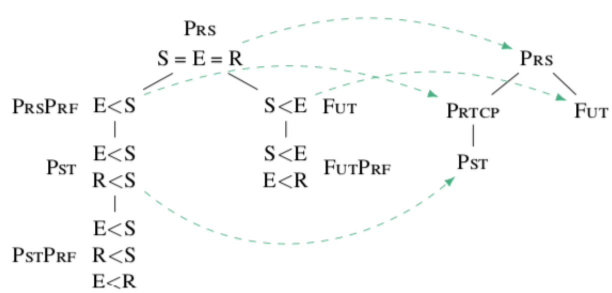


Fig. 2. The hierarchy of morphological tense motivated by Reichenbach's tense system

There are three reasons for identifying participle with present perfect. 1) The present tense refers to the default situation from which other tenses represent deviations [5]. Hence the past perfect and the future perfect follow from the

[2] There is always a great danger resulting from terminology. It is likely that, in some descriptive traditions, the term perfect is used for an aspectual rather than a tense distinction. This is true in the Semitic tradition, for example, where perfect and imperfect are used for what is likely perfective and imperfective. The *-ive* distinction is usually aspectual.

semantics of the present perfect, combined with an account of the past tense and the future tense [20].

2) The claim in (1) is verifiable by comparing the frequency rates of the perfects. The future perfect seems to be the least frequent among the perfects. A corpus-based study of English perfect constructions show that the present perfect is the most frequently used type of perfect in English [4].

3) The hierarchy of tense is an implicational hierarchy; if a language has a past perfect or a future perfect, it is very likely that it also has a present perfect (whereas the reverse does not necessarily hold). In this hierarchy, the present perfect has the least distance from the default point ($E = R = S$) with only one shift ($E < R, S$). The future perfect ($S < R < E$) and the past perfect ($E < R < S$) both undergo two shifts from the default. This results in the hierarchical ordering of the tenses.

In sum, I have shown that the future tense does give rise to apparent *ABA violations in verbal paradigms. But these are expected if one combines monotonicity [13]—a more general notion of *ABA— with a partial order of tenses in the spirit of Reichenbach [22]. This establishes a strong upper bound on the range of typological variation, with the only permitted but unattested pattern being syncretism of the past and future to the exclusion of other tenses.[3]

In the next section, I will introduce the variations of gender resolution rules as yet another instance where monotonicity sets a boundary on the attestability of certain morphological patterns.

4 Gender Resolution Rules

Resolved agreement is a term used to describe the predicate agreement with a subject made up of coordinated elements. The rules that determine the forms to be used are called resolution rules. Gender resolution rules are very diverse. This is mainly because they do not always have a unified semantic justification [11]. In French for instance, if two nominal heads, one feminine and one masculine, are conjoined, the resolved form is always masculine. Thus the resolution rules in French favor masculine agreement as the default gender. This is different from Icelandic or inanimate coordination in Rumanian where neuter and feminine are favored, respectively.

(2) [le garçon et la fille] sont compétents
[the boy.M and the girl.F] are competent.M.PL
'The boy and the girl are competent.' French

(3) [frægð-∅ og fram-i] eru tvíeggj-uð
[fame.F.SG and success.M.SG] are double.edged-N.PL
'Fame and success are double-edged.' Icelandic (Friðjónsson 1991: 90)

[3] Like the absence of AAB patterns in adjectival gradation, this might be due to independent factors [7].

(4) [usa si pretele] ele...
[door.F.the and wall.M.the] theyF.PL...
'The door and the wall, they...' Rumanian (Corbett 1991: 288)

4.1 Possibilities and Patterns

Just like tense syncretism, resolved gender stands out for how small the number of realized systems is relative to how many logically conceivable options there. In order to fully appreciate this point, let us take a moment to look at the combinatorics of resolved gender. Given k possible genders, there are k ways for any two genders and k^{k^2} resolution systems. Assuming that the order of elements in a coordination does not matter, the number of resolution systems equals $k^{k(k+1)/2}$. This is explained below using *triangular numbers.*

Assume that (a + b) is our coordination and the number of gender values in different languages are the exponents. In each line, the binomial expansion of each expression is given. We then abstract out of the mathematical details and replace them by a dot (•).

