The canonical approach in typology¹

Greville G. Corbett
Surrey Morphology Group
University of Surrey

To appear in: **Zygmunt Frajzyngier, David Rood & Adam Hodges (eds)**
***Linguistic Diversity and Language Theories*. Amsterdam: Benjamins.**
[Proofs returned September 2004.]

Introduction

We are invited to consider ‘the proper object for theories of language structure’. Following Chomsky’s suggestion, we should aim to define the notion ‘possible human language’. As a part of that we need to develop the notion ‘possible word’. We were also asked to examine ‘the rich variation of forms and functions observed in the languages of the world’. This is the realm of typology, and I shall suggest that for typologists to make a full contribution we may need to refine our methods, as we move into more difficult areas of the typology of possible words.

As one way forward I suggest a ‘canonical’ approach. This approach sidesteps two potential dangers in typology, namely ‘premature statistics’ and ‘not comparing like with like’. The first danger is that something which is frequently found may be treated as uninteresting, whereas there are linguistic phenomena which are common yet which, I believe, should surprise us. The second danger is that we fail to take sufficient care over our terminology and so do not see that phenomena labeled identically are in fact distinct (conversely we miss identities because of different traditions of labeling).

In a canonical approach, we take definitions to their logical end point and build theoretical spaces of possibilities. Only then do we ask how this space is populated. There are of course precedents for this approach; for instance Greenberg (1959) can be read in this light. It follows that canonical instances (the best examples, those most closely matching the canon) may well not be the most frequent. They may indeed be extremely rare, or even non-existent. However, they fix a point from which occurring phenomena can be calibrated, and it is then significant and interesting to investigate frequency distributions.

I shall illustrate this canonical approach, and present some practical outcomes. The point is to create theoretical spaces, to populate them while the languages are still there to be investigated, and to study both what is frequent and what is rare. The paper brings together material from different projects, to illustrate the canonical approach in various areas.

¹ The support of the ESRC under grant R00027135 is gratefully acknowledged. I wish to thank the following for helpful discussion of some of the issues: Matthew Baerman, Dunstan Brown, Marina Chumakina, Andrew Hippisley and Carole Tiberius. Versions of the paper were given at the University of Kentucky 9 May 2003, and at the conference “Linguistic Diversity and Language Theories”, Boulder, Colorado, 14-17 May 2003. I am indebted to both audiences for helpful discussion.

1. Syntax

As an instance of syntactic typology we consider the complex phenomenon of agreement. Here we are interested in canonical agreement, agreement in gender between adjective and noun in the noun phrase. Recall that by canonical, we mean ‘best’, ‘clearest’, ‘indisputable’ (according to the ‘canon’). We shall also see weakenings of the criteria, giving instances which are less canonical, and others which are on the fringe of agreement. Just some of the criteria are given here for illustration: full justification is given in Corbett (2003c).

1.1. Definitions

The term agreement commonly refers to some systematic covariance between a semantic or formal property of one element and a formal property of another. Steele (1978: 610)

We call the element which determines the agreement (say the subject noun phrase) the **controller**. The element whose form is determined by agreement is the **target**. The syntactic environment in which agreement occurs is the **domain** of agreement. And when we indicate in what respect there is agreement, we are referring to agreement **features**. Thus number is an agreement feature, it has the values: singular, dual, plural and so on. This is diagrammed in Figure 1.

Figure 1: Framework of terms

If agreement is determined by factors which do not themselves mark agreement, then these are agreement **conditions**. Note that the controller-target relation is logically asymmetrical, but this does not imply asymmetry in the syntactic model. In fact, most modern approaches use some form of unification.

