Chapter 2

One lexeme, many classes: Inflection class systems as lattices

Sacha Beniamine

Department of Linguistic and Cultural Evolution, Max Planck Institute EVA

This paper discusses the nature of inflection classes (ICs) and provides a fully implemented methodology to conduct typological investigations into their structure. ICs (conjugations or declensions) are sets of lexemes which inflect similarly. They are often described as partitioning the set of lexemes, but similarities across classes lead some authors to favor hierarchical descriptions. While some formalisms allow for multiple inheritance, where one class takes after two or more others, it is usually taken as an exceptional situation.

I submit that the structure of ICs is a typological property of inflectional systems. As a result, ICs are best modelled as semi-lattices, which by design capture non-canonical phenomena. I show how these monotonous multiple inheritance hierarchies can be inferred automatically from raw paradigms using alternation patterns and formal concept analysis. Using quantitative measures of canonicity, I compare six inflectional systems and show that multiple inheritance is in fact pervasive across inflectional systems.

1 Introduction

In some inflectional systems, the same morphosyntactic properties can be expressed differently across lexemes. Descriptions of the resulting inflection classes (declensions or conjugations) can take several forms. The simplest possibility is to use a partition of the set of lexemes into classes, as in Figure 2.1a. Possible partitions will differ in their granularities. Pedagogical grammars are often content with giving a broad classification in major classes. At the other end of the spectrum, various studies (e.g. Stump & Finkel 2013) presuppose a classification into numerous fine-grained classes.

Sacha Beniamine. 2021. One lexeme, many classes: Inflection class systems as lattices. In Berthold Crysmann & Manfred Sailer (eds.), *One-to-many relations in morphology, syntax, and semantics*, 23–51. Berlin: Language Science Press. DOI: 10.5281/zenodo.4729789

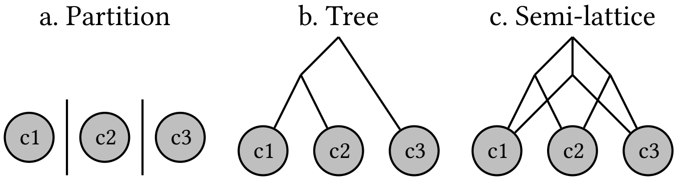

Figure 2.1: Three types of classification structures

Broad and fine-grained classifications can be linked by assuming a hierarchically-organized system of classes (Corbett & Fraser 1993; Dressler & Thornton 1996). In recent years, various efforts have been made towards inferring inflection class hierarchies automatically from paradigms (Brown & Hippisley 2012; Lee & Goldsmith 2013; Bonami 2014). While they use very different methodologies, most of these approaches converge on the use of tree-shaped hierarchies (Figure 2.1b.). Network morphology (Corbett & Fraser 1993; Brown & Hippisley 2012) uses richer structure through default inheritance and multiple inheritance of orthogonal properties, but does not allow for multiple inheritance in a single dimension (e.g. affixes).

In this paper, I argue that while "inflection classes" (IC) usually refers to either partitions (Figure 2.1a.) or trees (Figure 2.1b.), these make simplifications which overlook numerous relations between lexemes and hide structural properties that are in fact pervasive. I show that semi-lattices (Figure 2.1c.), where one subclass may belong to more than one superclass, are more faithful models of inflectional systems. I use formal concept analysis (Ganter & Wille 1998, hereafter FCA) to automatically infer semi-lattices of inflection classes for the verbal systems of French, English, Modern Standard Arabic, European Portuguese and Zenzontepec Chatino; as well as for the nominal system of Russian.[1]

I compare these systems using canonical typology. To do so, I provide formal definitions of inflectional structure and precise quantitative measures of inflectional canonicity, which can be computed automatically from a large inflected lexicon.

Inflection classes are usually taken as classes of lexemes or stems related by common affixes (Carstairs 1987; Carstairs-McCarthy 1991; Stump & Finkel 2013). However, alternations between stems also contribute to the expression of inflectional information. Segmentation in stems and affixes is useful to produce systems in constructive approaches (in the sense of Blevins 2006), where the goal

[1]The methodology described in this paper is fully implemented as part of the Qumín toolkit (Beniamine 2018) which can be accessed at: https://github.com/XachaB/Qumin. Qumín is distributed under GPLv.3.

is to generate the forms from a minimal grammar. Instead, I adopt here the abstractive approach (Blevins 2006) and attempt to account for all interesting generalizations. As a consequence, I take INFLECTIONAL BEHAVIOR to be relations between word-forms, or ALTERNATION PATTERNS, rather than affixes (Bonami & Luís 2014; Bonami & Beniamine 2016).

In the first section, I present partition- and tree-based accounts of ICs. Next, I motivate the need for multiple inheritance hierarchies as a more truthful model of ICs. In Section 3, I present FCA, which can be used to infer a semi-lattice of classes. The last section contrasts the properties of the IC lattices of six languages.

2 The structure of inflection class systems

IC systems are often described as a partition of a few broad classes of lexemes which share some of their inflectional behavior. Partitions of ICs are used both in pedagogical grammars and in many descriptive accounts. They usually count only a few classes. They are, as Matthews (1991: 129) puts it, "classes of lexemes that go together in respect of some inflection". This definition relies on the inflectional similarity between lexemes.

Corbett (1982) counts six nominal ICs (declensions) in Russian, which Table 2.1 illustrates by showing the full paradigm of one exemplar lexeme per class. I indicate frequencies based on counts in a lexicon of 1,239 nouns (Beniamine & Brown 2019) described in more detail in the appendix and in Beniamine 2018.

While it is usually thought that there is only one correct inventory of ICs in a given system, the number of classes is in fact often disputed, even in very well-documented languages. Corbett (1982: 202) highlights such disagreements in the case of Russian nouns: "The reader not familiar with the literature will quite reasonably expect a straightforward account of the paradigms in Russian. Tradition answers three, some writers claim four, and more recently it has been suggested that only two paradigms are required". The situation of Russian nouns is far from exceptional. One reason is that constructive and pedagogical analyses both usually strive for the shortest possible description. This leads to the merging of classes wherever possible, for example where distinct surface realizations can be abstracted away as allomorphy or predicted using semantic or grammatical properties of the lexemes. For example, Corbett shows that most descriptions of the ICs of Russian nouns merge together the classes ZAKON and VINO. The classes KOST' and PUT' are also usually merged, sometimes with the class VREMJA. In a similar fashion, Plénat (1987) provides a two-class analysis of the French verbal inflectional system, which is usually described as having three conjugations. To

Table 2.1: Six broad inflection classes of Russian in Roman transliteration, according to Corbett (1982: 203)

lexeme	zakon	vino	škola	kost’	put’	vremja
gloss	‘law’	‘wine’	‘school’	‘bone’	‘way’	‘time’
frequency	874	96	428	112	1	6
nom.sg	*zakon*	*vino*	*škola*	*kost’*	*put’*	*vremja*
acc.sg	*zakon*	*vino*	*školu*	*kost’*	*put’*	*vremja*
gen.sg	*zakona*	*vina*	*školy*	*kosti*	*puti*	*vremeni*
dat.sg	*zakonu*	*vinu*	*škole*	*kosti*	*puti*	*vremeni*
ins.sg	*zakonom*	*vinom*	*školoj*	*kost’ju*	*putem*	*vremenem*
loc.sg	*zakone*	*vine*	*škole*	*kosti*	*puti*	*vremeni*
nom.pl	*zakony*	*vina*	*školy*	*kosti*	*puti*	*vremena*
acc.pl	*zakony*	*vina*	*školy*	*kosti*	*puti*	*vremena*
gen.pl	*zakonov*	*vin*	*škol*	*kostej*	*putej*	*vremen*
dat.pl	*zakonam*	*vinam*	*školam*	*kostjam*	*putjam*	*vremenam*
ins.pl	*zakonami*	*vinami*	*školami*	*kostjami*	*putjami*	*vremenami*
loc.pl	*zakonax*	*vinax*	*školax*	*kostjax*	*putjax*	*vremenax*

do so, he merges the second and third conjugation using abstract phonological representations. Blevins (2004) reports that the nominal system of Estonian has been described as having between 26 and 400 “paradigms”, which can be merged in 6 to 12 ICs.

Going back to the data presented in Table 2.1, two shades of gray indicate some similarities across classes in each cell. All the classes share realizations for the dative, instrumental and locative plural. The class zakon shares the same endings as the class vino for the genitive, instrumental and locative singular. The locative singular is also identical to that of škola. zakon and škola also share the same endings in the nominative and accusative plural, while vino and škola both present no affixes in the genitive plural. The nominative and accusative singular of zakon, like those of kost’ and put’, show no affixes on the stem, etc. To these similarities in terms of endings or affixes, one could add similarities in terms of alternations, such as syncretisms: for example, the classes zakon, vino, kost’, put’ and vremja (but not škola) all present a syncretism between nominative and accusative singular. All these lexemes share a syncretism between the nominative and accusative plural.

