PART II

PARADIGMS AND THEIR VARIANTS

CHAPTER 5

INFLECTIONAL PARADIGMS

JAMES P. BLEVINS

5.1 Introduction

The status of paradigms varies widely across approaches, ranging from the Post-Bloomfieldian view that they are 'epiphenomena', to classical word and paradigm (WP) conceptions of paradigms as fundamental units of lexical organization. This variation rests in turn on different views of the structure of morphological systems. Within the Post-Bloomfieldian tradition, paradigms represent collections of forms which are based on a common root and/or which have a partially overlapping derivational history. In line with the rigidly syntagmatic perspective of these accounts, it is the relation between formatives and words (or larger units) that determines the structure of a morphological system. The sets defined by the inflected (or derivationally related) forms of an item have no grammatical status or coherence as 'units'. At the other extreme are classical WP approaches, which factor grammatical systems into sets of exemplary paradigms and inventories of principal parts. Between these extremes lies a range of approaches that incorporate a notion of 'paradigm' that corresponds to sets of forms, or to sets of abstract 'cells', which are usually defined in terms of feature bundles.

This chapter sets out some of the dominant conceptions of paradigms that have been developed in different morphological traditions. Section 5.2 identifies the kinds of elements that have been organized into paradigms, and contrasts alternative notions of paradigmatic structure. Section 5.3 summarizes types of relations that have been argued to require reference to a notion of paradigm. Section 5.4 reviews recent applications of information theory to model the implicational structure of paradigms. Section 5.5 concludes by considering how conceptions of paradigms reflect general assumptions about inflectional systems.

5.2 The Organization of Paradigms

In parallel to ontological questions about the types of units that make up paradigms, there are issues concerning WHICH units of whatever kind are grouped together into paradigms and about the organizational principles that bind these units into a structure. Inflectional paradigms are the core case of morphological paradigms, though some accounts develop a notion of derivational paradigm or 'morphological family' (de Jong 2002) that groups derivationally-related forms or items.

Within the class of inflectional paradigms, there are different sets of cells or forms that can be identified as constituting sub-paradigms. There is relatively little ambiguity in most declensional systems, since the cells or forms of a nominal item can usually be organized into a single matrix of elements that differ for the features (case, number, gender, definiteness, etc.) that are distinctive in a given language. Yet even nominal paradigms may exhibit splits of various kinds. The paradigm of an item can be organized into sub-paradigms, whose members are more closely associated with each other than with members of other sub-paradigms. German provides a simple case of this type of split, as discussed in Section 5.3.2. Singulars and plurals each form sub-paradigms whose members are more similar to each other than they are to members of the other sub-paradigm. In languages with a dual, such as Slovene, the dual and plural forms likewise tend to show a stronger affinity to each other than to singulars. In declensional systems with large case inventories, such as Finnish or Estonian, it is common for there to be a split between the core 'grammatical' cases and the set of 'local' or 'semantic' cases. In a language such as Tundra Nenets, nominal paradigms may be organized into multiple clusters, each based on a common stem (Ackerman *et al.* 2009). The even larger form inventories found in Daghestanian languages are likewise organized into multiple 'series' of local cases (Kibrik 1998). However, there is some debate about the precise number of case distinctions in these languages Comrie and Polinsky (1998), and some uncertainty about the combinatorics of spatial series markers and case endings.

In each of these cases, the traditional paradigm represents a unit of organization between the forms in a sub-paradigm and the full morphological family of an item. Conjugational systems add a further level, since the forms that define a verbal lexeme may be organized into multiple paradigms, defined by different tense, aspect mood, and voice features.[1] The fact that forms tend to be more strongly associated within than across sub-paradigms underlies the classical WP practice of describing paradigms in terms of multiple principal parts.

[1] There is a fairly robust consensus for organizing conjugational paradigms into subsets with common tense/aspect/mood/voice features and variable agreement features, as opposed to an inverse organization into sets of forms that share common agreement features and have contrasting tense/aspect/mood/voice features. Hence it is more customary to talk of 'the paradigm' of a verb or 'the present indicative (sub-)paradigm' than of 'the first person singular paradigm'.

In sum, a form in a given language may be assigned to successively larger collections, from sub-paradigms through to morphological families. The familiarity of the inflectional paradigm, in the sense of the complete set of cells or forms of an item, is due in large part to the relevance of this notion to grammatical descriptions and the usefulness of the notion for pedagogical purposes.

The structure associated with paradigms tends to depend fairly directly on assumptions about the types of elements that are grouped together into paradigms. Approaches that treat paradigms as sets of forms often define this set in terms of derivation from common bases. The model of A-Morphous Morphology (Anderson 1992) is typical in defining the paradigm of an item as the 'set of surface word forms that can be projected from the members of its stem set' (1992: 134), where the stem set contains the root of an item and any stem allomorphs. Approaches that treat cells as abstract feature bundles regard paradigms in turn as n-dimensional matrices, in which each dimension is defined by a distinctive feature. Within models such as Paradigm Function Morphology (PFM) (Stump 2001b), this abstract paradigm space is defined by the feature distribution constraints of a language, independent of the forms that come to be associated with cells by 'spell-out' rules.[2] The fact that A-Morphous Morphology and Paradigm Function Morphology are both realizational models shows that the precise characterization of paradigms is something of a free choice in this class of models.

Assumptions about the organization of paradigms are more deeply embedded in the analytic practices of classical WP models. This tradition proceeds from 'the general insight that one inflection tends to predict another' (Matthews 1991: 197), with the primary locus of prediction identified as 'words as wholes, arranged according to grammatical categories … distinguished by their endings' (1991: 187).[3] Paradigms consist of cell-form pairs that are related not by shared bases or derivational histories but networks of interdependencies. In some cases, sets of interpredictable cells may be morphosyntactically coherent. However, they need not be, and in the patterns that Matthews (1972, 1991) terms 'parasitic' and Aronoff (1994) calls 'morphomic', the features of interdependent cells may serve solely to pick out pairs of forms that alternate or covary in some way.

The classic WP factorization of inflectional systems into principal parts and exemplary paradigms embodies a strong hypothesis about the interdependency of forms in a paradigm. A language whose paradigms consisted entirely of mutually independent forms could not be broken down in this way, since no set of forms smaller than a full paradigm would identify the class of an item. This holds independently of the complexity of a system, as one can see by considering a system with just two cells, each associated with two distinct patterns of exponence. If each pattern of exponence that

[2] Later elaborations of PFM introduce a split between 'form paradigms' and 'content paradigms', in part to describe patterns of heteroclisis. See Stump (2006a).

[3] Although Matthews (1991) here specifies 'endings' in the context of a discussion of Latin, predictive relations are understood to apply to any systematic patterns of form variation.

realizes the first cell can co-occur with both of the patterns that realize the second cell, the system will contain a total of four patterns. There is no difficulty in exemplifying paradigms, associating them with classes, or in assigning individual items to classes. However, this classification is purely taxonomic, and the paradigms and classes it defines cannot play the role that exemplary paradigms do in a classic WP analysis. Since the forms of an item are inflectionally independent, the system does not define principal parts, and each item must be represented by both forms. If a system has weakly interdependent forms, then a factorization into exemplary paradigms and principal parts may also achieve negligible economy and merely distribute patterns across a large and unilluminating collection of paradigms that 'multiply out' the possibilities of co-occurrence. Hence patterns of mutual implication in a system must be relatively tight in order for a classic WP factorization to simplify rather than complicate the description of a system.