1.2. Canonical examples

As an instance of canonical agreement, consider agreement in gender in the Italian noun phrase:

Italian (Pierluigi Cuzzolin, personal communication)

(1)	il	nuov-o	quadr-o
	DEF.SG.M	new-SG.M	picture(M)-SG ²
	‘the new picture’		
(2)	i	nuov-i	quadr-i
	DEF.PL.M	new-PL.M	picture(M)-PL
	‘the new pictures’		
(3)	la	nuov-a	tel-a
	DEF.SG.F	new-SG.F	painting(F)-SG
	‘the new painting’		
(4)	le	nuov-e	tel-e
	DEF.PL.F	new-PL.F	painting(F)-PL
	‘the new paintings’		

Put briefly, the canonical features of these examples are as follows (they are listed in Appendix 1):

controller: is present, has overt features, and is consistent in the agreements it takes, part of speech is not relevant (this is a vacuous criterion here)

target: has bound expression of agreement, obligatory marking, doubling the marking of the noun, marking is regular, alliterative, productive; the target has a single controller and its part of speech is not relevant

domain: agreement is asymmetrical (the gender of the adjective depends on that of the noun), local, and the domain is one of multiple domains

features: lexical (in one instance), matching values, not offering any choice in values

conditions: no conditions

The different canonical aspects of agreement converge so that agreement in gender of the modifier with the noun in the noun phrase emerges as the canonical instance. Phenomena which extend the instances ‘outwards’ are grouped under the five components of our account of agreement. For illustration we will consider just the criteria which relate to controllers.

1.3. Controller criteria

Several criteria relate to the controller. An important one is that canonical controllers are present.

Criterion 1: controller present > controller absent
 (‘>’ is to be read as ‘is more canonical than’)

Compare these two similar examples:

² Glossing follows the Leipzig glossing rules
 (<http://www.eva.mpg.de/lingua/index.html>)

Russian

- (5) ty čitaš'
 you read.2SG
 'you are reading'

Serbian/Croatian/Bosnian

- (6) čitaš
 read.2SG
 'you are reading'

In such sentences in Russian the controller is typically present, while in Serbian/Croatian/Bosnian typically it is not. Serbian/Croatian/Bosnian is a pro-drop language. We treat as canonical what is sometimes called 'grammatical agreement' rather than 'anaphoric agreement' (Bresnan, and Mchombo 1989, Siewierska 1999, Bresnan 2001). An effect of adopting Criterion 1 is that the canonical type is restricted to relatively few languages, since pro-drop is common.

While discussions of 'dropping' concentrate on pronouns, we are making a more general point here: it is more canonical for any controller to be present rather than absent. For agreement of the adjective with the noun within the noun phrase, it is more canonical for the noun to be present; similarly in possessor-possessed agreement it is more canonical for the 'possessed' to be present.³

Criterion 2: controller has overt features > controller has covert features

Compare these Russian examples:

Russian

- (7) ona spa-l-a
 she sleep-PST-SG.F
 'she was sleeping'
- (8) ja spa-l / spa-l-a
 I sleep-PST[SG.M] / sleep-PST-SG.F
 'I was sleeping' (man/woman speaking)

In (7) the controller is overtly feminine: the pronoun *ona* 'she' contrasts with *on* 'he'. In (8) the controller *ja* 'I' does not distinguish gender. We treat examples like (7) as canonical in this respect, rather than those like (8). Another way of stating this criterion is that a canonical controller marks at least as many distinctions as the target. It does so in two respects: in terms of the number of features, and in terms of their values.

On the basis of these criteria (and others not discussed here), a more general principle may be suggested (compare Moravcsik 1988: 90):

Principle-I: Canonical agreement is redundant rather than informative

³ I am grateful to Andrew Spencer for this observation.

In the Russian example (7), the feminine feature is available from the controller (criterion 2). In (8) it is not. Agreement in the canonical example is redundant. Similarly, English examples like *the horse is/the horses are* are more canonical than *the sheep is/are*. The situation where there is no controller present, and hence the only information about the controller is that supplied by the target (as in pro-drop constructions) is non-canonical (though, as we noted, it is commonly found); this is the point of Criterion 1.

Let us continue with other criteria relating to controllers.

Criterion 3: consistent controller > hybrid controller

A consistent controller is one which controls a consistent agreement pattern. This is more canonical than one which controls different feature values. The notion ‘consistent agreement pattern’ is intuitively easy, but not quite so easy to define (for the details see Corbett 1991: 176-181). As a basic characterization, a consistent agreement pattern is the set of agreements controlled by a typical regular controller. In Table 1, it is represented by a row.

attributive adjective	predicate	relative pronoun	personal pronoun	values
-yj	-Ø	-yj	on-Ø	singular masculine
-aja	-a	-aja	on-a	singular feminine
-oe	-o	-oe	on-o	singular neuter
-ye	-i	-ye	on-i	plural

Table 1: Agreement patterns in Russian

A hybrid controller, on the other hand, takes agreements from more than one such pattern. It controls different feature values on different targets.