A look at the Russian lexicon described in the appendix shows that the behavior of lexemes inside each class is less homogeneous than suggested by the table of exemplars. While all the exemplars shown above are inanimate and present the accusative-nominative syncretism, I found several lexemes with an accusative-genitive syncretism (typical of animates): 163 in the class ZAKON, 8 in the class VINO, 47 in the class ŠKOLA and 6 in the class KOST' (see Corbett & Fraser 1993: 129). Moreover, 76 lexemes of the class ZAKON, 3 of the class VINO and 6 of the class ŠKOLA have genitives in *-ej* rather than *-ov* or the bare stem.

Since similarity is gradient, it is difficult to determine how similar lexemes need to be to belong to the same class. Recent works in computational linguistics have attempted to decide on the best partition using minimal description length, either by comparing hand-written analysis (Walther & Sagot 2011) or by generating the analysis automatically from the data (Beniamine et al. 2017). But even when selected very rigorously, the resulting partitions are simplifications. They can be useful as pedagogical tools, or as compact constructive descriptions, but they do not account for all similarities between classes, nor for the internal variation in each class.

At the other end of the descriptive spectrum, various studies take ICs as very fine-grained partitions, where each distinction in inflectional behavior warrants a separate class. IC membership is then defined in terms of identity. Aronoff (1994: 64) defines an IC as "a set of lexemes whose members each select the same set of inflectional realizations". Carstairs-McCarthy (1994: 739) provides two definitions of a paradigm:

> (1) PARADIGM$_1$: the set of combinations of morphosyntactic properties or features (or the set of "cells") realized by inflected forms of words (or lexemes) in a given word-class (or major category or lexeme-class) in a given language.
> (2) PARADIGM$_2$: the set of inflectional realizations expressing a paradigm$_1$ for a given word (or lexeme) in a given language.

Based on these definitions, he offers a very similar definition of ICs: "a set of words (lexemes) displaying the same paradigm$_2$ in a given language". Applied to realistic datasets, these definitions yield a high number of classes, many of which are often very small. Stump & Finkel (2013) report 72 ICs for French verbs, while Bonami (2014), Beniamine et al. (2017) and Beniamine (2018) find up to 97 classes.[2] For Russian nouns, Beniamine (2018) identifies 159 ICs based on identity of surface segmental inflectional behavior (not counting stress patterns). While,

[2]While they all base their computations on the Flexique lexicon (Bonami et al. 2014), differences across accounts are due both to different methodologies and to corrections that have been made in the lexicon since its publication.

by definition, these classes do not show any internal heterogeneity, enumerating them does not account for any similarities across classes.

Descriptive grammars often make use of explicit or implicit tree-shaped hierarchies when they provide several granularity levels. For example, the French pedagogical grammar Bescherelle (Arrivé 2012) describes three ICs, each exemplified by numerous verbal exemplars (one per page) and finer variations in footnotes. These can be interpreted as a three-level hierarchy. Campbell (2011) describes the ICs in Zenzontepec Chatino, an Oto-Manguean language spoken in Oaxaca, by a three-level hierarchy presented in Figure 2.2. Zenzontepec Chatino expresses inflection through prefixes and has only four paradigm cells: potential, habitual, progressive and completive. Figure 2.2 shows common prefixes for each node of the hierarchy. The notation "[lam]" marks the laminalization of initial [t] in class Bt. Campbell (2011) shows identical underlying prefixes for classes Au and Ac, but they differ on the surface. Class Bc presents a stem-initial alternation between *y-* and *ch-*. Since class C2 presents several distinct affixes, it could be further divided in two distinct classes. The first level of Campbell's (2011) classification is not based on similarity alone, but inherits from Kaufman's (1989) description of Zapotec ICs.

Dressler & Thornton (1996), Kilani-Schoch & Dressler (2005) and Dressler et al. (2008) use the term "macroclass" for the broad ICs based on similarity and "microclass" for the fine-grained ICs based on identity of inflectional behavior. They link both in tree-shaped hierarchies, in which any node can be seen as an IC. Microclasses form the leaves of the hierarchy, while macroclasses form the first level below the root. Any number of intermediate classes is possible. In Kilani-Schoch & Dressler's (2005) approach to French, the macroclasses are not based on similarity alone, but instead they constitute a bipartition between productive and unproductive patterns. Each IC is motivated by common inflectional patterns, written as implicative statements which the authors call "paradigm structure conditions". These conditions are inherited by default.

In network morphology (Corbett & Fraser 1993; Brown & Hippisley 2012), ICs are also represented by a tree-shaped default inheritance hierarchy. The analyses are constructive: couched in the DATR formalism, each node specifies affixal rules. The grammar is designed to generate surface forms. Default inheritance has two main advantages. First, it allows for more compact representations by limiting repetitions and the overall number of nodes in the hierarchy. Second, it gives the notion of regularity a natural status: a node which rewrites a default is exceptional relative to the ancestor which stipulated the default rule.

Going back to Russian nouns, Brown (1998) count four main ICs which correspond to the first four declensions described by Corbett (1982): ZAKON (I),

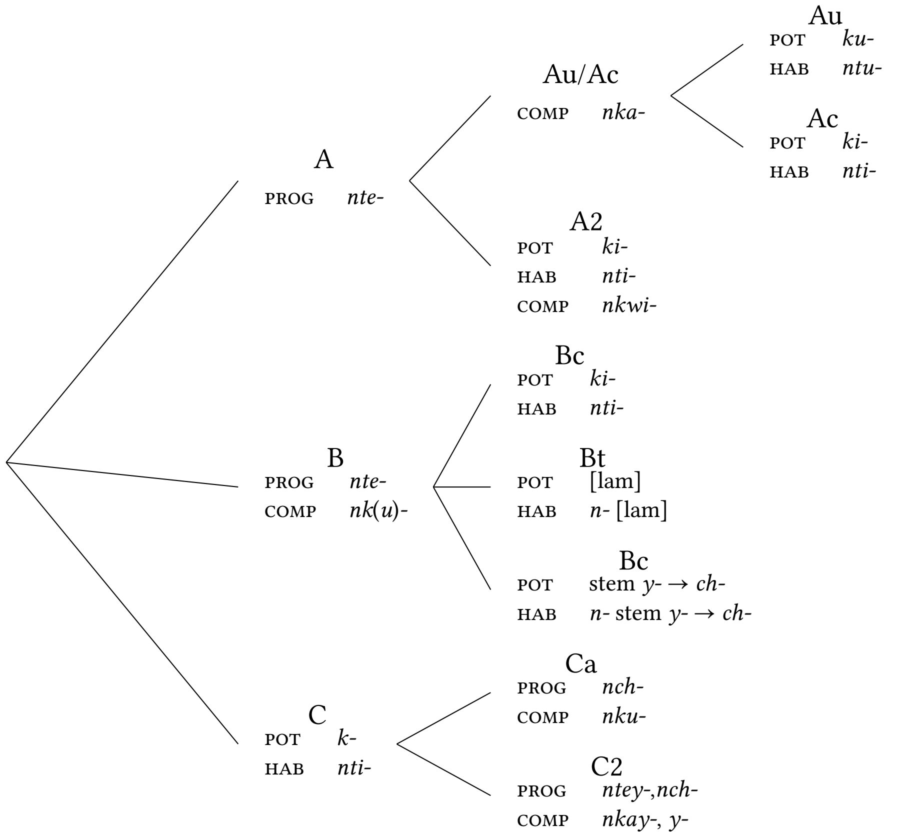

Figure 2.2: Inflection class tree in Zenzontepec Chatino verbs according to Campbell (2011: 229)

ŠKOLA (II), KOST' (III) and VINO (IV). Brown (1998) argues in favor of the hierarchical structure summarized in Figure 2.3. In the inflectional tree, the leaves N_I to N_IV stand for each of the four ICs. The root is the node MOR_NOMINAL, which also spans adjectives (which I will ignore for the purpose of this paper). It defines common properties between nouns and adjectives, as well as two default values: a zero affix in the nominative singular and an *-i* ending in the nominative plural. The term EVALUATION denotes the usage of a realization function which takes as input morphological properties of a lexeme and can assign distinct values to lexemes belonging to the same class. The node MOR_NOM specifies a thematic vowel characteristic of all nouns, a default affixal value for the locative singular and a default syncretism between dative and locative singular. There is only one intermediate node, N_O. It manifests properties shared between classes I and IV.