There is an intimate connection between the assumptions that an approach adopts about the elements in a paradigm and the organization imposed on those elements, and the usefulness of the resulting structure. Sets of simple forms are relatively uninformative; the members of such form sets can be organized in terms of their morphotactics or their derivational history, but shared morphotactic or derivational structure rarely play any direct grammatical role. So it is unsurprising that little if any grammatical significance is attached to paradigms in models that treat them merely as sets of simple forms. Sets of abstract cells intrinsically define a much richer structure. It is within this essentially closed and uniform space that notions like suppletion, syncretism, and paradigm gap are defined most clearly. The features that define this space likewise provide the vocabulary for the markedness proposals of Jakobson (1932, 1936), and the property co-occurrence restrictions and other types of constraints developed in approaches such as Stump (2001b). However, pairings of forms and abstract cells define the most informative elements, and determine the richest structure, since relations may refer both to the features and to the form of a cell.

Implicational relations highlight the contrast between these conceptions of paradigms. Words or sub-word units are of limited predictive value *qua* forms. Knowing that the paradigm of an item contains a particular word form is often relatively uninformative unless one knows what cell the form realizes and, consequently, one knows where the form fits into the paradigm. Abstract cells support predictions based on markedness or other types of relations between unrealized features. Pairings of forms and abstract cells are again of the greatest predictive value, as they sanction deductions between forms and between cells but also deductions between forms based on their place in a paradigm. In setting out the evidence for treating paradigms as central components of the inflectional system of a language, this chapter will attempt to present the strongest case for a classic WP perspective by adopting the most informative conception of paradigms, one in which they are structured sets of form–cell pairs.

As in the case of words, the role of paradigms in pedagogical and reference grammars attests to their practical usefulness. The pedagogical relevance of paradigms reflects their descriptive value, at least for languages of the 'flectional' type (Sapir 1921), as Matthews (1991) emphasizes.

> Pupils begin by memorising paradigms. These are sets of words as wholes, arranged according to grammatical categories. This is not only traditional, it is also effective. They learn that different members of a paradigm are distinguished by their endings.... They can then transfer these endings to other lexemes, whose paradigms they have not memorised.... It seems unlikely that, if a structuralist method or a method derived from structuralism were employed instead, pupils learning Ancient Greek or Latin—or, for that matter, Russian, Modern Greek or Italian—would be served nearly so well. (Matthews 1991: 187f.)

Comprehensive descriptions of languages with intricate morphological systems also tend to exhibit exemplary paradigms and patterns at different levels of generality. No practical alternative has yet been devised that would replace exemplary patterns by something like morpheme inventories related to surface realizations by rules or constraints. In part, this reflects the fact that Post-Bloomfieldian models were not designed with practical or comprehensive description in mind, so that the output of their analyses may range from nontransparent to indeterminate (Karttunen 2006). This tradition is accordingly sceptical about the relevance of the pedagogical or descriptive utility of paradigms.

It is of course true that the notion of paradigms represented in pedagogical and reference grammars is strongly influenced by the uses to which paradigms have been put in these sources. It is usually assumed that inflectional systems can be factored into a discrete number of inflection classes, each represented by the full paradigm of some exemplary member of the class. Principal part inventories are normally 'static' in the sense of Finkel and Stump (2007), in that each non-exemplary item of a given class is represented by the same forms or sets of forms, for example the nominative singular or first person singular. The deduction of new forms from exemplary paradigms and principal parts is likewise attributed to symbolic processes of the kind expressed by proportional analogies.

Each of these assumptions creates problems that remain largely unaddressed, and mostly unacknowledged, in the descriptive and pedagogical sources that make use of paradigms. The difficulty of motivating the choice of principal parts (or 'leading forms') is perhaps the best known of these problems.

> One objection to the Priscianic model ... was that the choice of leading form was inherently arbitrary: the theory creates a problem which it is then unable, or only partly able, to resolve. (Matthews 1972: 74)

The other assumptions raise similar problems. Although pedagogical traditions tend to agree broadly on the number of classes in a language (and even on the choice of exemplary members of each class), this consensus often reflects established practices or considerations of utility. There are simply no generally agreed criteria for class identification, let alone for selecting exemplary items.

The optimal solution to these problems avoids them altogether by recognizing that they derive from practical idealizations rather than core properties of paradigms or paradigm-based models. For pedagogical purposes, it is useful to draw the most informative cells of a paradigm to the attention of language learners. However, there is no reason to assume that a single form will always identify the inflectional pattern of an item.[4] Conversely, there is no reason to ignore the partial information supplied by other forms. There are also no grounds for assuming that a language can be organized into some fixed set of classes, independent of the purposes or level of detail of the classification. Instead, different sets of interdependent forms may be defined based on their predictive value. For pedagogical purposes, it will often be useful to bundle these sets of interdependent forms into larger collections that specify the shape of each form of a single item, irrespective of how loosely the larger forms of different sets are connected to each other. In most languages, it is of course likely that there is at least some member of each class whose forms are frequent enough that the full paradigm of that item could serve as an exemplar for the class. However, the contrast between forms that are taken to be resident in the mental lexicon of a speaker and those that must be deduced from previously-encountered forms will depend on frequency, and it is known that the different forms of an item may have very different frequencies of occurrence. There is also no principled reason to assume that the analogical processes that deduce new forms of an item should be representable symbolically, rather than sub-symbolically.

What remains after these idealizations have been stripped away is a conception of paradigms as structured networks of interdependent elements. Although it is possible to generalize over these networks, traditional paradigm-based models incorporate a distinctively exemplar-based conception. The leading idea is that form variation within a morphological system is exhibited by exemplary patterns that serve a dual role within a grammatical system. On the one hand, the paradigms resident in a speaker's mental lexicon (whether full or partial) specify forms of particular lexical items. On the other hand, these structures provide a model for the inflection of as yet unencountered forms of other items. The pedagogical practice of identifying a single exemplary paradigm for each inflection class represents an idealization over the sets of resident forms that collectively reinforce the patterns exhibited by the exemplary paradigms.

5.3 Paradigmatic Relations

The value of this classical WP notion of paradigm depends on the types of generalizations and insights that it supports. Within the linguistic literature, a number of

[4] Languages in which regular items can be represented by a single diagnostic form will of course conform to a pedagogical ideal of inflectional economy. A similar conception underlies more theoretical principles, such as the Single Base Hypothesis of Albright (2002).

principles and constraints have been proposed that operate in a distinctive way within paradigms. Although not always directly formulated in terms of paradigms, these constraints exploit the fact that members of a paradigm tend to have numerous properties in common, including 'inherent' grammatical features, a core lexical meaning, and a shared stem, or stem set. These properties put the elements of a paradigm in close 'competition' for syntagmatic slots, which may give rise to grammatical reflexes. The most familiar reflex is illustrated by cases in which competition between alternatives is taken to be resolved in favour of the most specific alternative. This interaction is expressed by a range of principles, including Morphological Blocking (Aronoff 1976), the Elsewhere Condition (Kiparsky 1982b), Pāñini's Principle (Janda and Joseph 1992; Stump 2001b), and the Disjunctive Ordering Principle (Anderson 1986).