An example can be found in the Talitsk dialect of Russian (Bogdanov 1968). In this dialect, a plural verb can be used with a singular noun phrase, to indicate reference to a person or persons besides the one indicated directly. That is to say, we have an ‘associative’ construction, but it is indicated not by a marker on the nominal, but by plural agreement:

Talitsk dialect of Russian (Bogdanov 1968)

- (9) moj brat tam tóža žýl'-i⁴
 my[SG] brother[SG] there also lived-PL
 ‘my brother and his family also lived there’

The plural agreement is found in the verbal predicate, but not in the noun phrase, and so we have different agreements according to the target. Of course, consistent controllers take either singular or plural agreements, irrespective of the target: that is, they have consistent agreement patterns. The inconsistent pattern does not represent a row in Table 1.

The possible patterns of agreement with hybrid controllers are tightly constrained by the Agreement Hierarchy (Corbett 1979, 1991: 225-260, 2000: 188-192; Cornish

⁴ Bogdanov’s transcription has been transliterated here.

1986: 203-211; Barlow 1991, 1992: 136-137; Kirby 1999: 92-96). The Agreement Hierarchy distinguishes four types of agreement target:

attributive - predicate - relative pronoun - personal pronoun

Figure 2: The Agreement Hierarchy

The Agreement Hierarchy allows us to constrain possible agreement patterns as follows:

For any controller that permits alternative agreement forms, as we move rightwards along the Agreement Hierarchy, the likelihood of agreement forms with greater semantic justification will increase monotonically (that is, with no intervening decrease).

The considerable range of data covered by the constraints of the Agreement Hierarchy can be found in the references above.

Criterion 4: controller's part of speech irrelevant > relevant (given the domain)

The idea is that given a domain, for instance, subject-predicate agreement, in the canonical case we do not need further information on the part of speech. For instance, in Russian we do not need to have different rules for a subject noun phrase headed by a noun as compared to one headed by a pronoun. Sometimes, however, the difference is substantial. A good example is Bayso, where the rules are rather different for pronouns as compared with nouns. For this complex situation see Hayward (1979), Corbett and Hayward (1987), Corbett (2000: 181-183).

These two criteria fall under a second general principle:

Principle II: Canonical agreement is syntactically simple

Agreement varies from examples which can be captured by a relatively simple rule, to those which are exceptionally complex. The two criteria, Criterion 3 and Criterion 4, both point to agreement phenomena which can be captured by simple and general rules.

1.4. Principles of canonical agreement

Several further criteria can be grouped under a third principle:

Principle III: The closer the expression of agreement is to canonical (i.e. affixal) inflectional morphology, the more canonical it is as agreement.

We return to canonical inflectional morphology in §3.1 below. Note that the principles, and the criteria which they summarize, never conflict. They therefore define a multi-dimensional space. The convergence of all the criteria is a point at which truly canonical agreement is described.

The canonical approach allows us to clarify some of the conceptual problems and misunderstandings that characterize the problem of agreement (see Evans 2003, Mithun 2003, Polinsky 2003, for further discussion). Having seen the gradient nature of many of the properties (as well as the ways in which they overlap), the question of

‘drawing the line’ between agreement and other phenomena appears secondary. It is more important to understand agreement and its related phenomena than to draw a precise line at which we might claim agreement ‘stops’ and some other phenomenon begins.

1.5. Relevance for a typological database

For practical purposes, however, such as the construction of a typological database, we do need to draw a line between what is included and what is not. Here the results of our approach can be of use: we can be clear where and why such a line is drawn, and users of the database can aware how the data relate to their own conceptions and analyses of the area. If they wish to define the area of investigation more tightly, the means of querying the database allow this to be done.