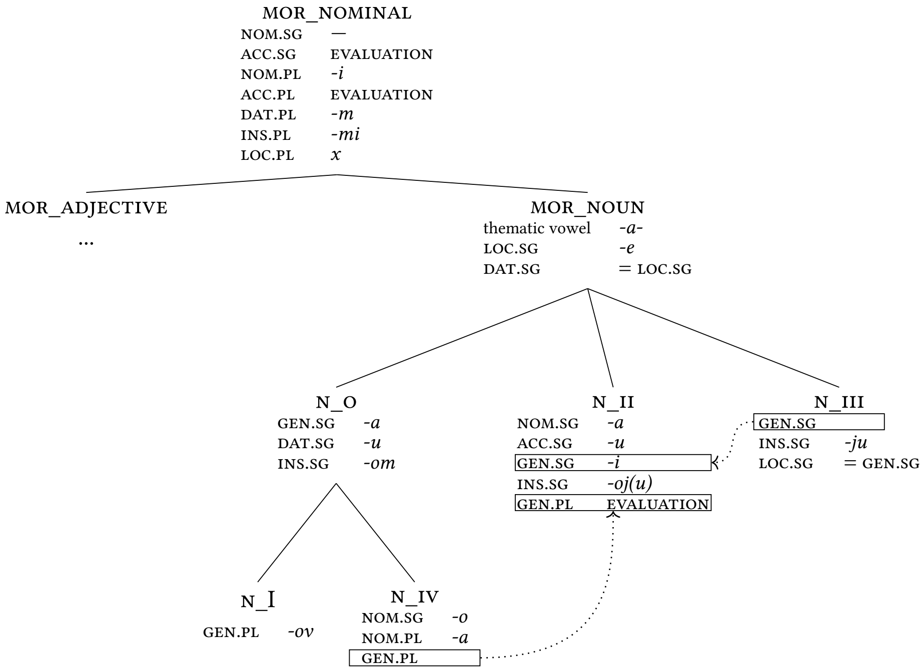

Figure 2.3: DATR hierarchy for Russian nouns according to Brown (1998: Theory B, 128 et seq.)

In Brown's (1998) account, some commonalities between classes are not modeled through the tree structure itself but by direct references across classes for specific cells. These references are indicated in Figure 2.3 by dotted arcs between framed cells. For example, the genitive plural of class IV is formed by using the evaluation functions of the genitive plural in class II. The need for this second mechanism highlights the inadequacy of a tree structure to express all similarities between ICs. In addition, while default inheritance is useful for producing a compact hierarchy, it hides the exact span of the default rules. In the following section, I show how a richer hierarchy can account more naturally for IC structure in an abstractive approach.

3 Noncanonical systems as inflection class lattices

In the previous section, I showed that partitions and tree structures have been used to describe inflectional systems even when their similarity structure is more complex than these descriptive devices can account for. It is, however, conceivable that some inflectional systems do conform to the structure of either a partition or a tree.

Corbett (2009) chooses this particular ideal structure as a canonical point of comparison for typological investigation. He defines canonical IC systems as following the principle of distinctiveness (Corbett 2009: 3), which can be evaluated using four criteria:

> PRINCIPLE I (distinctiveness): Canonical inflection classes are fully comparable and are distinguished as clearly as is possible. [...]
>
> *criterion 1* In the canonical situation, forms differ as consistently as possible across inflectional classes, cell by cell.
>
> *criterion 2* Canonical inflectional classes realize the same morphosyntactic or morphosemantic distinctions (they are of the same structure).
>
> *criterion 3* Within a canonical inflectional class each member behaves identically.
>
> *criterion 4* Within a canonical inflectional class each paradigm cell is of equal status.

From these criteria, it follows that in a canonical system, there are no similarities between classes. If two classes were to have a common exponent or alternation pattern, they would violate criterion 1. Moreover, the cells affected by common patterns would then be less predictive of the ICs than other cells, which violates criterion 4. According to criterion 2, a canonical system of ICs can have only one form per paradigm cell and lexeme. Defective lexemes, which lack forms for certain cells and overabundant lexemes, which have more than one possible form for certain cells, violate criterion 2. Finally, criterion 3 means that all classes are microclasses: they are based on identity. In a canonical system, micro- and macroclasses coincide. The system then truly has the shape of a partition (or a one-level tree, with classes as leaves and the whole system as root).

If real systems mostly conformed to the canonical ideal – which is not usually expected – then it would be adequate to model them using partitions. If, however, noncanonicity is the norm, then more expressive models are required. Since partitions and trees make the assumption of a certain degree of canonicity, these models are not suited to evaluating a system's position in the canonical space.

Figure 2.4 shows the same four ICs of Russian nouns as in Figure 2.3, now arranged as a partition, with each class characterized by affixes. While the shape of this classification is that of a partition, it is obvious from the numerous repetitions that it is not the structure of the data. The use of a partition masks the system's noncanonicity.

N_I		N_IV		N_II		N_III	
NOM.SG	—	NOM.SG	*-o*	NOM.SG	*-a*	NOM.SG	—
ACC.SG	—	ACC.SG	*-o*	ACC.SG	*-u*	ACC.SG	—
GEN.SG	*-a*	GEN.SG	*-a*	GEN.SG	*-i*	GEN.SG	*-i*
DAT.SG	*-u*	DAT.SG	*-u*	DAT.SG	*-e*	DAT.SG	*-i*
INS.SG	*-om*	INS.SG	*-om*	INS.SG	*-oj*	INS.SG	*-ju*
PREP.SG	*-e*	PREP.SG	*-e*	PREP.SG	*-e*	PREP.SG	*-i*
NOM.PL	*-i*	NOM.PL	*-a*	NOM.PL	*-i*	NOM.PL	*-i*
ACC.PL	*-i*	ACC.PL	*-a*	ACC.PL	*-i*	ACC.PL	*-i*
GEN.PL	*-ov*	GEN.PL	—	GEN.PL	—	GEN.PL	*-ej*
DAT.PL	*-am*	DAT.PL	*-am*	DAT.PL	*-am*	DAT.PL	*-am*
INS.PL	*-ami*	INS.PL	*-ami*	INS.PL	*-ami*	INS.PL	*-ami*
PREP.PL	*-ax*	PREP.PL	*-ax*	PREP.PL	*-ax*	PREP.PL	*-ax*

Figure 2.4: Partition of four Russian inflection classes

The tree structure in Figure 2.3 assumes an intermediate level of canonicity and is also insufficient to express all the similarities between these ICs. The analysis in Figure 2.5 accounts for each point of similarity between the four classes in Figure 2.4. This analysis does not allow any other inheritance mechanism than the hierarchy itself: as a consequence, it does not contain defaults, rules of referral, or evaluation functions.[3]

In contrast to a tree, the hierarchy in Figure 2.5 displays multiple inheritance. For example, class I has two parents. From one parent, it inherits the absence of affix in the nominative and accusative singular, and from the other parent, it inherits its values for its genitive, dative and instrumental singular affixes. This structure is a lattice. Lattices have been used to model linguistic structures, for example in the type hierarchy of HPSG (Flickinger 1987; Pollard & Sag 1994; Ginzburg & Sag 2000) or in phonological feature hierarchies (Chomsky & Halle 1968; Frisch 1997). Since ICs can be seen as "classes of lexemes that share similar morphological contrasts" (Brown & Hippisley 2012: 4), I call any node of this hierarchy an inflection class, not only its leaves. In consequence, one lexeme can belong to many inflection classes.

[3]For this small example, in the interest of legibility, I take classes I to IV to be microclasses, and I exclude some lexemes which Brown (1998) accounts for using evaluation functions. The hierarchy can, however, be extended to account for all microclasses of a system. For the same reason, I ignore adjectives in this example.

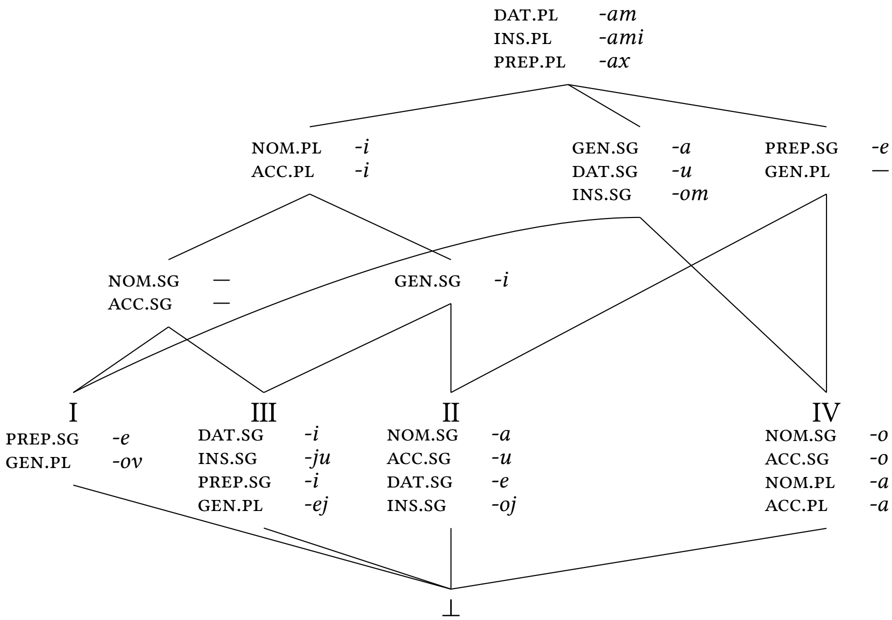

Figure 2.5: Lattice of four Russian inflection classes

In the hierarchy in Figure 2.5, each intermediate node represents a similarity point between lower nodes. All the similarities are represented.

In this hierarchy, classes are ordered by increasing generality. Higher nodes hold more general information than lower nodes: their value is less specified and they encompass more classes. Information specified on the leaves, labeled here with Roman numerals, is entirely distinctive: it is specific to each microclass.

All the information relating to a class can be read by going through each of its ancestors. The common information shared by any two classes can be found by searching for their least upper bound, also called JOIN. If any values are common to all ICs, they are specified at the highest node, which is called the SUPREMUM.