(5) *Triangular numbers*

$(a+b)^0$	1	•
$(a+b)^1$	a b	• •
$(a+b)^2$	a^2 2ab b^2	• • •
$(a+b)^3$	a^3 $3a^2b$ $3ab^2$ b^3	• • • •

The number of dots in each triangular pattern is its *Triangular Number*. The first triangle, a gender-less system (g=0) has just one dot. The second triangle (g=1) has another row with 2 extra dots, making 1 + 2 = 3 dots. The third triangle (g=2) has another row with 3 extra dots, making 1 + 2 + 3 = 6 dots. The fourth (g=3) has 1 + 2 + 3 + 4 = 10 dots.

The rule for calculating any triangular number is as follows. First, we rearrange the dots as bellow:

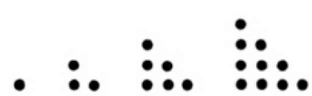


Fig. 3. Triangular numbers are the number of dots in each triangular pattern.

Then double the number of dots, and form them into a rectangle which has the same number of rows but has one extra column (to make this clear the two triangles are shown in green and red).

Now it is easy to see that the number of dots in a rectangle is $n(n+1)$ and the number of dots in a triangle is half that, i.e., $n(n+1)/2$.

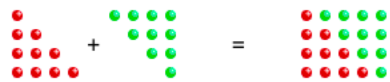
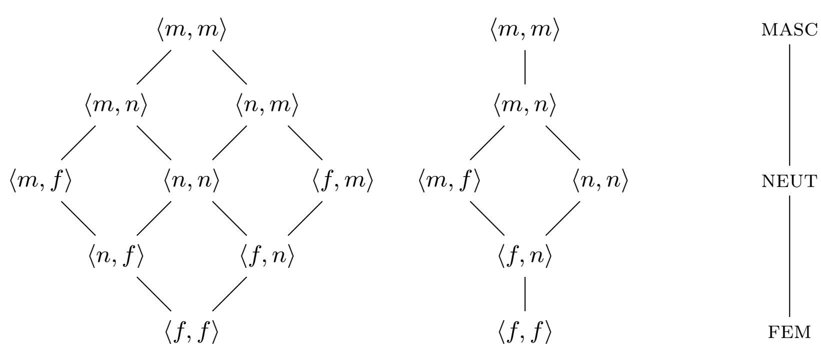


Fig. 4. Doubling the number of dots in each triangular pattern to form a rectangle.

For languages with 2 genders, this yields $2^{2(2+1)/2} = 2^{6/2} = 2^3 = 8$ possibilities. Yet only two patterns are attested in our sample of seven 2-gender languages (French, Spanish, Latvian, Hindi, Panjubi, Modern Hebrew and Rumanian). The same happens in 3-gender languages: out of $3^{3(3+1)/2} = 3^{6/2} = 3^6$ $= 729$ possibilities only 6 are realized. The space of logical possibilities quickly becomes quite large, as more and more genders are added: 8, 729, $1,048,576$ (million), $30,517,578,125$ (billion), $21,936,950,640,377,856$ (quadrillion), etc.

By definition, if A is some algebraic structure, the set of all functions X to the domain of A can be turned into an algebraic structure of the same type in an analogous way. Let us assume an underlying hierarchy of $f < n < m$ and construct a pointwise algebra to represent various gender combinations. At the top of the algebra $\langle m, m\rangle$ stands for the combination of two masculine genders. At the bottom, $\langle f, f\rangle$ represents the coordination of two feminine noun phrases. All other combinations are ordered between these two nodes. Since in a coordination the order of the coordinated elements does not matter (i.e., $\langle m, n\rangle = \langle n, m\rangle$), we remove all the symmetrically repeated nodes from the previous algebra to arrive at a simplified hierarchy.

(6) *The algebra of gender combinations*



Gender assignment lacks an overall semantic justification; thus the fact that there is an overwhelming uniformity of hierarchies across the available data sample is quite impressive.

4.2 Gender Resolution Patterns

The resolution systems discussed here are primarily based on Corbett's 1991 textbook on gender, which maps out the known variation in the gender systems of the world. It includes a comprehensive survey of gender systems with data from over 200 languages, which makes for a great typological study. I have filled the gaps in data from other sources on individual languages. Here I present five representative languages: French, Slovene, Latin, Tamil and Archi. French and Slovene are representative of languages that are argued to have syntactic resolution rules. Tamil and Archi are examples of a semantic type resolution. And finally Latin is described as a mixed type system where meaning and form are both involved in the patterns of resolution [11].