The Surrey Database of Agreement (Tiberius, Brown, Corbett and Barron 2002) includes detailed information on a small, carefully chosen set of genetically diverse languages: Basque, Chichewa, Georgian, Hungarian, Kayardild, Mayali, Ojibwa, Palauan, Qafar, Russian, Tamil, Tsakhur, Turkana, Yimas and Yup’ik. The database was designed and implemented in ACCESS by Dunstan Brown and Roger Gentry, and is described in Tiberius, Brown and Corbett (2002).

Figure 3: Structure of the Surrey Database of Agreement

The design incorporates the five elements of our account in §1.1 above, which means that the researcher can query the database according to controllers, targets, domains, features (categories) and conditions. In each of these the researcher may choose to restrict the search further (and we anticipate that those who restrict their search

because of a different view of agreement will often choose the more canonical instances and omit others).

The database has a relatively small number of languages. It is nevertheless quite large, because for each agreement phenomenon in each language we give a wealth of information wherever possible (thus besides straightforward subject-verb agreement, we include quantified subjects, subjects consisting of conjoined noun phrases and so on, giving the features involved, and any conditions or options). It is intended to ‘open doors’: to enable researchers to test hypotheses and to find data relevant to their research questions, and the sources to go further. There are links to examples in each instance; this means that the researcher can check back to the original data and so can be sure what the entries in the database mean.

For each language included there is also a prose report, written by the researcher who entered the data for the language, giving sources, and enabling the user to see how decisions were made. Since some of the analytical decisions are difficult (we are dealing with areas where there is considerable uncertainty in the field), these reports are valuable in allowing the user to see the approach of the researcher, and to treat the data accordingly. In some instances experts on the particular languages contributed a good deal and their generous help significantly strengthened the database. For more discussion of the underlying philosophy see Corbett (2003a); the database is freely available at <http://www.smg.surrey.ac.uk/>.

In our discussion of canonicity we have stressed that what is canonical need not be frequent. The Italian examples cited earlier (1)-(4) are familiar, but not particularly common in the world’s languages. While they are familiar, it is important that we should not lose sight of their surprising nature. They show information ‘in the wrong place’, that is information about the noun expressed on the adjective. The information is ‘displaced’ (Moravcsik 1988: 90). In assigning them to a type, we should not cease to be surprised by such examples.

When discussing canonical agreement we asserted without discussion that one of the principles is that canonical agreement is expressed by canonical inflectional morphology. We take up this topic in the next section

2. Morphology

We start from the notion of a canonical paradigm, that is, an idealized morphosyntactic paradigm, for distinct word classes in individual languages. Setting this up is not always straightforward. We first have to establish the morphosyntactic features and their values (on the basis of some distinct inflectional behaviour). Then we establish which are valid for each word class. Finally we project all the logically possible feature value combinations to produce the underlying morphosyntactic paradigm. We thus ‘multiply out’ the logical possibilities to produce an idealized morphosyntactic paradigm. A theoretical example for nouns is given in Figure 4.

	SINGULAR	DUAL	PLURAL
NOMINATIVE			
ACCUSATIVE			
GENITIVE			
DATIVE			
INSTRUMENTAL			
LOCATIVE			

Figure 4: An abstract (morphosyntactic) paradigm

2.1. Four dimensions for a morphological typology

Suppose that the features relevant to inflection in a given language are established and we have a set of paradigms, like the one in Figure 4. We might then expect that for any given lexeme, every cell of its paradigm will be filled by the inflectional system (completeness), each form will be different (distinctiveness), the stem will be predictable, and the inflections will be predictable. A straightforward view of the function of inflectional morphology could lead to that expectation, and it provides a useful definition of canonical morphology. It can also be seen as the situation for which an Item and Arrangement model would be adequate. (See Stump 2001: 1-30 for discussion of the necessary characteristics of a realistic theory of inflectional morphology.)

The mismatches between these canonical expectations and the variety of real inflectional systems lead to our typology.

canonical expectation	phenomenon
completeness of paradigm	defectiveness
distinctiveness of cells	syncretism
regularity of stems	suppletion
regularity of inflection	'deponency'

Figure 5: Canonical expectations and corresponding morphological phenomena

These four phenomena are investigated in Corbett (1999). For illustration here we shall discuss just two of them, the most widespread – syncretism (§2.2), and perhaps the most challenging for a typologist – suppletion (§2.3).