Another reflex of competition between forms of an item is a type of 'paradigmatic deduction' that allows a form to be interpreted based on the absence of a marker. Noun paradigms in English provide a simple example. Plural nouns in English are marked by the suffix /z/ (represented orthographically as *-s*). There is, however, no marking of singular number, and none is needed, given that a singular noun is unambiguously identified by the lack of a plural marker. A zero morph adds no information to what speakers can already deduce from the absence of any realized exponent.[5] Similar patterns are even more typical of more intricate paradigms. As Anderson (1992) notes, Georgian verbal paradigms provide a striking illustration of the fact that 'information may sometimes be conveyed not by constituents that are present in the structure of a given word, but precisely by the fact that certain other material is ABSENT'.

> Consider the Georgian Verb form [*dagxat'e*] in [Table 5.1], for example, ... This form represents agreement with a first-person singular Subject and a second-person singular Direct Object. ... But while an overt affix (/g/) is present to signal agreement with the second-person Object, no affix marks the fact that the Subject of this Verb is (and must necessarily be) first-person singular. This agreement can be inferred from the following information. The Subject cannot be second person, because if it were, the sentence would be reflexive—but reflexive forms in Georgian are grammatically third person, and this Verb has a second-person Object. Similarly, the Subject cannot be third person, since, if it were, there would be a suffix [/-a/] at the end of the Verb. Thus ... the Subject must be first person. But it must be singular, rather than plural, since a first-person plural Subject would trigger the introduction of a suffix /-t/) at the end of the Verb. We know therefore that the Subject of this Verb must be first-person singular, but this fact is not signaled by the presence of any overt affix in the word. (Anderson 1992: 87)

[5] Early Post-Bloomfieldian accounts introduced a 'zero' singular marker in such cases (Harris 1942: 110), and the practice survives in some contemporary work. Hockett (1947: 230) recognized that this type of element had 'a very dubious status', and the coherence of a 'zero marker' is questioned by Matthews (1991: 124). From a classic WP perspective, the appeal to 'zeros' can be seen as a means of compensating for the rigidly syntagmatic character of the post-Bloomfieldian model, which has no provision for any type of paradigmatic comparison.

Table 5.1 Aorist indicative paradigm of XAT'VA 'paint'

	Object				
Subject	1 Sg	1 Pl	2 Sg	2 Pl	3
1 Sg	–	–	dagxat'e	dagxat'et	davxat'e
1 Pl	–	–	dagxat'et	dagxat'et	davxat'et
2 Sg	damxat'e	dagvxat'e	–	–	daxat'e
2 Pl	damxat'et	dagvxat'et	–	–	daxat'et
3 Sg	damxat'a	dagvxat'a	dagxat'a	dagxat'at	daxat'a
3 Pl	damxat'es	dagvxat'es	dagxat'es	dagxat'es	daxat'es

Source: Tschenkéli (1958: §31).

This traditional deduction interprets an inflected form in the context of its paradigm. The deduction is, as Anderson shows, effective, provided that two conditions are met. First, the speaker must be able to identify the paradigm of *dagxat'e* in order to compare this form against relevant alternatives. Second, given that the deductions involve whole forms, not localized patterns of exponence, the speaker requires access to the full aorist indicative paradigm of XAT'VA—or sufficient information to reconstruct a paradigm of this class.[6]

5.3.1 Paradigm Economy

More direct evidence for paradigms takes the form of constraints or generalizations that apply to paradigms. In the diachronic domain, notions like 'levelling' and 'extension' are principally defined with reference to inflectional paradigms (Hock 1991). In the synchronic domain, various constraints have been proposed that exploit, directly or indirectly, the implicational structure that binds the cells of a paradigm into a cohesive whole. This structure underlies the organization and economy of classic WP models, as Matthews (1991) notes.

> The most general insight is one inflection tends to predict another.
> This insight can be incorporated into any model. Traditionally, it is the basis for the method of exemplary paradigms. If the alternations were independent, these would have to be numerous. (Matthews 1991: 197)

Matthews' observation that predictive relations prevent alternations from becoming too 'numerous' also gives the classic WP explanation for what have come to

[6] Interestingly, both conditions rely on a notion of paradigm that is a more coherent and cohesive structure than the simple set of output forms that Anderson (1992) proposes elsewhere.

be known as ‘paradigm economy’ effects. The contemporary interest in paradigm economy effects stems largely from the discussion in Carstairs (1983), which draws attention to these effects and attributes them to a dedicated Paradigm Economy Principle (PEP). Subsequent elaborations of this approach attempt to subsume the effects of the PEP under a range of more general principles (Nyman 1986; Carstairs-McCarthy 1991, 1994). However the general conception of economy that underlies these proposals is articulated most clearly in the initial formulation of the PEP. In contrast to a traditional model, the PEP is not concerned with distinctive inflectional patterns in general but with ‘inflectional resources’, by which Carstairs (1983) means ‘affixal resources’.

> Paradigm economy provides at least a partial answer to a question which, so far as I can discover, has not been asked before—a question about how, in any inflected language, the inflexional resources available in some word-class or part of speech are distributed among members of that word class. (Carstairs 1983: 116)

The distribution of inflectional resources is characterized in terms of upper and lower bounds on the space of possible inflection classes, where these bounds are determined by the number of different affixal strategies available for realizing paradigm cells in a given word class. This can be illustrated by a schematic example. Consider a simple declensional system with two contrastive properties, number and case, two distinct number features, N_1 and N_2 and four case features, C_1, C_2, C_3 and C_4. Since each combination of number and case features defines a paradigm cell, these number and case features define the family of eight-cell paradigms summarized in Table 5.2. The actual number of realizations for each cell are determined mainly by the size of the nominal lexicon, which is independent of the economy of the inflectional system. Hence to describe inflectional economy one must first isolate patterns of inflectional exponence. For each cell (N_i, C_j), let $\xi(N_i, C_j)$ represent the set of patterns that realize the cell and $|\xi(N_i, C_j)|$ the number of patterns in that set. There will be at least one inflectional pattern associated with every cell, since the existence of the pattern is a precondition for recognizing the cell in the first place. In every class system, there will also be a maximum number of patterns realized by any cell, though there need not be a unique ‘maximally allomorphic’ cell.

Table 5.2 Schematic paradigm structure

	N_1	N_2
C_1	(N_1,C_1)	(N_2,C_1)
C_2	(N_1,C_2)	(N_2,C_2)
C_3	(N_1,C_3)	(N_2,C_3)
C_4	(N_1,C_4)	(N_2,C_4)

The space of possible inflection classes can then be defined with reference to the allomorphic variation exhibited by individual cells. It is usually assumed that two inflection classes can be distinguished if they exhibit (morphologically conditioned) allomorphy in any cell. From this assumption, it follows that the maximally allomorphic cell defines the minimal number of classes. This minimum just reflects the fact that there must be at least as many classes as there are realizations for the maximally allomorphic cells, since otherwise some class would have multiple realizations for those cells. At the other extreme, the largest space of classes corresponds to the CELL PRODUCT, that is, the product of all of the realizations of the individual cells. In the case of the system in Table 5.2, the cell product is defined as $|\xi(N_1, C_1)| \times |\xi(N_1, C_2)|, \ldots, \times |\xi(N_2, C_4)|$. The cell product defines the largest class system, since it exhaustively enumerates all of the combinations of patterns exhibited by individual cells.