French Let us start with the simplest gender system we can consider. In French there are two genders, feminine and masculine. If in a coordination the conjuncts are of the same gender, then that gender will be used as the resolved form. If one conjunct is masculine and another is feminine, then a masculine form is used. Languages like French are quite common, e.g., Spanish, Latvian, Hindi, Italian, Panjabi, and Modern Hebrew, etc..

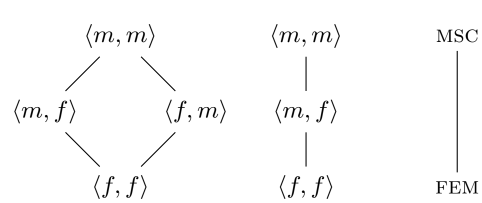
Table 4. Gender values and resolution in French

SG	PL
MSC	MSC
FEM	FEM

	MSC	FEM
MSC	M	M
FEM	M	F

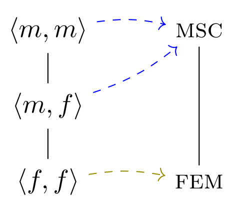
We start by building a hierarchical algebraic construction based on an underlying hierarchy of gender. Assuming $f < m$, we construct a pointwise algebra to represent the possible gender combinations. At the top of the algebra $\langle m, m\rangle$ stands for the combination of two masculine genders. At the bottom, $\langle f, f\rangle$ represents the coordination of two feminine noun phrases. The two other combinations are ordered between these two nodes. These two sets are the same, so we remove one of them to arrive at a more simplified hierarchy.

(7) *The gender hierarchy in French*



Elements of this algebra are then mapped into a hierarchy of plural genders. As we can see these mappings are all monotonic. In languages with a gender structure like French, it does not matter which gender is higher in the hierarchy. If you flip this structure you get the same kinds of mappings but in the reverse order.

(8) *Monotonic gender mappings in French*



Slovene Slovene has three numbers and three genders. The predicate agreement forms are given below. In this table, *bil* is the past active participle of the verb 'be' [11]. The dual forms will result only if the two conjoined noun phrases are singular. The gender resolution works the same for both dual and plural conjunctions. The gender system in Serbo-Croatian is similar, except that there is no dual there.

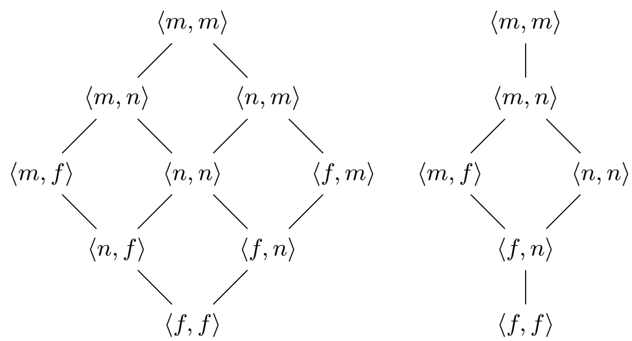
Table 5. Gender values in Slovene

	SG	DL	PL	
MSC	∅	a	i	MSC
FEM	a	i	e	FEM
NEUT	o		a	NEUT

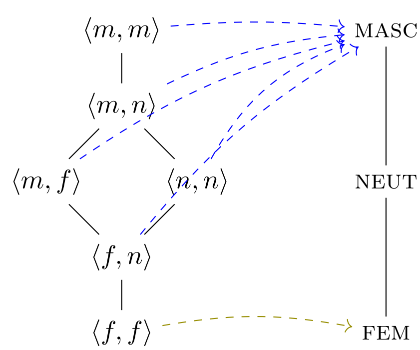
A masculine noun conjoined with a masculine will resolve in masculine. The same way, a feminine noun conjoined with a feminine will resolve in feminine. But a masculine noun conjoined with a feminine or with a neuter resolves in a masculine predicate. If a feminine and a neuter are conjoined, you will still find the masculine agreement on the predicate.

In order to explore the hierarchical structure of Slovene, once again we start from an underlying hierarchy of $f < n < m$ to construct a pointwise algebra and represent the gender combinations. In the simplified structure, all repeated nodes are removed.