Symmetrically, one can find the common subclass of two nodes by searching their greatest lower bound, also called MEET. There is only one such child. For example, the node {NOM.PL *-i*, ACC.PL *-i*} and the node {PREP.SG *-e*, GEN.PL -} have the class II for greatest lower bound. The lowest node in the hierarchy, or INFIMUM, noted ⊥, is the MEET between any pair of the leaves, because no lexeme can belong to more than one of these microclasses. Since the infimum is always present and never brings any relevant information, I will sometimes omit it.

This hierarchy displays precisely what distinguishes this system from the canonical situation. While canonical ICs have only microclasses and a supremum (root) as is the case in Figure 2.4, the structure in Figure 2.5 has five more intermediate classes. A hierarchy of canonical ICs has a depth of 1, but the lattice from Figure 2.5 has a depth of 3 (the longest path from the root to a microclass follows three edges). Finally, while the canonical situation shows only simple inheritance, classes in this hierarchy have on average 1.4 direct parents.

This section showed that a partition model makes the prediction that the classes are canonical, which isn't the case of the partial systems previously discussed. A tree structure allows some sharing across microclasses, but still makes a prediction on their canonicity. It assumes that while classes can share some properties, there is no heteroclite sharing. HETEROCLISIS is usually taken to occur when the paradigm of a small IC is split in such a way that it follows two or more separate distinct ICs (Corbett 2009). The term can be extended in order to describe any class which displays multiple inheritance. Modeling IC systems as lattices will allow us to observe the amount of heteroclite sharing and quantify IC canonicity.

4 Inferring inflection class lattices with formal concept analysis

To automatically produce an inflectional lattice, I use formal concept analysis (Ganter & Wille 1998). This mathematical formalism allows us to study all interesting relationships between sets of objects (in this case lexemes, or microclasses) and their properties by ordering them in a CONCEPTUAL HIERARCHY. This section describes the basics of FCA, illustrated on a few sub-paradigms of English verbs shown in Table 2.2.

Table 2.2: Some sub-paradigms of English verbs

lexeme	PST	PST.PART	PRS
DRIVE	/drəˑʊv/	/drɪvn̩/	/draˑɪv/
RIDE	/rəʊd/	/rɪdn̩/	/raˑɪd/
BITE	/bɪt/	/bɪtn̩/	/baˑɪt/
FORGET	/fəgɒt/	/fəgɒtn̩/	/fəgɛt/

In the previous sections, I took inflectional attributes to be affixes. However, using affixes to automatically assess similarity of inflectional behavior is problematic (Beniamine 2018): first, they do not account for all similarities between paradigms (Beniamine et al. 2017), second, ignoring stem alternations excludes a large number of relevant inflectional properties (Bonami & Beniamine 2016).

Last but not least, there is no consensual method for segmenting wordforms into affixes (Spencer 2012). For these reasons, I prefer to rely on alternation patterns (Bonami & Luís 2014; Bonami & Beniamine 2016). Using the Qumín software (Beniamine 2018; 2017), they can be automatically inferred from raw forms in a language-agnostic way. Qumín takes as its input a fully inflected lexicon structured as a paradigm table (as in Table 2.2). Forms are transcribed in phonemic notation, and the lexicon is accompanied by a decomposition of each phoneme into minimal features (see the appendix). Both the structure of the paradigm table and the transcription constitute idealizations.

Table 2.3 shows the alternation patterns deduced from pairwise alternations from Table 2.2. For example, the alternation between /fəgɛt/ (PRS) and /fəgɒt/ (PST) follows the bidirectional alternation pattern _ɛ_ ⇌ _ɒ_, where "_" indicates the presence of constant material in the form.[4] The empty string is written ϵ .

Table 2.3: Alternation patterns for the subparadigms from Table 2.2

lexeme	PST.PART ⇌ PRS	PST.PART ⇌ PST	PRS ⇌ PST
RIDE	_ɪ_n̩ ⇌ _aˑɪ_	_ɪ_n̩ ⇌ _əˑʊ_	_aˑɪ_ ⇌ _əˑʊ_
DRIVE	_ɪ_n̩ ⇌ _aˑɪ_	_ɪ_n̩ ⇌ _əˑʊ_	_aˑɪ_ ⇌ _əˑʊ_
BITE	_ɪ_n̩ ⇌ _aˑɪ_	_n̩ ⇌ _ϵ	_aˑɪ_ ⇌ _ɪ_
FORGET	_ɒ_n̩ ⇌ _ɛ_ϵ	_n̩ ⇌ _ϵ	_ɛ_ ⇌ _ɒ_

Table 2.3 defines a relationship between lexemes and alternation patterns. It can be written as an incidence matrix, that is, a cross table where objects are indicated in rows and attributes in columns. A cross in a cell indicates that the object in this row instantiates the property in this column. Such a table is called a FORMAL CONTEXT. Table 2.4 shows the context for the subparadigms of English verbs from Table 2.2. I take objects to be lexemes and attributes to be combinations of a pair of cell and an alternation pattern.

A FORMAL CONTEXT is a triplet $\langle X, Y, I\rangle$, where X and Y are non-empty sets and I is a binary incidence relation between X (objects, in row) and Y (attributes, in column): $I \subseteq X \times Y$. For all objects $x \in X$ and all attributes $y \in Y$:

- $\langle x, y\rangle \in I$ indicates that the object x has the attribute y,
- $\langle x, y\rangle \notin I$ indicates that x does not have y.

[4] I report here a simplified view of alternation patterns, specifying only the alternating material as well as its position in the word. Qumín (Beniamine 2017; 2018) also extracts a detailed set of phonotactic constraints on the context of the changes. I omit it here in all examples for simplicity.

Table 2.4: Formal context for Table 2.3.

	PST.PART⇌PRS		PST ⇌ PRS			PST.PART ⇌ PST	
	_ɪ_n̩ ⇌ _aɪ_	_ɒ_n̩ ⇌ _ɛ_ϵ	_aɪ_ ⇌ _əʊ_	_aɪ_ ⇌ _ɪ_	_ɛ_ ⇌ _ɒ_	_ɪ_n̩ ⇌ _əʊ_	_n̩ ⇌ _ϵ
DRIVE	×		×			×	
RIDE	×		×			×	
BITE	×			×			×
FORGET		×			×		×

In the context table $\langle X, Y, I \rangle$, there is a cross at coordinates i, j if and only if $\langle x_i, y_i \rangle \in I$. Ganter & Wille (1998) write $\langle x, y \rangle \in I$ as xIy.

For any subset of objects $A \subset X$, we are interested in the attributes they have in common. For any subset of attributes $B \subset Y$, we are interested in the objects which instantiate them. Let us define two operators, "↑" and "↓" (Bělohlávek 2009: 6–7), such that:[5]

- The operator ↑ maps objects (subsets of X) to attributes (subsets of Y). $A \uparrow$ is defined as the subset of all attributes shared by the objects in A:

$$\uparrow : 2^X \to 2^Y \text{ and } A \uparrow = \{y \in Y | \text{ for each } x \in A : xIy\}$$

- The operator ↓ maps attributes (subsets of Y) to objects (subsets of X). $B \downarrow$ is defined as the subset of all objects which share all attributes in B:

$$\downarrow : 2^Y \to 2^X \text{ and } B \downarrow = \{x \in X | \text{ for each } y \in B : xIy\}$$

If the objects in A have no common attribute, then $A \uparrow = \varnothing$. Similarly, if no object shares all the attributes from B, then $B \downarrow = \varnothing$. Consequently, $\varnothing \uparrow = Y$ and $\varnothing \downarrow = X$.

[5]This notation is that of Bělohlávek (2009). Ganter & Wille (1998) represents both operators by $'$, writing the sets $A \uparrow$ and $B \downarrow$ as A' and B', respectively. I prefer Bělohlávek's (2009) more explicit convention.

For example, the following equalities can be deduced from Table 2.4:[6]

(1) {RIDE, DRIVE}↑ = {_ɪ_n̩ ⇌ _aˑɪ_ , _aˑɪ_ ⇌ _əˑʊ_, _ɪ_n̩ ⇌ _əˑʊ_}

(2) {_ɪ_n̩ ⇌ _aˑɪ_, _aˑɪ_ ⇌ _əˑʊ_, _ɪ_n̩ ⇌ _əˑʊ_}↓ = {DRIVE, RIDE}

(3) {_ɪ_n̩ ⇌ _aˑɪ_}↓ = {DRIVE, RIDE}

(4) {_aˑɪ_ ⇌ _ɪ_, ɛ ⇌ ɒ}↓ = ∅

These equalities can be read directly in Table 2.4. The lexemes DRIVE and RIDE share all of their attributes (1). The three patterns they share are only shared by them (2). The pattern _ɪ_n̩ ⇌ _aˑɪ_ is also shared by only DRIVE and RIDE (3). Finally, the operator ↓, applied to the concurrent contradictory pattern for PST ⇌ PRS, produces the empty set (4) unless there are overabundant lexemes instantiating these patterns.