Continuing with the schematic system in Table 5.2, let *D* represent the system and $\mathcal{C}_D$ be the number of inflection classes in *D*. The assumptions adopted so far dictate only that the value of $\mathcal{C}_D$ must fall somewhere between the number of realizations associated with the maximally allomorphic cell and the cell product. However, as Carstairs (1983) remarks, inflection class systems are not normally distributed within this space. Instead, they appear to cluster closely around the number of realizations associated with the maximally allomorphic cell:

> when we apply the traditional notion of 'paradigm', we find that the actual total of paradigms is at or close to the minimum logically possible with the inflexional resources involved, and nowhere approaches the logical maximum. (Carstairs 1983: 127).

To account for this clustering, Carstairs (1983: 127) proposes the PEP as 'an absolute constraint on the organization of the inflexional resources for every word-class in every language' which has the effect of 'keeping the total of paradigms for any word-class close to the logical minimum'. The PEP thus requires that the class size of any word-class system is equal to ('or close to') the number of realizations associated with the maximally allomorphic cell. In principle, the PEP would appear to impose a maximally restrictive constraint on class size, since it effectively requires that each inflectional system must be organized into the smallest possible set of classes. In practice, the restrictiveness of the PEP depends on how close is 'close enough' for compliance and, more fundamentally, on how classes are individuated. One immediate qualification is introduced by the decision to 'disregard ... stem alternations' (Carstairs 1983: 120) and restrict the notion of 'inflectional resources' relevant to the PEP to 'affixal exponents'. This exclusion reduces the overall number of inflection classes in languages with distinctive, class-specific stem alternations, and contributes to the goal of bringing the count closer to the logical minimum.

Yet even as a constraint solely on the distribution of affixal resources, the PEP places an extremely tight constraint on the relationship between cells in a paradigm. In a

Table 5.3 Exponents of principal parts in German

Gen Sg	Ø, -(e)s, -(e)n, -(e)ns
Nom/Acc/Gen Pl	Ø∼-e, -¨(e), -er∼-¨er, -s, -(e)n

Source: Carstairs (1983: 125).

system that conforms to the PEP, each maximally allomorphic cell partitions the system into classes, and every pair of maximally allomorphic cells partition the system into the same classes. No cell in a paradigm can have realizations that vary independently of the realizations of any maximally allomorphic cell. This means that for every realization of every maximally allomorphic cell there will be a unique realization in every other cell (including other maximally allomorphic cells). This entails that every maximally allomorphic cell is diagnostic of class. Moreover, since every class contains a maximally diagnostic cell, it follows that every class will be identifiable from the realization of a single cell. So a system that conforms to the PEP will realize the pedagogical ideal in which a single principal part suffices to identify class.

5.3.2 Gender-driven Economy in German

There are *prima facie* counterexamples to any principle that imposes this kind of tight organization on paradigm structure. As Carstairs acknowledges, traditional descriptions of German noun declensions do not appear to conform to strictures of the PEP. Table 5.3 lists the exponents that Carstairs isolates from the traditional principal parts of these declensions.[7] Since the smallest class space is defined by the most maximally allomorphic cell, the five exponents that realize the non-dative plural in Table 5.3 define a minimum class size of five. This corresponds closely to the number of plural classes recognized in traditional sources such as Duden (2005) (provided that stem alternations are again disregarded).

The five basic plural classes P1–P5 are illustrated in Table 5.4, in which each pattern is classified by the shape of its plural ending and—for completeness—whether it is, or can be, based on an umlauted variant of the singular stem. A comparison of the forms in the first two blocks of rows shows that masculine and neuter nouns both follow all five patterns, although comparatively fewer neuter nouns follow the 'weak' pattern P2 and only a very few exhibit umlaut in P3. The bottom block of rows shows that feminine nouns do not form plurals in -*er* and that they tend to require umlauted stems in patterns P3 and P5.

[7] Parentheses and dashes mark alternations treated as phonologically conditioned or stylistic, and exponents of the form '-¨er' indicate an umlauted stem vowel.

Table 5.4 Plural patterns in German

	P1	P2		P3	P4		P5
Ending	-s	-(e)n	-e	-e	-er	0	Ø
Stem	–uml	–uml	–uml	+uml	+uml	–uml	+uml
N/A/G	Uhus	Prinzen	Hunde	Bünde	Münder	Balken	Gärten
Dat	Uhus	Prinzen	Hunden	Bünden	Mündern	Balken	Gärten
(Masc)	'owls'	'princes'	'dogs'	'waistbands'	'mouths'	'beams'	'gardens'
N/A/G	Autos	Ohren	Jahre	Flöße	Länder	Muster	Klöster
Dat	Autos	Ohren	Jahren	Flößen	Ländern	Mustern	Klöstern
(Neut)	'cars'	'ears'	'years'	'rafts'	'countries'	'patterns'	'cloisters'
N/A/G	Bars	Regeln	–	Hände	–	–	Töchter
Dat	Bars	Regeln	–	Händen	–	–	Töchtern
(Fem)	'bars'	'rules'	–	'hands'	–	–	'daughters'

Source: cf. Duden (2005: 226).

Table 5.5 Singular declensional patterns in German

	Masc			
			Neut	Fem
	S1	S2	S1	S3
Nom	Pegel	Prinz	Segel	Regel
Acc	Pegel	Prinzen	Segel	Regel
Dat	Pegel	Prinzen	Segel	Regel
Gen	Pegels	Prinzen	Segels	Regel
	'level'	'prince'	'sail'	'rule'

Source: cf. Duden (2005: 197).

To conform to the PEP, the choice of plural ending must determine the regular singular patterns, illustrated in Table 5.5.[8] Pattern S3 characterizes feminine nouns, which are invariant in the singular. Neuters likewise follow just the 'strong' pattern S1, in which the genitive singular is marked by -*(e)s*. Masculine nouns exhibit the only variation, as they may follow either the strong pattern S1, marked by -*(e)s* or the weak pattern S2, marked by -*(e)n*.

It is sometimes claimed for German that 'the choice of plural formation depends largely on gender and/or inflection class as manifested ... in the singular' (Laaha *et al.*

[8] This table omits the proper name class listed in Duden (2005: 197), which patterns with S1 nouns. It also excludes the pattern exhibited by NAME 'name' and a small handful of other nouns, which have accusatives and datives in -*en* and genitives in -*ens*.

Table 5.6 Combinations of singular and plural patterns in German

			S1	S2	S3
	uml	Masc	Neut	Masc	Fem
P1	–	UHU	AUTO	—	KAMERA
P2	–	STAAT	OHR	PRINZ	REGEL
P3	–	HUND	JAHR	—	—
P3	+	BUND	(FLOSS)	—	HAND
P4	+	MUND	LAND	—	—
P5	–	BALKEN	MUSTER	—	—
P5	+	GARTEN	(KLOSTER)	—	TOCHTER

Table 5.7 Co-occurrence of plural and genitive singular endings in German

		S1		S2	S3
	Pl	Masc	Neut	Masc	Fem
P1	*–s*	*–s*	*–s*	—	Ø
P2	*–(e)n*	*–s*	*–s*	*-(e)n*	Ø
P3	*–e*	*–s*	*–s*	—	Ø
P4	*–er*	*–s*	*–s*	—	—
P5	Ø	*–s*	*–s*	—	Ø

2006: 279). This claim is unfounded, as Table 5.6 shows. Matching the singular patterns in Table 5.5 with the plural patterns in Table 5.4 defines nineteen combinations, which highlight the dissociation of plural from singular patterns.