2.2. Syncretism

The assumption of completeness is that each cell in a paradigm should be filled by a different form (which is what a functional view might predict). Thus in our abstract paradigm there 'should' be 18 different forms. What then if only 17 are found? The identity of form may be due to syncretism. Various fairly similar definitions of syncretism have been offered:

Identity in form between two grammatically different inflections.

Trask (1997: 215)

The relation between words which have different *morphosyntactic features but are identical in form. ... Used especially when the identity is regular across all paradigms. [* see also 'case syncretism']

Matthews (1997: 367)

Russian adjectives display **syncretism**, that is, a single inflected form may correspond to more than one morphosyntactic description.

Spencer (1991:45)

In instances of syncretism, two or more cells within a lexeme's paradigm are occupied by the same form. Stump (2001: 212)

Let us look more closely at why the cells should be claimed to be different when the forms are the same. There may be two different reasons:

1. other items of the same class have different forms. This is 'internal' syncretism, indisputably relevant to the grammar of the particular language.
2. values of a feature have a separate realization in one environment, and the projection of these values to the morphosyntactic paradigm produces the additional cell(s). This is 'external' syncretism, essential to the typologist, but not necessarily to be reflected in the grammar of the particular language.

The difference can be made clear by asking which is the shape of the paradigm of Russian past tense verbs:

	SG	PL		SG	PL
MASC	byl	byl-i	MASC	byl	byl-i
FEM	byl-a	byl-i	FEM	byl-a	
NEUT	byl-o	byl-i	NEUT	byl-o	

Figure 6: 'External' syncretism (Russian *byl* 'was')

We see that Russian verbs in the past tense distinguish two numbers and three genders. This gives an abstract paradigm with six forms, as on the left. This shows that there is syncretism for gender in the context of plural number. However, no verb has more than four forms in the past tense, and our morphological description of Russian may operate with these (implying the paradigm on the right), see Brown (1995).

On the basis of this canonical approach, including external syncretism, we have created a typological database of syncretism (Baerman, Brown and Corbett 2002a, available at: <http://www.smg.surrey.ac.uk/>). The design is given in Figure 7, and is described in Brown (2001).

Figure 7: Structure of the Surrey Syncretism Database

The database records all instances of inflectional homophony in thirty genetically diverse languages, comprising 1256 separate entries. It is based on the notion of abstract morphosyntactic paradigms (discussed above) and every type of cell identity, as compared with that canonical expectation, is recorded in the database. Note that the shape of the theoretical paradigm could vary according to other assumptions (and links to language reports in the database make clear what the assumptions were) but even then the notion of all cells being canonically different still holds (Matthew Baerman, personal communication).

For recent discussion of syncretism see Evans, Brown and Corbett (2001), Stump (2001: 212-241), Baerman, Brown and Corbett (2002b), which highlights the unusual syncretisms found in Indo-European, and Corbett, Baerman and Brown (2002), and for an extended account see Baerman, Brown and Corbett (forthcoming).

2.3. Suppletion

Given a further assumption that canonically a lexeme will have a single stem or a series of predictably related stems (§2.1), then we must address the issue of suppletion. Suppletion has been defined as follows:

The use of two or more entirely different **stems** in forming the inflections of a single lexical item, as in English *good/better, go/went, person/people*.
Trask (1997: 212)

Morphological process or alternation in which one form wholly replaces another. Thus in *went* either the whole form, or a stem *wen-*, is in a suppletive relationship to *go*.

In **partial suppletion** only a part of the form is replaced. E.g. in *thought* (or *though-t*), the *th-* of *think* is unchanged and only *-ink* is affected. Matthews (1997: 364)

We may think of suppletion from two slightly different points of view. One view treats it at the end case of a cline of inflectional irregularity, an instance of particularly severe irregularity (naturally with partial suppletion as less irregular than full suppletion). Another view is that suppletion represents an extreme, rather different from more minor irregularities. We shall take both points of view, since each provides illustration for a canonical approach.