Using these operators, we can define a FORMAL CONCEPT. A formal concept in the context $\langle X, Y, I\rangle$ is a pair $\langle A, B\rangle$ of a set of objects $A \subseteq X$ called the EXTENSION of the concept and a set of attributes $B \subseteq Y$ called the INTENSION of the concept, such that $A \uparrow = B$ and $B \downarrow = A$. In other words, the objects from A have in common exactly the attributes from B, no more, no less. Reciprocally, the attributes from B are common to all objects in A, no more, no less.

For example, ⟨{DRIVE,RIDE}, {_ɪ_n̩ ⇌ _aˑɪ_, _aˑɪ_ ⇌ _əˑʊ_, _ɪ_n̩ ⇌ _əˑʊ_ }⟩ is a formal concept, because we have both (1) and (2). However, ⟨{DRIVE,RIDE}, {_ɪ_n̩ ⇌ _aˑɪ_ }⟩ is not a formal concept, because despite (3), the opposite isn't true, as {_ɪ_n̩ ⇌ _aˑɪ_ } is only a subset of {RIDE, DRIVE}↑ (1).

From the incidence table, it is possible to produce a list of all the formal concepts. Examples (5) through (11) list all the concepts present in Table 2.4:

(5) ⟨ ∅, {_ɒ_n̩ ⇌ _ɛ_ϵ, _ɪ_n̩ ⇌ _aˑɪ_, _n̩ ⇌ _ϵ, _ɪ_n̩ ⇌ _əˑʊ_, _aˑɪ_ ⇌ _əˑʊ_, _aˑɪ_ ⇌ _ɪ_, _ɛ_ ⇌ _ɒ_ } ⟩

(6) ⟨ {BITE}, {_ɪ_n̩ ⇌ _aˑɪ_, _n̩ ⇌ _ϵ, _aˑɪ_ ⇌ _ɪ_} ⟩

(7) ⟨ {FORGET}, {_ɒ_n̩ ⇌ _ɛ_ϵ, _n̩ ⇌ _ϵ, _ɛ_ ⇌ _ɒ_} ⟩

(8) ⟨ {RIDE, DRIVE}, {_ɪ_n̩ ⇌ _aˑɪ_, _ɪ_n̩ ⇌ _əˑʊ_, _aˑɪ_ ⇌ _əˑʊ_} ⟩

(9) ⟨ {BITE, FORGET}, {_n̩ ⇌ _ϵ} ⟩

[6] In all examples below and in Figures 2.6 and 2.7, morphosyntactic attributes for the alternation patterns are not repeated. This is a shortcut, as our attributes are actually combinations of a pair of cells and an alternation pattern. In our small example, where only seven patterns are considered, this omission does not lead to ambiguity. However, due to syncretism, this would not be the case for most real systems.

(10) ⟨ {RIDE, DRIVE, BITE}, {_ɪ_n̩ ⇌ _aˑɪ_} ⟩

(11) ⟨ {RIDE, DRIVE, BITE, FORGET}, ∅ ⟩

I noted, when observing the lattice in Figure 2.5, that classes were ordered by specificity. Concepts can also be ordered according to their specificity. Given two concepts $\langle A_1, B_1\rangle$ and $\langle A_2, B_2\rangle$ in $\langle X, Y, I\rangle$, $\langle A_1, B_1\rangle$ is more specific than $\langle A_2, B_2\rangle$ if and only if A_1 is a subset of A_2, which entails that B_2 is a subset of B_1. Let us call $\langle A_1, B_1\rangle$ a subconcept of $\langle A_2, B_2\rangle$:

$$\begin{aligned}\langle A_1, B_1\rangle \leq \langle A_2, B_2\rangle &\iff A_1 \subseteq A_2\\ &\iff B_2 \subseteq B_1\end{aligned}$$

In other words, the subconcept contains only some of the objects (lexemes) from the more general concept, but more attributes (patterns). For example, the concept in example (8) is a subconcept of the concept in example (10). The subconcept has one fewer lexeme and two more patterns.

If $\langle A_1, B_1\rangle \leq \langle A_2, B_2\rangle$ and there are no concepts $\langle A_i, B_i\rangle$ in $\langle X, Y, I\rangle$ such that $\langle A_1, B_1\rangle \leq \langle A_i, B_i\rangle \leq \langle A_2, B_2\rangle$, then $\langle A_1, B_1\rangle$ is an immediate lower neighbor of $\langle A_2, B_2\rangle$, which is written: $\langle A_1, B_1\rangle \prec \langle A_2, B_2\rangle$.

The collection of all formal concepts of a context $\langle X, Y, I\rangle$, together with the order relation $\leq$, form the CONCEPT LATTICE of $\langle X, Y, I\rangle$, written $\mathcal{B}\langle X, Y, I\rangle$. A finite ordered set can be represented by a Hasse diagram in which each element of the set is a node in a hierarchical structure. If an element is a subconcept of another, it is written lower in the diagram. Edges link immediate neighbors. For any pair of concepts c_1, c_2 in $\langle X, Y, I\rangle$, we have $c_1 \leq c_2$ if c_2 can be reached from c_1 by an ascending path.

Figure 2.6 shows the hierarchical representation of the context lattice from Table 2.4 as a Hasse diagram. Each node is annotated by its concept.

However, this notation is redundant. It is not necessary to repeat on higher nodes objects that have been defined by lower concepts, as they can be deduced from the hierarchical structure. Symmetrically, it is not necessary to repeat on lower nodes attributes that have been defined by higher concepts. The reduced notation only writes objects and attributes in the structure on those concepts which define them. Figure 2.7 shows the same lattice as Figure 2.6, in reduced notation. Concept lattices written in reduced notation can be read as monotonous multiple inheritance hierarchies. The resulting hierarchy is unique. It is entirely deduced from the context table and there are no possible alternative structures which fit with the above definitions.

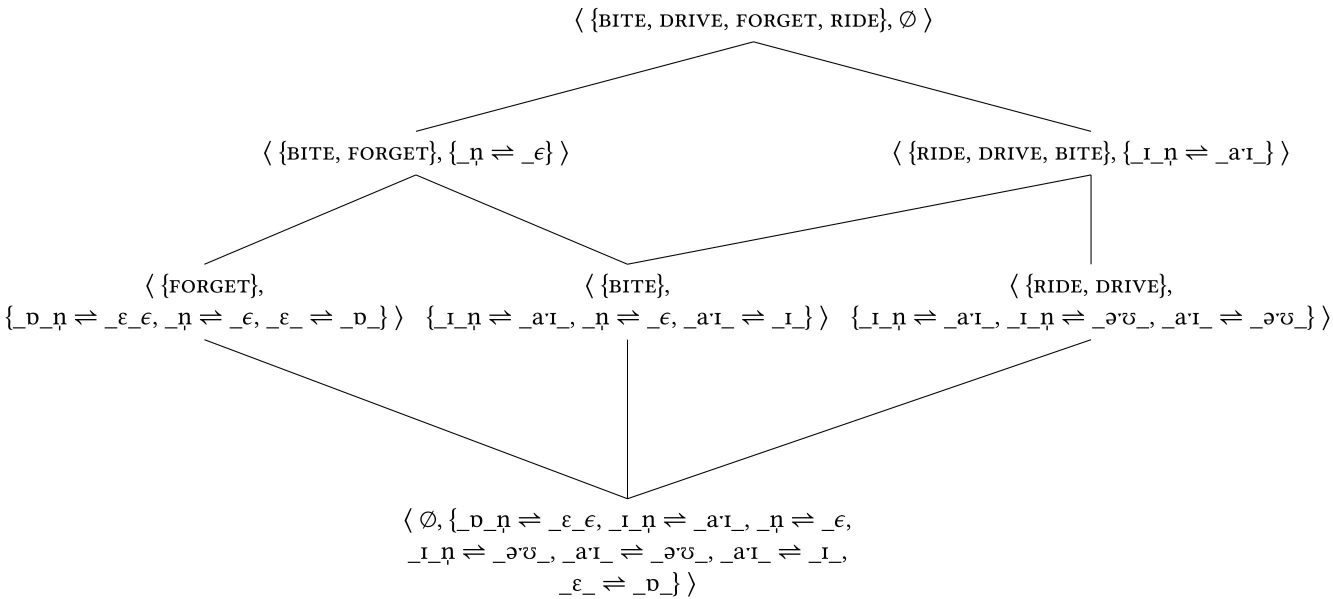

Figure 2.6: Concept lattice for the context in Figure 2.4

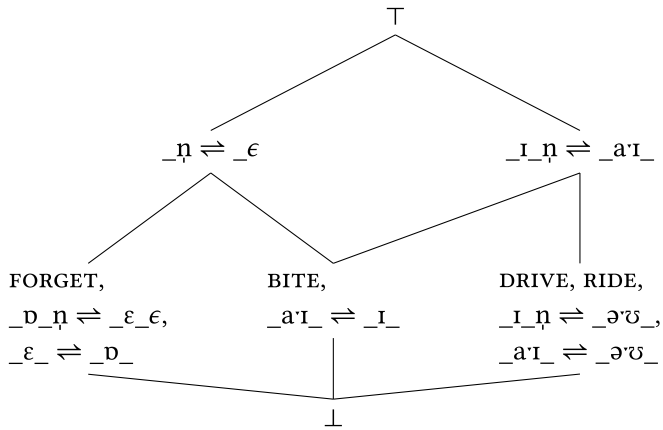

Figure 2.7: Concept lattice for the context in Figure 2.4, reduced notation

5 Properties of inflection class lattices

In this section, I apply the methodology described in the previous section to a few inflectional systems and investigate the similarity structure across their paradigms. I build IC lattices for the verbal systems of Modern Standard Arabic, English, French, European Portuguese and Zenzontepec Chatino, and for the nominal system of Russian. These languages are chosen for their variety and the availability of the computational resources needed for a quantitative investigation. The selection does not constitute a typologically representative sample, but it illustrates a variety of inflectional strategies.