The affixal variation is isolated in Table 5.7, which plots the co-occurrence of plural and genitive singular endings.[9] These patterns are obtained from Table 5.6 by collapsing pairs of rows that differ solely in stem umlaut, and replacing the exemplary lexemes by plural and genitive endings.

At first glance, this pattern appears incompatible with the PEP, since every plural ending except -*er* in the second column corresponds to two or even three distinct genitive singular endings. Yet on closer inspection the system is in fact almost maximally economical. As Carstairs argues, traditional 'declensions' (and 'conjugations') can be brought into closer conformance with the PEP by restricting attention to genuine inflectional variants, and grouping paradigms into common classes if they differ

[9] Though omitting -*(e)ns*, which occurs with a small and declining set of nouns.

Table 5.8 Mildly uneconomical class space in German

	P1	P2		P3	P4	P5
Masc	S1	S1	S2	S1	S1	S1
Neut	S1	S1		S1	S1	S1
Fem	S3	S3		S3	–	S3

solely with respect to lexically-conditioned properties. In the specific case of German, Carstairs (1983: 126) observes that 'Gender is lexically determined for German nouns; and we can readily combine each of the Feminine-only "declensions" ... with some non-Feminine "declension", just as we traditionally combine Latin Neuters and non-Neuters in Declension II'. Returning to the classes defined by the P1, P3, and P5 plural endings in Table 5.7, one can see that they exhibit a general contrast between masculines and neuters, which are marked by -*s* in the genitive singular, and feminines, which are uninflected in the singular. The distribution of umlauted vowels (suppressed in Table 5.7) is likewise a lexical property of nouns in classes P3 and P5.

The strongest challenge to the PEP is posed by class P2. There is, first of all, some artificiality in collapsing what are arguably two distinct plural strategies. Plurals in -*(e)n* are the default for 'native' feminine nouns, as well as for feminine nouns formed with productive endings such as -*ung* and -*heit*, as expressed in the second 'basic rule' (*Grundregel*) in Duden (2005: 193). The formation of masculine and neuter plurals in -*(e)n* is much more restricted and cannot be regarded as productive in the modern language.[10] Even within the non-productive subclass, neuters have unambiguous singular forms. So it is just the masculines that show inflectional variation, with weak nouns such as PRINZ following the weak singular pattern S2, and mixed nouns like STAAT following the strong singular pattern S1. The resulting class space is exhibited in Table 5.8.

Hence by consolidating gender-conditioned variation, one can arrive at a space of six classes, assuming just two P2 subclasses, or seven classes, if productive feminines in -*(e)n* are distinguished from the masculines and neuters. This exceeds, but is indisputably close to, the minimum of five determined by the patterns of plural exponence. One can in principle bring the system into conformance with the PEP by coercing plurals in -*(e)n* into a single class and making a 'specific exemption' for mixed paradigms, as Carstairs (1983: 127) suggests. This kind of exception is of course very different from the consolidation of lexically-conditioned variation. However, it is also true that the evaluation of a general principle like the PEP should not hinge too directly on patterns that even traditional sources regard as mixed or hybridized.

[10] There is also psychological and neurolinguistic evidence for distinguishing feminines in -*(e)n* from the corresponding masculines and neuters (Clahsen 1999; Penke *et al.* 1999).

It is arguably more instructive to diagnose the general characteristics of German that allow the PEP to work, or at least to work as well as it does. One pivotal property is the clean dissociation of inflectional and lexical variation. The patterns of plural affixation P1–P5 can reasonably be regarded as inflectional, while affixal variation between the singular patterns S1 and S3 can, as Carstairs argues, be attributed to lexical factors. Hence there is only one case, involving masculines of class P2, in which there is inflectional variation in both the plural and singular. It is precisely this case that is problematic for the PEP. Although the separation of lexical and inflectional variation promotes economical paradigm organization in German, this separation is less a feature of inflection class systems in general than a symptom of the near-complete loss of singular contrasts in German. Hence a useful test case for the PEP comes from declensional systems that exhibit inflectional contrasts in the singular and plural.

5.3.3 Uneconomical Partitives in Finnish

One immediate consequence of the relation between paradigm economy and predictability is that the PEP will tend to be violated by any language whose description requires more than one principal part, since multiple parts imply that the realization of multiple cells must vary independently. Finnish provides a useful test case, given that descriptions of declensions often list up to five principal parts (typically the nominative, genitive, and partitive singular, along with the partitive plural and a plural 'local' case such as the inessive). Some of these forms identify stem alternations that are disregarded in determining compliance with the PEP, while others identify affixal patterns. Since the local case forms have mostly invariant endings, the grammatical case exponents in Table 5.9 exhibit nearly all of the inflectional variation that is relevant to the PEP.

Examination of Table 5.9 identifies the partitive singular as the most highly differentiated cell, with three distinct patterns -*A*, -*tA*, and -*ttA*. The archiphoneme '*A*' ranges over the vowels *a* and *ä*, which show harmony with the final stem vowel, so that the

Table 5.9 Grammatical case exponents in Finnish

	SG	PL
Nom	Ø	-t
Gen	-n	-den∼-tten, -en
Part	-A, -tA, -ttA	-A, -tA

Source: Karlsson (1999).

Table 5.10 Grammatical case forms in Finnish

Part Sg	Part Pl	Part Sg	Part Pl	(Gen Pl)	
A	A	asemaa	asemia	(asemien)	'position' (13) [15]
A	tA	perunaa	perunoita	(perunoiden)	'potato' (17) [3]
tA	A	lohta	lohia	(lohien)	'salmon' (33) [24]
tA	tA	leikkuuta	leikkuita	(leikkuiden)	'haircut' (25) [18]
ttA	tA	huonetta	huoneita	(huoneiden)	'room' (78) [4]
A	A∼tA	karitsaa	karitsoja	(karitsojen)	
			karitsoita	(karitsoiden)	'lamb' (15) [6]
A∼tA	A	lahtea	lahtia	(lahtiin)	
		lahta			'bay' (34) [12]

Source: Pihel and Pikamäe (1999).

three patterns in Table 5.9 have six surface realizations. The partitive plural is realized by the first two endings, -*A* and -*tA*, but these endings are usually described as varying independently of the partitive singular endings. The genitive plural is realized by three endings, but two of these, -*den* and -*tten* are regarded as variants.[11] The grammatical case cells in Finnish thus have at most three realizations, and determine a cell product of 12 (3 × 2 × 2).

The variation exhibited by grammatical case forms (according to standard descriptions) is set out in Table 5.10.[12] The first two columns in Table 5.10 plot the co-occurrence of partitive singular and plural exponents, followed in the next two columns by forms of exemplary items (including predictable genitive plurals). We can with no loss of generality restrict attention to word types with unique partitive singular and plural affixes, exhibited in the first five rows in Table 5.10.