(9) *The gender hierarchy in Slovene*



(10) *Monotonic mappings in Slovene*



In a sense, the resolved agreement in Slovene (and similar languages like Serbo-Croatian) favors the masculine. Feminine is only used if all conjuncts are feminine, and the neuter is not used at all. Interestingly, we will have the same monotonic mappings if we flip over the structure along with the hierarchy. As long as the neuter is in the middle, all the mappings are indeed monotonic.

Latin There are three genders in this language: masculine, feminine and neuter. Conjuncts of the same gender resolve in a form from the same gender. If conjuncts are of different genders, though, the criterion is purely semantic. Here the resolved form to be used depends on whether the nouns denote persons or not.

(11) *Resolution Rules in Latin*
a. Masculine is used if all conjuncts are masculine;
b. Feminine is used if all conjuncts are feminine;
c. Masculine is used if all conjuncts are human;
d. Otherwise, neuter is used.

The rules are ordered in this way because the masculine and the feminine genders are not semantically restricted to humans. This means that a human feminine in conjunction with a human masculine resolve in masculine rather than the default neuter.

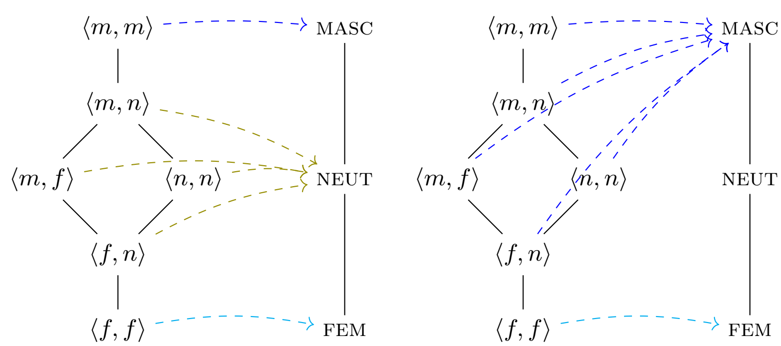
Table 6. Non-human and human resolution in Latin

Non-human	MSC	FEM	NEUT
MSC	M	N	N
FEM	N	F	N
NEUT	N	N	N

Human	MSC	FEM	NEUT
MSC	M	M	M
FEM	M	F	M
NEUT	M	M	M

In order to show these mappings, we divide the rules into two sets of human and non-human rules. Within each sub-system, all the mappings are monotonic.

(12) *Monotonic mappings in Latin non-human (left) and human (right)*



Tamil Dravidian languages are clear examples of semantic resolution. Tamil has three genders: masculine (for nouns denoting male rationals), feminine (for nouns denoting female rationals) and neuter (for non-rationals). The resolved forms, however, result in two forms only: rational and neuter. [4]

Table 7. Gender values in Tamil

SG	PL
MSC	RATIONAL
FEM	
NEUT	NEUT

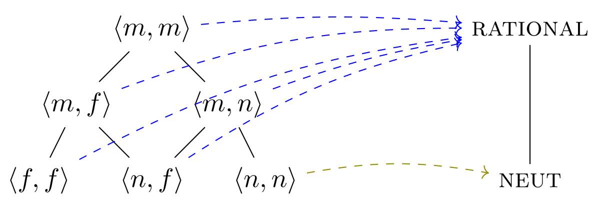
[4] The resolution rules in Telugu, another Dravidian language, is the same as Tamil. This happens despite the fact that in Telugu, feminine and neuter are not distinguished in the singular.

If, in a coordination structure, all conjuncts denote rationals, the rational form is used. If all conjuncts denote neuters, the neuter form should be used. The combination of a rational (feminine or masculine) with a neuter is generally avoided. But if ever allowed, the rational form is used.

(13) *Resolution Rules in Tamil*
a. Rational is used if all conjuncts are rational;
b. Neuter is used if all conjuncts are non-rational;
c. Otherwise, rational is used, although an alternative construction is preferred.

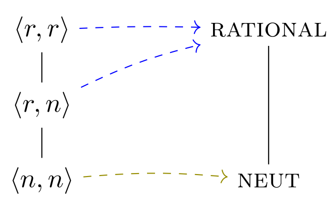
Over a hierarchy that places rational (including masculine and feminine) over neuter, all the mappings from controller genders to target genders are monotonic.