For a description of the input datasets, see the appendix. As a first step before inferring IC lattices, I compute alternation patterns between all pairs of cells automatically from surface forms using the Qumín software (Beniamine 2017; 2018).

Russian declensions have been described as the conjunction of two separate systems: one affixal and one made of stress alternations (Brown & Hippisley 2012). Similarly, Campbell (2016) described Zenzontepec Chatino inflection as consisting of "two orthogonal layers, the prefixal system and the tone alternation system, simultaneously at play". Because alternation patterns describe change in a holistic way, inferring alternation patterns on whole forms in these datasets leads to a multitude of rare patterns which represent the many possible intersections of two more general phenomena, one on each dimension. As a solution, I divided the datasets into two parts, then joined the two resulting tables before inferring the classifications. For Russian, I created one table containing solely phonological segments and one containing solely stress information. For Zenzontepec Chatino, I created separate segmental and tonal tables. Ideally such decisions would be made automatically, but this enterprise is left to future work. For more discussion on the subject, see Beniamine (2018).

I define microclasses as the partition of lexemes which instantiate exactly the same alternation patterns for all pairs of cells: these are identical rows in the alternation pattern table. I keep only one entry representative of each microclass, which I call the EXEMPLAR lexeme. The choice of the exemplar is arbitrary. To build inflectional context tables, I take objects to be microclass exemplars and attributes to be combinations of a pair of cell and alternation pattern. The resulting contexts are very large. I use the python library CONCEPTS (Bank 2016) to generate all concepts from the context table and order them by specificity.

I obtain very large lattices. As an example, Figure 2.8 shows the overall structure of French and English lattices. Objects are labelled on the structure next to the concept which defines them. For legibility purposes, alternation patterns are not labelled. These examples are typical of the situation for all observed languages: the structures are by far too large for manual exploration and multiple inheritance is pervasive.

This fact in itself invalidates the hypothesis according to which real inflectional systems could be appropriately described as either partitions or trees. Computing the whole similarity structure now allows us to quantify precisely how far from the canon these systems fall. I operationalize three measures described in Section 3:

Number of concepts: in the canonical situation, if a lattice has b leaves, there are exactly $b+1$ concepts in the system (ignoring the infimum), the only other concept being the supremum. The higher the number of concepts, the more an inflectional system violates criterion 1 (distinctivity).

Depth of the hierarchy: In the canonical situation, the longest path (and in fact, all paths) from the root to a leaf passes through only one edge. Evaluating the depth of the hierarchy gives us information regarding the type of sharing between classes. A deep hierarchy is organized in successive classes and subclasses. Because concepts imply their ancestors, a deep hierarchy has more implicative structure than a shallower one. The deeper the hierarchy, the more it violates criterion 4 (flat implicative structure).

Mean degree: A canonical IC hierarchy is a one-level tree. A multi-level tree is a minor deviation from the canon. In a tree, the mean in-degree is 1 (ignoring the root, which has no incoming edges). Mean degree indicates the amount of multiple inheritance in the hierarchy. The higher the mean degree, the more the structure violates criterion 1 through heteroclite sharing.

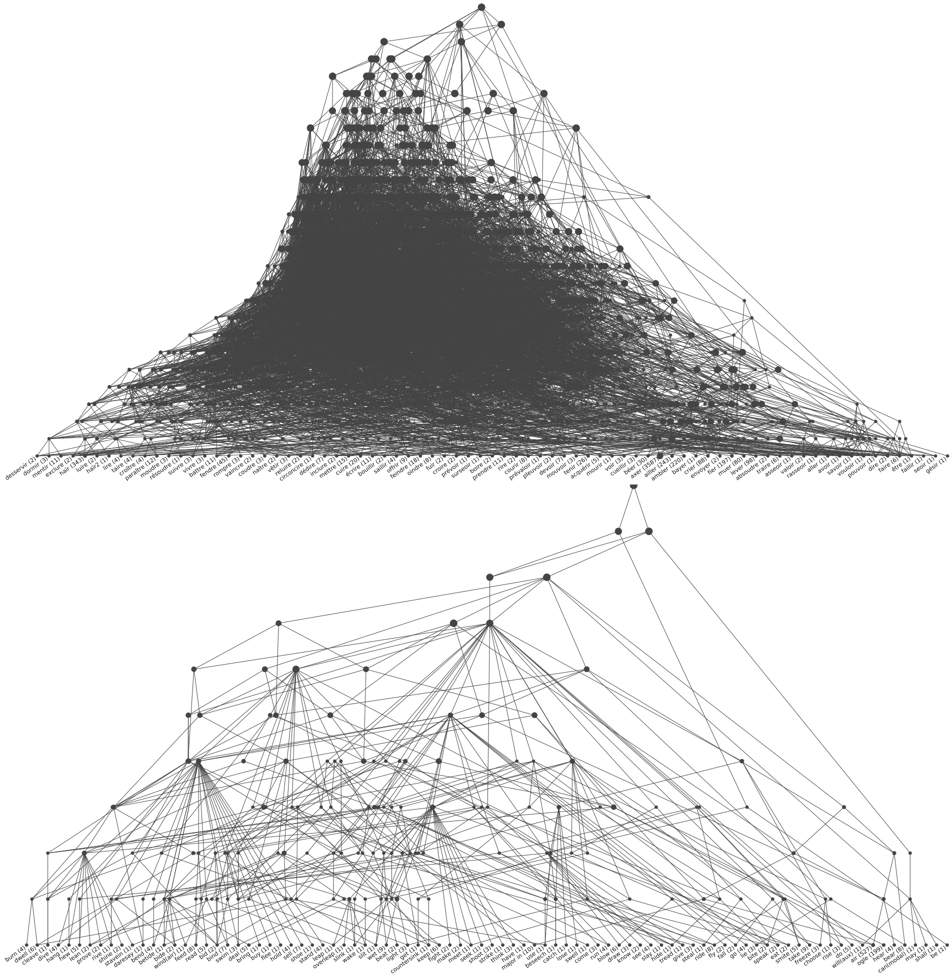

Figure 2.8: Inflection class lattices for French (top) and English (bottom) verbs.

Table 2.5 shows these measures for each system, as well as the number of lexemes in the dataset and the number of microclasses based on inflectional patterns. It is notable that the number of concepts found in each dataset is often comparable to the number of lexemes. In modern standard Arabic, there are 10 times more concepts than lexemes and in Russian, there are 35 times more concepts than lexemes. In French and Zenzontepec Chatino, the number of concepts and lexemes are of the same order. In English and European Portuguese, there are fewer concepts than lexemes, though the number of concepts is still high. This shows an important deviation from the conception according to which ICs provide a summary of inflectional behaviors.

Table 2.5: Canonicity measures of inflection class lattices based on alternation patterns

	Lexemes	Microclasses	Leaves	Depth	Degree	Concepts
MSA[a]	1018	367	302	33	3.65	10125
English	6064	118	88	11	1.91	244
French	5249	97	77	27	3.96	4845
Russian	1529	226	208	26	5.19	53858
EP[b]	1996	60	60	21	2.79	677
ZC[c]	324	99	98	8	2.65	524

[a]Modern Standard Arabic
[b]European Portuguese
[c]Zenzontepec Chatino

The mean in-degree in all systems is close to or higher than 2, indicating that heteroclisis is the general case. Depth and number of concepts are always much higher than in the canonical situation, although it is difficult to compare these raw numbers from one dataset to another, given that the number of leaves varies.