If we start from the idea that suppletion is part of a cline of irregularity this might lead us to flesh out the scale, arranging items from fully regular to suppletive. This was attempted by Corbett, Hippiusley, Brown and Marriott (2001) where such an irregularity scale was used to investigate in detail the distribution of the regular and irregular expression of number in Russian texts. We examined the nouns in the one million word Uppsala corpus of Russian (see Lönnegren 1993 for details) and recorded all the lexemes occurring at least five times. Our dataset contains around 5440 lexemes, accounting for some 243 000 word forms from the entire corpus. Each noun in our dataset was assessed according to the irregularity scale proposed.

Having made explicit what would count as regular, less regular and fully irregular (suppletive) we had to ask what is meant by frequency. Frequency can be viewed in two ways. We might compare lexemes one with another or we could compare regular and irregular forms within lexemes. For the first approach, we could count up how many times each lexeme occurs in the plural. We call this the *absolute frequency* of a lexeme's plural. We can then compare the absolute frequency of plural of different lexemes, regular and irregular, to see if there is a relationship between irregular plurals and their absolute frequency. Alternatively we could analyse the plural by comparing it, within the lexeme, with the other available forms. For a given lexeme, we can count how often it occurs in the plural as compared with the number of times in the singular. This is the *relative frequency* of the plural. We can then compare the relative frequency of the plural in lexemes where it is irregular with that in lexemes where it is regular.

We found relations between frequency and irregularity and a certain degree of correspondence with the Irregularity Scale. Seven out of eight groups of nouns confirm the hypothesis of a relation between irregularity and absolute plural anomaly. A clear result is that the three nouns of modern Russian which show suppletion stand out dramatically. We found some evidence that the frequency of occurrence of the irregular forms, and not just frequency of occurrence of the lexeme as a whole, does relate to irregularity of the forms in question (see the cited paper). As far as (full) suppletion is concerned, though the median plural proportion is high, the result is not statistically significant. The difficulty is that there are so few nouns with suppletive stems. For the full statistical details see Corbett, Hippiusley, Brown and Marriott (2001)

We now consider suppletion as an extreme, and since this is a particular challenge for typologists we devote a separate section to it.

3. Typology of the extreme

There is also good reason for treating suppletion as an extreme phenomenon. It requires us to allow for lexical items whose constituent parts have no shared phonology. If we attempt a typology of suppletion, concentrating on ‘full’ suppletion, this will indeed be an unusual sort of typology, a typology of extremes. The object will be lexemes, not languages. For a canonical approach here, a good starting point is Mel’čuk’s definition:

For the signs X and Y to be suppletive their semantic correlation should be maximally regular, while their formal correlation is maximally irregular. Mel’čuk (1994: 358)

Starting from this general definition, we can establish dimensions along which the phenomenon may vary. We can in many cases point to the ‘canonical’ or best instances, those which are maximally transparent in semantic terms and maximally opaque in formal terms (cf. Mel’čuk 1994: 342, forthcoming). We write ‘a > b’, for ‘a is a more canonical example than b’. This means that we can recognise, for example, that some restrict suppletion to inflectional morphology, while others including Mel’čuk allow for suppletion in derivational morphology. Semantic correlations are typically more regular in inflectional than in derivational morphology, hence the clearer (and for some linguists the only) instances of suppletion will be found in inflectional morphology.

Corbett (2003b) proposed 15 criteria for canonical suppletion. We shall consider just some of them here, to illustrate the good perspectives that a canonical approach offers (§3.1). We shall then investigate two of the interactions of suppletion with other morphological phenomena (§3.2).

3.1. Some criteria for a typology of suppletion

For illustration we consider some of the criteria for canonical suppletion, those which deal with the elements which can be suppletive, their phonology and, most difficult, their distribution.⁵

3.1.1. Element involved

Criterion 1: fused exponence > stem

This criterion is discussed in various terms. The basic notion is that if the suppletive form combines stem and inflection, this is ‘more suppletive’ than if the form is the stem to which appropriate inflections are added. Thus *worse* is a more canonical instance of suppletion than *better*, being more opaque, since *better* may be analysed as a suppletive stem *bett-* with the comparative marker *-er*. Nübling (1998: 78-79) talks of Segmentierbarkeit ‘segmentability’ versus Nichtsegmentierbarkeit, and makes

⁵ For discussion of the rise of suppletion and for the complication of alternating suppletion, see Chumakina, Hippisley and Corbett (forthcoming).

the point that items with a segmentable affix are ‘etwas weniger suppletiv’ ‘somewhat less suppletive’ than those without.