(14) *Monotonic mappings in Tamil*



The resolution rules in Tamil are not based on formal gender values but rather follow the two semantic values, RATIONAL and NEUTER. This means that there are only two classes of nouns in the plural. Hence we can reconstruct a hierarchy that only includes those two values in a linear order. The mappings over this hierarchy are all still monotonic.

(15) *Monotonic semantic mappings in Tamil*



Archi Caucasian languages are also famous for the semantic distinctions they make. Archi is a North-East Caucasian language.

(16) *Archi gender system*
I. male humans: God and other spiritual beings
II. females
III. most animals and some inanimate nouns
IV. some animals and most inanimate nouns

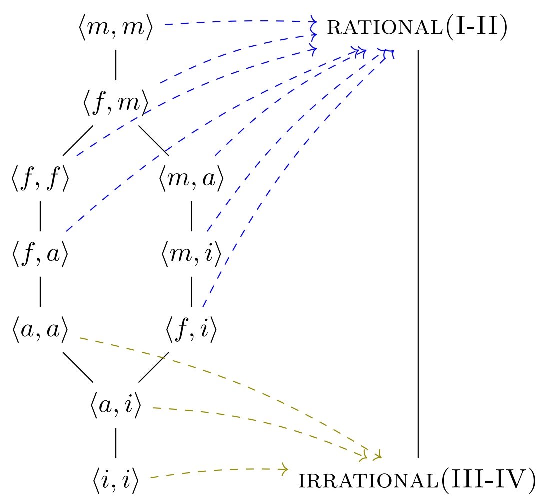
Table 8. Gender values in Archi

	SG	PL	
MSC	w	b/ib	rational
FEM	d		
ANIMATE	b	ib	irrational
INANIMATE	t		

(17) *Resolution Rules in Archi*
a. I/II is used, if there is at least one rational conjunct (R);
b. Otherwise, III/IV is used (IR).

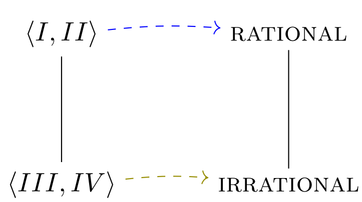
If the rational gender (including masculine and feminine) resides higher on the hierarchy relative to the irrational, then the mappings are monotonic.

(18) *Monotonic mappings in Archi*



Even though this account seems to work, gender and animacy are not the defining factors in Archi resolution. If we reduce the structure of conjoined noun phrases to rational and irrational entities, then a simple pattern emerges.

(19) *Monotonic semantic mappings in Archi*



In sum, I have shown that even though gender assignment in different languages greatly vary, the emerging gender hierarchies are substantially the same. The distinction made between the syntactic and semantic gender systems boils down to those gender values that are used in the resolved plural forms. In a syntactic system, e.g., French and Slovene, resolution rules are based on formal gender values. These systems mostly include feminine and masculine genders. As we saw, the nature of mappings remains the same as long as the feminine and masculine values reside on the two end nodes of the gender hierarchy. In a semantic system, e.g., Tamil and Archi, resolution rules are based on semantic values (RATIONAL vs IRRATIONAL), which results in a condensed hierarchy of gender that only includes those two values. Regardless of the hierarchy, in both system types, the resolution rules follow monotonic mappings from the base hierarchies to the output forms. Similarly, Latin, as a mixed gender type, combines two subsystems based on a semantic feature (the property of being human). Essentially for our account, both semantic sub-types use monotonic mappings.

5 Conclusion

In this article, we saw that the restrictions on morphosyntactic paradigms are systematically formalizable and have extralinguistic explanations. I have used a broad range of cross linguistic data to show this within two specific domains: tense syncretism in verb paradigms and resolved gender agreement. To this end, I have used the monotonicity account of Graf [13] that is based on an underlying hierarchy and the simple requirement that the mappings from these hierarchies to output forms are monotonic. I have derived the tense hierarchy from the logically rigorous framework of Reichenbach [22], while the gender hierarchy is directly motivated by typological data [11]. The findings reported in this article lends further empirical support to the idea that monotonicity is a linguistic universal that extends beyond semantics.