To be able to compare these values across datasets, I calculate a relative depth and a relative number of concepts (or density). Given a lattice with b leaves and a depth of h, I normalize this depth by the maximal possible depth over b leaves, which is $b - 1$ (ignoring the infimum):

$$\text{relative depth}(\mathcal{B}\langle X, Y, I\rangle) = \frac{h}{b-1}$$

The maximal depth $b - 1$ corresponds to the least possible canonical situation, where the lattice is the power set over the b leaves. In that case, there are

$n = 2^b - 1$ concepts. I thus normalize the number of concepts in the lattice by this maximal value, and I call the resulting measure DENSITY. If a lattice $\mathcal{B}\langle X, Y, I\rangle$ has n concepts over b leaves, then its density is:

$$\text{density}(\mathcal{B}\langle X, Y, I\rangle) = \frac{n}{2^b - 1}$$

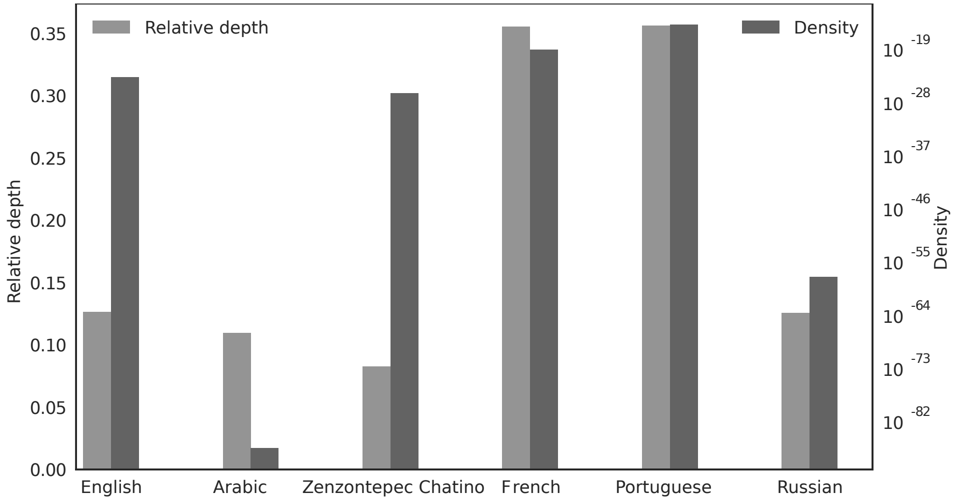

Figure 2.9: Relative canonicity measures on alternation pattern lattices

Figure 2.9 shows these values for each system. The growth of 2^b is such that compared to the maximum non-canonicity conceivable, our lattices have very few nodes, resulting in very low densities (all below 10^{-10}), even when the absolute number of nodes is high. The differences in density in Figure 2.9 are very small (they are shown on a log scale to make them perceptible) and depend mainly on the number of leaves. There is more variation in relative depth. In Zenzontepec Chatino, Modern Standard Arabic, Russian and English, relative depth is lower than 0.15, while European Portuguese and French verbal systems have densities around 0.35, indicative of a more hierarchical system. It is interesting to note that the absolute depth in Russian, French and European Portuguese is similar, but results in a higher density for Portuguese and French because they have fewer than 100 microclasses, while Russian counts over 200. It appears that the French and European Portuguese verbal systems, both Romance languages, would be especially poorly accounted for by a partition, despite a tradition of doing so in Romance linguistics.

Globally, these results show that the resulting classifications are visually very complex and far from the canon. This allows us to reject without hesitation the hypothesis according to which either partitions or tree structures would be appropriate models of ICs. However, these systems are also orders of magnitude less complex than the theoretical maximum.

6 Conclusion

In this chapter, I argued that while "inflection classes" usually refers to either partitions or trees, the similarity structure of inflectional systems is usually more complex and should rather be modeled as a lattice. Following the intuition according to which ICs are sets of lexemes distinguished by common inflectional properties, I put forward that any such maximal set is a relevant IC. FCA allows us to build automatically the ordered set of all these classes, or *concepts*, from paradigms of alternation patterns inferred over a large lexicon.

Using this methodology, I investigated the verbal systems of Modern Standard Arabic, English, French, European Portuguese and Zenzontepec Chatino, as well as the nominal system of Russian. I find that in all cases, the similarity structure between inflectional paradigms is undoubtedly hierarchical and that heteroclisis (multiple inheritance) is pervasive. These facts hold strongly even in systems like English which are usually seen as having a trivial inflectional structure.

The resulting classifications are much larger than what is suggested by traditional accounts and far too large for manual analysis. Usually, ICs are taken to be convenient summaries of an inflectional system. Our investigation shows that this is not the case when taking into account the entire IC structure: the number of concepts is often of the same order, if not higher, as the size of the lexicon. While one can always choose a small subset of classes for pedagogical or constructive purposes, there is no prominent such subset in the hierarchies. This can certainly explain why there are so many alternative analyses of known inflectional systems into partitions of ICs.

I defined precise quantitative measures of inflectional canonicity, taking partitions and trees as two degrees of inflectional canonicity. I showed that while the systems are much larger than they would be in the canonical situation, they are much closer to that ideal than they are to the theoretical maximum. This indicates that these systems are certainly not arbitrarily complex. This finding goes along with known observations that inflectional complexity, while surprisingly high in appearance, is usually bounded (Carstairs 1987; Carstairs-McCarthy 1991; Ackerman et al. 2009; Ackerman & Malouf 2015).

In conclusion, this study highlights the fact that the distribution of inflectional behaviors in a realistic lexicon is both highly structured and much more intricate than hand-crafted descriptions suggest.

Appendix

To compute IC lattices, I take as input paradigm tables of full, non segmented, raw forms in phonemic notation. The algorithm I use to infer alternation patterns (Beniamine 2018; 2017) also requires a decomposition of each phoneme into distinctive features. These serve as a basis to weight phoneme similarity in order to find linguistically sound alternations. They are also used to choose alternation patterns which lead to better generalizations over the whole lexicon. Unless specified otherwise, the definition of these features was based on Hayes (2012). The datasets and their constitution are described in more detail in Beniamine (2018).

Arabic is a Semitic language. Modern Standard Arabic is the standardized variety of Arabic used in writing in Arabic speaking countries. The lexicon was extracted and normalized from Wiktionary entries as part of the UNIMORPH project (Kirov et al. 2016). The UNIMORPH lexicon provides orthographic forms. I transcribed them phonemically in a semi-automatical way (for more details, see Beniamine 2018). The resulting lexicon counts 1,018 lexemes, inflected for 109 possible combinations of mode, tense, voice, gender, person and number.

English is a West Germanic language spoken primarily in the United Kingdom, the United States, Australia, Canada and globally as a lingua franca. Our lexicon is a subset of the CELEX2 database (Baayen et al. 1995). The original SAMPA notations were transcribed into IPA automatically (Beniamine 2018). The original lexicon often includes unlabelled regional variants, which leads to paradigms where overabundance (more than one form for a given lexeme and paradigm cell) is frequent. Most verbs are inflected for five paradigm cells: present third person, other present forms, past participle, present participle, past. However, because of the verb *be*, which is overdifferentiated, I count eight paradigm cells: infinitive, present first person, present third person, present other persons, past participle, present participle, past first person, past third person, other past persons. The lexicon counts 6,064 verbal lexemes. Distinctive phonological features are based on Halle & Clements (1983) and Chomsky & Halle (1968).

French is a Romance language spoken primarily in France. French verbs are inflected for 51 paradigm cells, structured in seven finite tenses, each inflected for six persons, the imperative inflected for only two persons and six nonfinite cells. I use the verbal entries from the lexicon Flexique (Bonami et al. 2014), itself based on Lexique (New et al. 2001). Phonological features are based on Dell (1973). The lexicon counts 5,249 lexemes.

European Portuguese is a Romance language spoken in Portugal. Our lexicon is based on frequent verbs from Veiga et al.'s (2013) pronunciation dictionary. It counts 1,996 lexemes inflected for 69 combinations of mood, tense and person. Phonological features originate from Bonami & Luís (2014).

Russian is an East Slavic language spoken in Russia and neighboring countries. Beniamine & Brown's (2019) lexicon was generated by a DATR fragment Brown & Hippisley 2012 as romanized forms. The forms were then transcribed phonemically semi-automatically (Beniamine 2018). The nominal paradigm of Russian counts six combinations of case and number. A small number of lexemes are also inflected for second singular locative (see Brown 2007). The lexicon counts 1,529 lexemes.

Zenzontepec Chatino is a Chatino language of the Zapotecan branch of Oto-Manguean, spoken in Oaxaca, Mexico. The dataset I use comes from Surrey's Oto-Manguean inflectional class database (Feist & Palancar 2015) and is based on data provided by Eric Campbell. Explicit low tones were added automatically in the dataset (Beniamine 2018). Zenzontepec Chatino verbs are inflected for only four paradigm cells, with aspect/mood values: completive, potential, habitual and progressive. The dataset counts 324 lexemes.

Abbreviations

ACC	accusative
COMP	completive
DAT	dative
FCA	formal concept analysis
GEN	genitive
HAB	habitual aspect
IC	inflection class
INS	instrumental
LOC	locative
NOM	nominative
PL	plural
POT	potential
PROG	progressive
PRS	present
PST.PART	past participe
PST	past
SG	singular

Acknowledgements

This work was supported by the operation Morph 1 of the axis 2 of the Laboratory of Excellence "Empirical Foundations of Linguistics" (EFL). I thank Olivier Bonami for the many conversations which led to the idea of this paper, and his valuable feedback and advice. I am very grateful to the two anonymous reviewers for their comments and suggestions, and to the proofreader and the copy-editor for their help. All remaining errors are mine.

References

Ackerman, Farrell, James P. Blevins & Robert Malouf. 2009. Parts and wholes: Implicative patterns in inflectional paradigms. In James P. Blevins & Juliette Blevins (eds.), *Analogy in grammar*, 54–82. Oxford: Oxford University Press.

Ackerman, Farrell & Robert Malouf. 2015. The No Blur principle effects as an emergent property of language systems. In *Proceedings of the Annual Meeting of the Berkeley Linguistics Society*, vol. 41. DOI: 10.20354/B4414110014.

Aronoff, Mark. 1994. *Morphology by itself*. Cambridge: MIT Press.