These word types define the five classes C1–C5 in Table 5.11. A class size of five is closer to the maximum of six defined by the product of the partitive exponents than to the minimum of three dictated by the PEP. But it is still far from the cell product of 12, due to the fact that genitive plural endings do not add any new classes. Instead, variation in the genitive plural largely conforms to the expectations of the PEP in aligning with the partitive plural, with -*A* implying -*en* and -*tA* predicting -*den* (Karlsson 1999: 92f.).

The challenge presented by Finnish is that the partitive plural is itself NOT predictable from the partitive singular. The four logical possibilities defined by the exponents -*A*, and -*tA* are illustrated in the first four rows in Table 5.10: singular -*A* co-occurs

[11] 'The ending -*den* can always be replaced by the ending -*tten*' (Karlsson 1999: 93).

[12] The numbers in parentheses identify the word type number from the authoritative *Nykysuomen sanakirja* 'Dictionary of modern Finnish' (Häkkinen 1990), which are used in Pihel and Pikamäe (1999), and the number in square brackets indicates how many of the 85 word types exhibit that pattern of partitive exponence.

Table 5.11 Classes defined by unique partitive exponents

	C1	C2	C3	C4	C5
Part Sg	A	A	tA	tA	tA
Part Pl	A	tA	A	tA	ttA
Gen Pl	en	den	en	den	den

with each of the plural endings -*A* and -*tA* and the same is true of singular -*tA*. Since there is no grammatical gender in Finnish, it is not possible to collapse the classes in Table 5.11 into 'macroparadigms' whose internal variation is conditioned by gender differences, as in the previous case of German. Nor are there any other morphosyntactic properties that condition the variation in partitive exponents. It is also not plausible to maintain the PEP by treating the patterns in Table 5.10 as conditioned by variation in the shape or prosody of stems. Instead, the independent variation of partitive singular and plural exponents define a product class space that violates the PEP.

As the discussions of German and Finnish clarify, the PEP is ultimately a constraint on the form of morphological DESCRIPTIONS rather than on the organization of the systems themselves. This is to some degree unavoidable, given that inflection classes are part of a descriptive frame of reference that is imposed on languages rather than 'discovered' in them. Questions about the number and type of inflection classes in a language cannot even be formulated without specifying the purposes for which the classes are defined. The status of classes is obscured to some degree by the fact that the agreement regarding class individuation in traditional sources derives from common descriptive or pedagogical goals. However, this consensus does not rest on any general task-independent criteria for distinguishing classes. Two paradigms that exhibit distinct inflectional patterns in each of their cells would be assigned to distinct classes in any system of classification. But how many cells must vary? Does one always suffice, or does partial overlap between two paradigms give rise to 'subclasses'? Does it matter if variation involves suppletion or some other form of irregularity? How many items must follow an inflectional pattern in order to constitute a class? Presumably single suppletive items do not form classes. But what then is the item threshold that separates classes from residual suppletive patterns?

The issues raised by these questions are not addressed by dictating a consistent policy governing the definition of classes, subclasses, item-specific patterns, etc. The true challenge lies in arriving at a principled, task-independent, basis for this definition. The fact that this challenge is seldom if ever confronted reflects an understanding that there is not some fixed number of classes to be discovered in a language, but rather that the number assigned depends on the level of precision to which the classes are described or the purposes for which they are defined. The following quotation describes the situation in Finnish; the status of Estonian is comparable (Blevins 2007, 2008).

> There is no consensus on how many inflectional classes there are for nominals and verbs. Traditional Finnish lexicography as manifested in *Nykysuomen sanakirja* (Dictionary of modern Finnish, 1951–1961) postulates 82 inflectional classes for nominals, whereas at the other extreme, a generative description such as Wiik (1967) operates with none but a wealth of ordered (morpho)phonological rules. A surface-oriented morphological approach would recognize at least 10 nominal inflectional classes. (Karlsson 2006: 476)

This indeterminacy highlights limitations of economy conditions which impose a numerical bound on constructs, in this case 'inflection classes', that are part of the description of a morphological system. The most direct development of the PEP is the No-Blur Constraint (Carstairs-McCarthy 1994), which expresses economy in terms of restrictions on the relation between affixal resources and classes. This alternative thus inherits any limitations that accrue from a purely affixal measure or from assumptions about the determinacy of inflection classes.[13] Nevertheless, these issues mainly pose challenges to the specific formulation of PEP and successor constraints. Constraints that attempt to characterize the highly economical and cohesive organization of inflectional paradigms draw attention to a pattern which 'by itself is enough to establish the paradigm as an important concept within linguistic theory' (Carstairs 1983: 127). The observation underlying the PEP also cannot be dismissed as a descriptive artifact. The 'tendency' that Carstairs (1983: 127) discerns 'towards keeping the total of paradigms for any word class close to the logical minimum' seems real enough, even if it does not appear to be susceptible to formalization directly in terms of numerical bounds on class size or affixal classification.

5.4 Modelling Paradigmatic Structure

From a classic WP perspective, paradigm economy effects do not derive from constraints that regulate the economical organization of paradigms or the 'inflectional resources' distributed across their cells. Rather, these effects are taken to be secondary reflexes of the implicational structure of an inflectional system.

This structure is not explicitly formalized in the classic WP tradition but merely exhibited by interdependencies between forms of exemplary items. The implicative Paradigm Structure Conditions proposed by Wurzel (1984) in the context of a model of Natural Morphology thus provide one of the first explicit representations of the structure of inflectional paradigms.

[13] Some of the empirical consequences of these assumptions are discussed in Stump (2005b).

> Observation of complicated paradigms shows that implicative relations do not only obtain between one basic inflexional form, either lexical (*sobaka* [Russian 'dog' JPB]) or non-lexical (*Männer* [German 'men' JPB]), and all the other inflexional forms, but exist throughout the whole paradigm: all paradigms (apart from suppletive cases) are structured on the basis of implicative patterns which go beyond the individual word, patterns of varying complexity. Of particular complexity in this respect is, for example, the implicative pattern of the *i*-declension in Latin: /im/ in the A.Sg. ⊃ /ī/ in the Abl.Sg. ⊃ /īs/ in the A.Pl. ⊃ /ium/ in the G.Pl. ⊃ ... From the A.Sg. form we can derive, via a number of steps, all other forms, but not vice-versa. Since the implicative patterns determine the structure of the paradigms of a language we call them implicative PARADIGM STRUCTURE CONDITIONS (PSCs). (Wurzel 1984: 208)

Yet the use of the material conditional '⊃' to model implicational relations constrains the descriptive scope of PSCs, restricting their application to exceptionless patterns. Some patterns are exceptionless, as in the case of the relation between Finnish partitive singulars and genitive plurals in Table 5.10. However, in many if not most patterns, one cell or form is only partially informative about another, and a general analysis of implicational structure must be able to represent this partial information, and the ways in which it can be combined.