The major advantage of the presented account is that it combines substantive universals (linguistic hierarchies) and formal universals (monotonicity) to give a tighter characterization of morphosyntactic phenomena. The future research on this topic will be pursued with two main goals. First is to expand the range of morphosyntactic domains which requires careful treatment of typological data and motivated hierarchies. And secondly to integrate monotonicity with notions of subregular complexity in order to better understand the properties of attestable linguistic patterns.

References

1. Aksenova, A., Graf, T., Moradi, S.,: Morphotactics as tier-based strictly local dependencies. In: Proceedings of the 14th sigmorphon workshop on computational research in phonetics, phonology, and morphology, pp. 121–130. (2016)
2. Andersson, S.: (*)ABA in Germanic verbs. Glossa: A Journal of General Linguistics, **3**(1), 119. (2018) https://doi.org/10.5334/gjgl.733

3. Bobaljik, J.: Universals in comparative morphology: Suppletion, superlatives, and the structure of words, volume 50. MIT Press. (2012)
4. Bowie, J., Aarts, B.: Change in the English infinitival perfect construction. The Oxford Handbook of the History of English 200–210. (2012)
5. Bybee, J., Perkins, R., Pagliuca, W.: The evolution of grammar: Tense, aspect, and modality in the languages of the world, volume 196. University of Chicago Press. (1994)
6. Caha, P.: How (not) to derive a *ABA: The case of Blansitt's generalisation. Glossa **2**(84), 1–32. (2017) https://doi.org/10.5334/gjgl.348
7. Chomsky, N.: Three factors in language design. Linguistic Inquiry **36**, 1–22. (2005)
8. Chomsky, N.: Aspects of the theory of syntax. MITPress, Cambridge. (1965)
9. Comrie, B.: On Reichenbach's approach to tense. In Proceedings of the 17th Meeting of the Chicago Linguistic Society (CLS), **17**, 24–30. (1981)
10. Cooreman A., Goyvaerts D.: Universals in Human Language. A Historical Perspective. In: Revue belge de philologie et d'histoire, tome 58, fasc. 3, pp. 615–638. (1980) Langues et litteratures modernes – Moderne taalen letterkunde. https://doi.org/10.3406/rbph.1980.3295
11. Corbett, G.: Gender. Cambridge University Press. (1991)
12. Goldsmith, J.: Autosegmental phonology. MIT Press, London. (1976)
13. Graf, T.: Monotonicity as an effective theory of morphosyntactic variation. FSMNLP/MOL special issue. (2019)
14. Graf, T.: A Subregular bound on the complexity of lexical quantifiers. In: Schloder, J., McHugh, D., Roelofsen, F. (eds) Proceedings of the 22nd Amsterdam Colloquium, pp. 455–464. (2019)
15. Graf, T.: Why movement comes for free once you have adjunction. In: Edmiston, D., Ermolaeva, M., Hakguder, E., Lai, J., Montemurro, K., Rhodes, B., Sankhagowit, A., Tabatowski, M. (eds) Proceedings of CLS 53, pp 117– 136. (2018)
16. Heinz, J.: The computational nature of phonological generalizations. In: Hyman, L., Plank, F. (eds) Phonological Typology, Phonetics and Phonology, chapter 5, pp. 126–195. Mouton De Gruyter. (2018)
17. Icard, T.F., Moss, L.S.: Recent progress on monotonicity. In: Linguistic Issues in Language Technology. (2014)
18. Klein, W.: The present perfect puzzle. Language **68**, 525–552. (1992)
19. Moradi, S. *ABA generalizes to monotonicity. In: Baird, M., Pesetsky, J. (eds) Proceedings of NELS 49, V.2. GSLA. (2019)
20. Musan, R.: The present perfect in german: Outline of its semantic composition. Natural Language & Linguistic Theory **25**, 355–401. (2001)
21. Pancheva, R.: The aspectual makeup of perfect participles and the interpretations of the Perfect. In Perfect Explorations, volume 2. Walter de Gruyter. (2003)
22. Reichenbach, H.: Elements of symbolic logic. MacMillan, New York. (1947)
23. Sapir, E.: Language: An introduction to the study of speech. Harcourt, Brace, New York. (1921)
24. Wiese, B.: Form and function of verb ablaut in contemporary standard German. In: Sackmann, R. (ed) Studies in Integrational Linguistics. John Benjamins. (2005)