Arrivé, Michel (ed.). 2012. *Bescherelle: la conjugaison pour tous*. Nouvelle édition. Paris: Hatier.

Baayen, R. Harald, R. Piepenbrock & L. Gulikers. 1995. *CELEX2 LDC96L14*. Philadelphia: Linguistic Data Consortium.

Bank, Sebastian. 2016. *Concepts: Formal concept analysis with Python*. https://pypi.org/project/concepts/ (23 February, 2021).

Bělohlávek, Radim. 2009. Introduction to formal concept analysis. Olomouc. https://phoenix.inf.upol.cz/esf/ucebni/formal.pdf (23 February, 2021).

Beniamine, Sacha. 2017. Un algorithme universel pour l'abstraction automatique d'alternances morphophonologiques. In *Actes de TALN 2017*, 77–85.

Beniamine, Sacha. 2018. *Classifications flexionnelles: étude quantitative des structures de paradigmes*. Université Sorbonne Paris Cité - Paris Diderot. (Doctoral dissertation).

Beniamine, Sacha, Olivier Bonami & Benoît Sagot. 2017. Inferring inflection classes with description length. *Journal of Language Modelling* 5(3). DOI: 10.15398/jlm.v5i3.184.

Beniamine, Sacha & Dunstan Brown. 2019. *Inflected lexicon of Russian nouns in IPA notation*. Online repository. DOI: 10.5281/zenodo.3428591. https://gitlab.com/sbeniamine/russiannounlexicon (23 February, 2021).

Blevins, James P. 2004. Inflection classes and economy. In L. Gunkel, G. Müller & G. Zifonun (eds.), *Explorations in nominal inflection*, 41–85. Berlin, Boston: De Gruyter. DOI: 10.1515/9783110197501.51.

Blevins, James P. 2006. Word-based morphology. *Journal of Linguistics* 42(3). 531–573. DOI: 10.1017/S0022226706004191.

Bonami, Olivier. 2014. La structure fine des paradigmes de flexion. Mémoire d'habilitation U. Paris Diderot.

Bonami, Olivier & Sacha Beniamine. 2016. Joint predictiveness in inflectional paradigms. *Word Structure* 9(2). 156–182. DOI: 10.3366/word.2016.0092.

Bonami, Olivier, Gauthier Caron & Clément Plancq. 2014. Construction d'un lexique flexionnel phonétisé libre du français. In Franck Neveu, Peter Blumenthal, Linda Hriba, Annette Gerstenberg, Judith Meinschaefer & Sophie Prévost (eds.), *Actes du quatrième congrès mondial de linguistique française*, 2583–2596.

Bonami, Olivier & Ana R. Luís. 2014. Sur la morphologie implicative dans la conjugaison du portugais: une étude quantitative. In Jean-Léonard Léonard (ed.), *Morphologie flexionnelle et dialectologie romane. Typologie(s) et modélisation(s).* (Mémoires de la Société de Linguistique de Paris 22), 111–151. Leuven: Peeters.

Brown, Dunstan. 1998. *From the general to the exceptional.* University of Surrey. (Doctoral dissertation).

Brown, Dunstan. 2007. Peripheral functions and overdifferentiation: The Russian second locative. *Russian Linguistics* 31(1). 61–76. DOI: 10.1007/S11185-006-0715-5.

Brown, Dunstan & Andrew Hippisley. 2012. *Network Morphology: A defaults-based theory of word structure* (Cambridge Studies in Linguistics). Cambridge University Press. DOI: 10.1017/CBO9780511794346.

Campbell, Eric. 2011. Zenzontepec Chatino aspect morphology and Zapotecan verb classes. *International Journal of American Linguistics* 77. 219–246.

Campbell, Eric. 2016. Tone and inflection in Zenzontepec Chatino. In Enrique L. Palancar & Jean-Léonard Léonard (eds.), *Tone and inflection*, 141–162. Berlin: Mouton de Gruyter.

Carstairs, Andrew. 1987. *Allomorphy in inflection.* London: Croom Helm.

Carstairs-McCarthy, Andrew. 1991. Inflection classes: Two questions with one answer. In F. Plank (ed.), *Paradigms: The economy of inflection* (Empirical Approaches to Language Typology [EALT]), 213–253. De Gruyter. DOI: 10.1515/9783110889109.213.

Carstairs-McCarthy, Andrew. 1994. Inflection classes, gender, and the principle of contrast. *Language* 70. 737–788.

Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English.* Harper & Row.

Corbett, Greville G. 1982. Gender in Russian: An account of gender specification and its relationship to declension. *Russian linguistics* 6. 197–232.

Corbett, Greville G. 2009. Canonical inflection classes. In Fabio Montermini, Gilles Boyé & Jesse L. Tseng (eds.), *Selected proceedings of the 6th Décembrettes: Morphology in Bordeaux*, 1–11. Somerville: Cascadilla Press.

Corbett, Greville G. & Norman M. Fraser. 1993. Network Morphology: a DATR account of Russian nominal inflection. *Journal of Linguistics* 29. 113–142. DOI: 10.1017/S0022226700000074.

Dell, François. 1973. *Les règles et les sons: Introduction à la phonologie générative* (Collection Savoir). Paris: Hermann.

Dressler, Wolfgang U., Marianne Kilani-Schoch, Natalia Gagarina, Lina Pestal & Markus Pöchtrager. 2008. On the typology of inflection class systems. *Folia Linguistica* 40(1-2). 51–74. DOI: 10.1515/flin.40.1-2.51.

Dressler, Wolfgang U. & Anna M. Thornton. 1996. Italian nominal inflection. *Wiener Linguistische Gazette* 55-57. 1–26.

Feist, Timothy & Enrique L. Palancar. 2015. *Oto-Manguean inflectional class database*. University of Surrey. DOI: 10.15126/SMG.28/1.

Flickinger, Dan. 1987. *Lexical rules in the hierarchical lexicon*. Stanford University. (Doctoral dissertation).

Frisch, Stefan. 1997. *Similarity and frequency in phonology*. Northwestern University. (Doctoral dissertation).

Ganter, Bernhard & Rudolf Wille. 1998. *Formal concept analysis: Mathematical foundations*. Springer. DOI: 10.1007/978-3-642-59830-2.

Ginzburg, Jonathan & Ivan A. Sag. 2000. *Interrogative investigations. The form, meaning, and use of English interrogatives*. Stanford: CSLI Publications.

Halle, Morris & George N. Clements. 1983. *Problem book in phonology: A workbook for introductory courses in linguistics and in modern phonology* (Bradford Books). Cambridge: MIT Press.

Hayes, Bruce. 2012. *Spreadsheet with segments and their feature values*. Distributed as part of course material for Linguistics 120A: Phonology I at UCLA. http://www.linguistics.ucla.edu/people/hayes/120a/index.htm (23 February, 2021).

Kaufman, Terrence. 1989. The phonology and morphology of Zapotec verbs.

Kilani-Schoch, Marianne & Wolfgang U. Dressler. 2005. *Morphologie naturelle et flexion du verbe français*. Tübingen: Gunter Narr Verlag.

Kirov, Christo, John Sylak-Glassman, Roger Que & David Yarowsky. 2016. Very-large scale parsing and normalization of Wiktionary morphological paradigms. In Nicoletta Calzolari (Conference Chair), Khalid Choukri, Thierry Declerck, Sara Goggi, Marko Grobelnik, Bente Maegaard, Joseph Mariani, Helene Mazo, Asuncion Moreno, Jan Odijk & Stelios Piperidis (eds.), *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)*. Portorož, Slovenia: European Language Resources Association (ELRA).

Lee, Jackson & John A. Goldsmith. 2013. Automatic morphological alignment and clustering. Presented at the 2nd American International Morphology Meeting.

Matthews, Peter H. 1991. *Morphology*. 2nd. Cambridge: Cambridge University Press.

New, B., C. Pallier, L. Ferrand & R. Matos. 2001. Une base de données lexicales du français contemporain sur internet: LEXIQUE. *L'Année Psychologique* 101. 447–462.

Plénat, Marc. 1987. Morphologie du passé simple et du passé composé des verbes de l' "autre" conjugaison. *ITL Review of Applied Linguistics* 77–78. 93–150.

Pollard, Carl & Ivan A. Sag. 1994. *Head-driven phrase structure grammar*. Stanford: CSLI Publications; The University of Chicago Press.

Spencer, Andrew. 2012. Identifying stems. *Word Structure* 5(1). 88–108. DOI: 10.3366/word.2012.0021.

Stump, Gregory T. & Raphael A. Finkel. 2013. *Morphological typology: From word to paradigm* (Cambridge Studies in Linguistics). Cambridge: Cambridge University Press. DOI: 10.1017/CBO9781139248860.

Veiga, Arlindo, Sara Candeias & Fernando Perdigão. 2013. Generating a pronunciation dictionary for European Portuguese using a joint-sequence model with embedded stress assignment. *Journal of the Brazilian Computer Society* 19(2). 127–134. DOI: 10.1007/s13173-012-0088-0.

Walther, Géraldine & Benoît Sagot. 2011. Modélisation et implémentation de phénomènes flexionnels non-canoniques. *Traitement Automatique des Langues* 52(2). 91–122.