5.4.1 Implication as Uncertainty Reduction

Standard notions from information theory permit a much more flexible representation of implicational patterns. The basic idea is to think of morphological variation first of all in terms of uncertainty; the more variants of a cell there are (and the more uniform their distribution), the harder it is to guess which variant realizes a particular cell. Then implicational dependencies correspond directly to uncertainty reduction; there is an implicational relation between one cell and another to the extent that knowing the realization of the first cell reduces uncertainty about the realization of the second. Unlike logical implication, this is not an all or nothing relation. Knowledge of one cell may merely reduce the uncertainty of another, without eliminating it altogether. For example, as Table 5.6 indicates, knowing that a German noun has the plural ending *-(e)n* implies that it is either a feminine noun that follows the (invariant) singular pattern S3 or a masculine noun that inflects according to singular pattern S2.

These intuitive notions have very straightforward formalizations. To begin with, the (self-)information carried by an exponent x can be expressed in terms of SURPRISAL, as in (1). The informativeness of the exponent, I_x, corresponds to the negative log of its probability $p(x)$, which just reflects the fact that the less likely x is, the more information it conveys when it does occur.

(1) $I_x = -\log_2 p(x)$

The uncertainty or ENTROPY (Shannon 1948) of a cell is then defined by averaging the surprisal values of the exponents that realize the cell. The entropy of a cell is defined in (2) in terms of a standard entropy definition by treating a cell C as a random variable that takes as values a set $\mathcal{X}$ of exponents $x_1 \ldots x_n$.

(2) $$H(C) = -\sum_{x \in \mathcal{X}} p(x) \log_2 p(x)$$

The surprisal value of an exponent cannot be estimated without information about its distribution, and the same information is required for an accurate measure of cell entropy. However, even without distributional information, a worst-case estimation of cell entropy can be obtained by treating the patterns of exponence as equiprobable. This reflects the fact that the uncertainty measured by entropy increases as a function of the number of alternatives and the uniformity of their distribution (since both factors make it more difficult to guess the value). In this limiting case, the probabilities 'cancel out' and entropy reduces to the $\log_2$ of the number of alternatives in the set $\mathcal{X} = x_1 \ldots x_n$, as in (3).

(3) $$H(C) = \log_2(n)$$

For the purposes of linguistic analysis, an extremely useful property of entropy is that estimations can be obtained from a range of descriptive sources. Entropy ceilings can be obtained for languages that are only described by traditional grammars, provided that the grammars contain a complete description of form inventories. The more balanced and detailed the corpus resources available for a language, the more accurate the corresponding entropy measures.

For a simple illustration, consider the Finnish grammatical case exponents, repeated in Table 5.12. The partitive singular is a maximally allomorphic cell, with three exponents. In the example in Table 5.12, the worst-case entropy of the partitive singular cell is $\log_2(3)$, or approximately 1.58 bits. This value assumes that the exponents are equally probable; any distributional skewing that makes any exponent more (or less) likely than the others will lower the entropy.

Entropy represents the uncertainty associated with a cell or paradigm. Once an uncertainty estimation has been obtained, implicational structure can be modelled by uncertainty reduction. Uncertainty reduction is expressed by CONDITIONAL ENTROPY,

Table 5.12 Grammatical case exponents in Finnish

	SG	PL
Nom	{Ø}	{-t}
Gen	{-n}	{-den∼-tten, -en}
Part	{-A, -tA, -ttA}	{-A, -tA}

$H(C_2|C_1)$, which measures the amount of uncertainty that remains about C_2 given knowledge of C_1. This relation can be defined as in (4).

(4) $$H(C_2|C_1) = -\sum_{x \in \mathcal{X}} p(x) \sum_{y \in \mathcal{Y}} p(y|x) \log_2 p(y|x)$$

As (2), the cells C_1 and C_2 are random variables that take as values the sets of exponents $\mathcal{X} = x_1 \ldots x_n$ and $\mathcal{Y} = y_1 \ldots y_m$. The only new element in (4) is $p(y|x)$, the conditional probability of an exponent y given an exponent x.

5.4.2 Diagnosticity and Economy

These basic notions of uncertainty and uncertainty reduction permit a straightforward representation of the implicational structure of a morphological system. The more information that a cell C_1 provides about another cell C_2, the lower the conditional entropy value $H(C_2|C_1)$ will be. If knowledge of C_1 completely identifies the value of C_2, then $H(C_2|C_1)$ will approach 0. If C_1 provides no information about C_2, then $H(C_2|C_1)$ will preserve the uncertainty of $H(C_2)$.

The selection of principal parts can be guided by entropy reduction. A set of cells function as principal parts within a word class whenever, for any paradigm in that class, knowledge of those cells reduces uncertainty about any other cells in the paradigm. There need not be a unique diagnostic cell or set of cells. The availability of a range of different diagnostic combinations clearly enhances the robustness of form deduction, since speakers are not dependent on encountering one uniquely informative form of a paradigm. The informal idea that diagnosticity correlates with variability across word or inflection classes can be expressed more precisely in terms of the uncertainty reduction measured by the conditional entropy. From the present perspective, one can see that the choice of leading forms IS to a large degree arbitrary. A pedagogical or reference grammar might use seemingly arbitrary criteria to select a particular cell or cell set. A description might select the smallest set of cells, the set with the most highly frequent members, or, more capriciously, the cell or cell set with the morphologically simplest members, etc. Since any fully diagnostic set of cells will do, all are equally suitable and the arbitrariness involved in selecting one is harmless.[14]

An uncertainty-based approach also offers a useful perspective on the central insight expressed by the PEP, namely that the variation exhibited by maximally allomorphic cells constrains the class-defining distribution of variants in other cells. Taking the number of variants in the maximally allomorphic cell to define class size (or distribution of variants) determines a tight constraint on the distribution of allomorphs—too tight as the example of Finnish suggests.

[14] There is also a trade-off between the number of cells required to identify class in a system and whether one uses the same cells to identify class, as Finkel and Stump (2007, 2009) show.

However, the same fundamental intuition can be expressed by means of a measure that defines economy more directly in terms of relations between maximally allomorphic cells and other cells. An economy measure based on the interdependence of cells neither requires an independent consensus about the number of classes in a language nor methods for imposing minimized descriptions, since interdependencies tend to be preserved irrespective of the number of classes or the 'granularity' of the classification. To some degree the central role of cell interdependence is obscured in the paradigm economy literature by the decision to view inflectional resources in terms of separate inventories of stems and inflections, given that the disassembly of forms disrupts implicational patterns based on whole words (termed 'gestalt exponence' in Ackerman *et al.* 2009).

Maximally allomorphic cells are significant within a system because they determine bounds on variation within the system. By definition, a maximally allomorphic cell is associated with the largest number of variants. Another way of putting this is that maximally allomorphic cells exhibit the greatest 'morphological uncertainty'; it is hardest to guess the realizations of these cells, because they are associated with the largest number of choices. The variation in a maximally allomorphic cell also serves a diagnostic function within a system. Either the variation in such a cell correlates with the variation in other cells, or else the system consists of mutually independent cells. This relation can again be phrased in terms of uncertainty. Knowing the realization of a maximally allomorphic cell reduces uncertainty about the realization of at least some other cells, except in the case where the cells of the system are mutually independent.

To express these relations schematically, let C_M be a maximally allomorphic cell and $H(C_M)$ represent the uncertainty associated with C_M. For a pair of cells C and D, let $H(C|D)$ represent the uncertainty that remains about C given knowledge of D. Then a paradigm consisting of cells $C_1, \ldots, C_n$ will exhibit the minimal interdependence assumed in a classic WP model if $H(C_i)$, the average uncertainty associated with the cells, falls within the range specified in (5).

(5) $H(C_i|C_M) < H(C_i) \leq H(C_M)$

The bounds associated with maximally allomorphic cells in (5) merely ensure that a system does not consist of mutually independent cells. The average uncertainty $H(C_i)$ is required to be greater than the average uncertainty given knowledge of a maximally allomorphic cell, $H(C_i|C_M)$, and no greater than the uncertainty of the maximally allomorphic cell, $H(C_M)$, itself. Economy principles impose further restrictions within these limits. The PEP requires in effect that knowledge of a maximally allomorphic cell reduces all uncertainty, that is, that $H(C_i|C_M)$ approaches zero. This follows from the assumption that the realization of a maximally allomorphic cell identifies the class of an item, together with the assumption that class determines the realizations of the other cells.

The formulation of economy in terms of uncertainty clarifies two issues. The first is that classes function essentially as proxies for patterns of interdependency in the PEP. A constraint that stipulates that maximally allomorphic cells eliminate uncertainty in other cells achieves the effect of the PEP without any reference to classes. The second issue is that the focus on maximally allomorphic cells exhibits the same bias as traditional principal part analyses. A system that conforms to the PEP is one in which a single diagnostic form from a maximally allomorphic cell determines all of the forms of a non-exemplary item. In much the same way that principal part analyses tend to ignore the informativeness of non-diagnostic forms, the PEP defines economy solely in terms of the correlation between maximally allomorphic cells and patterns of interdependency. From the perspective of uncertainty reduction, there is no reason to privilege maximally allomorphic cells in this way, since the interdependency between other cells will also determine the cohesion and economy of a system.

The idea that knowledge of some forms of an item reduces uncertainty about other forms of the item remains implicit in classic WP models. The first systematic attempt to measure this effect is proposed in the context of the Low Entropy Conjecture of Ackerman and Malouf (2013), which suggests that the average uncertainty of a cell given knowledge of another is not only lower than the uncertainty of the cell in isolation, but falls within a fairly narrow range.

5.5 Conclusion

This chapter has devoted considerable attention to the role played by implicational relations in inflectional paradigms. In part this focus reflects the fact that the implicational structure of inflectional paradigms is their most distinctive characteristic and also provides some of the strongest evidence for recognizing paradigms as fundamental units of lexical organization. Although implicational relations appear to be central to the organization of language in general, their effects are particularly salient in the domain of inflection. It is the fact that inflection operates over a closed, relatively uniform space that makes inflectional processes highly predictable, and in turn contributes to their productivity. Any speaker who knows that AUTO is a regular count noun in English can deduce that it has a 2-cell paradigm defined by the number contrast between singular *auto* and plural *autos*.[15] The two forms also predict each other, given that the regular plural marker is phonologically conditioned in English. From the fact that German *Auto* is neuter and ends in a vowel that does not conform to native nominal phonotactics, a German speaker can deduce that the plural forms and genitive singular of the noun AUTO (classes P1/S1 in Table 5.6) are realized by *Autos*, while all other singular cells are realized by *Auto*. The twenty-two-cell paradigm of the Finnish counterpart AUTO (class 1, represented by ASEMA in Table 5.10) is likewise defined by the two numbers and eleven regular cases in Finnish. The interpredictability of forms

[15] Though it does not affect the point at hand, the 'Anglo-Saxon' genitive marker *'s* is almost certainly best regarded as a phrasal affix (Hockett 1947: 242; Anderson 1992: 212; Blevins 2006a), not as a lexical case exponent, as suggested in Zwicky (1987) and Huddleston and Pullum (2002).

is more complex in Finnish, due to larger form inventories, consonant gradation and other sources of variation. Yet, as in English, interpredictability is facilitated by the uniformity of declensional paradigms. A speaker encountering a new noun in Finnish can assume that it has twenty-two case-number forms, and even deduce the shape of those forms with levels of confidence that depend on the morphological information expressed by the form they have encountered.[16]

Contributing to the regularity of inflection is the fact that each cell *C* in a feature space is related to other cells along two dimensions. On the one hand, *C* is related to other cells in the same paradigm, that is, to cells that have the same intrinsic features as *C*. On the other hand, *C* is related to counterpart cells in other paradigms that have different intrinsic features from *C* but contain the same (or similar) paradigmatic features. The uniform organization of inflectional paradigms contrasts starkly with the more variable structure exhibited by 'families' of derivationally-related forms. Derivational processes are traditionally classified as 'exocentric' in the sense that they define new items by changing the word class, valence, or other intrinsic properties of the items to which they apply. Accounts that attempt to demarcate inflection from derivation often focus on the fact that such 'property-changing' processes are non-monotonic. But from the present perspective, it is at least equally important that they also do not define a finite set of forms within a uniform feature space. From the word class of an item, one cannot in general predict the number and type of derivational formations in which it occurs. Given a list of derivational processes active in a language, it is of course possible to assign a uniform family of 'potential' forms to all of the members of a word class. Yet the uniformity achieved is deceptive, because it collapses a critical distinction between those forms that are established in a language and those that are merely possible in principle.[17] Defining expanded derivational paradigms does not make it any easier for a speaker to predict the derivational formations that are attested and in use within a language. The contrast with compounds is even clearer. Of the infinitely many possible noun compounds in English only a comparatively small number are established, and a speaker cannot predict the set of established compounds containing an item from the item itself. The lack of predictability is due in large part to the high degree of item-specific variation exhibited by derivational families, which does not give rise to form classes that sanction class-specific inferences.

The contrast between the cohesive implicational structure of an inflectional paradigm and looser confederations of derivational forms are also reflected in the ways that morphological systems are described and analysed. It is rare for an inflected form to be described as established on its own. The availability of a form tends to correlate instead with the productivity of a whole pattern. Conversely, notions like

[16] There may also be some residual comitative, abessive, or instructive forms.

[17] The available psycholinguistic evidence suggests that established forms, not potential forms, are psychologically relevant. A range of experimental studies have found that the processing of items in Germanic, Finnic, and Semitic, among other languages, is facilitated by the semantically transparent items that make up its morphological family and inhibited by semantically opaque items (Schreuder and Baayen 1997; de Jong 2002; Moscoso del Prado Martín *et al.* 2004a, 2004b).

'morphological gap', 'suppletion', and even, to a large degree, 'syncretism' are mainly or exclusively applied to inflectional paradigms. Derivational families do not usually have 'gaps' because they do not define implicational relations over a closed and uniform space that give rise to expectations about the existence of specific forms, even where the shape would be predictable. Suppletion can likewise only arise where there are definite assumptions about the shape of particular members of a form set. Whereas inflectional paradigms generate strong assumptions of this nature, derivational families do not. Syncretism presupposes a similar structure, as syncretic patterns imply the existence of independent cells that can be associated with fully or partially identical forms. But the notion of a derivational cell is not defined in the absence of a set of features that specify a morphosyntactic grid within which to place the cells.

More generally, descriptions of inflectional systems tend to assume a closed and uniform feature space that sanctions inferences about the existence and shape of forms. In many languages, there is evidence that this space is further partitioned into units that correspond to the traditional notion of a paradigm and that these units serve as the locus of form deduction and economy effects.