3.1.2. *Phonology of this element*

This section depends on the part of the definition stating that in suppletion the ‘formal correlation is maximally irregular’.

Criterion 2: full > partial

This is also known as strong versus weak suppletion (e.g. Nübling 1998:78)

(10) Russian: *idti* ‘go’ ~ *šel* ‘went’ (full suppletion)

(11) English *bring* ~ *brought* (partial suppletion)

For suppletion there should be no formal correlation. If we take that at its simplest, we could say that (10) shows suppletion and (11) does not. However, we can take the requirement to be that there is no rule, no regularity linking the two forms. Nevertheless, the partial phonological similarity of the English forms means that they are ‘less suppletive’ than the Russian forms. This brings us back to the issue of viewing suppletion both as the end point of a cline and as a special extreme phenomenon.

3.1.3. *Distribution of suppletive forms*

This is one of the key issues in the typology of suppletion. Given that a lexeme has two or more suppletive forms, the question is how they can be distributed. This criterion represents the conclusion of what will be some careful discussion:

Criterion 3: morphological > morphosyntactic

Here we follow the terms used in Stump’s (2003) account of heteroclisis. Alternative terms are ‘morphomic’ (Aronoff 1994: 22-29) or ‘morphologically systematic’ versus ‘featural’. The point is that the distribution of stems within a paradigm may conform to a morphological pattern or to a morphosyntactic feature specification. Stump (2002: 149-153) shows how the Network Morphology use of two default inheritance hierarchies allows us to capture this distinction. The link to the general definition is that morphosyntactic determination represents less pure semantic regularity than morphological determination (where there is no possible compromising of semantic regularity).

To see what it involved, consider this Russian lexeme:

	singular	plural
nominative	rebenok	deti
accusative	rebenka	detej
genitive	rebenka	detej
dative	rebenku	detjam
instrumental	rebenkom	det'mi
locative	rebenke	detjax

Table 2: The Russian noun *rebenok* ‘child’

A Slavist might well suggest that the stems are distributed according to the morphosyntactic feature of number. A Romance specialist might ask whether this example could not be described in purely morphological terms. This requires us to look at the stem patterns of Russian nouns. There are various patterns, but most give a singular-plural split (for details see Corbett 2000: 139-142).

As implied above, there is a tradition in Romance linguistics of treating the distribution of suppletive forms as determined by morphological patterns or templates (Matthews 1981, Vincent 1988: 297-298, Maiden 1992: 306-307 and Aski 1995). This is based on examples like the following:

	singular	plural
1 person	vais	allons
2 person	vas	allez
3 person	va	vont

Table 3: Present tense of French *aller* ‘go’

There are many French verbs which show various stem modifications, following this pattern of cells (verbs like *acheter* ‘buy’, to name just one). Of course, such a pattern can be described in featural terms, but the point of the Romance linguists’ claim is that the morphological pattern is established as a part of the morphological system (Maiden 2002). See Hippiusley, Chumakina, Corbett and Brown (2004) for further examples from outside Romance.

Before accepting morphological determination of suppletion, we should consider Carstairs-McCarthy’s statement (1994: 4410) that the distribution of suppletive elements will be (morpho)syntactic or phonological. He cites examples like Italian *va(d)- ~ and-* ‘go’, stating that the suppletion is determined by the stress: *va(d)-* is found where the stem is stressed (Carstairs 1988: 71). An account based on stress would be possible for the French example, as long as we stay with the present tense; for the rest of the distribution, this seems less plausible.

Given that the different potential determining factors frequently overlap, we should ask which examples are available which require us to acknowledge particular types of determining factor. The Russian example (Table 2) could be treated as morphological or morphosyntactic. The French example (Table 3) could be treated as morphological or possibly phonological. We need to establish which of these are required. Consider this Polish example:

	singular	plural
first	jestem	jesteśmy
second	jesteś	jesteście
third	jest	są

Table 4: Present tense of Polish *być* ‘be’

Here we find a suppletive relation between the third plural and all the remaining cells (Rothstein 1993: 717; Itkin 2002). There is no phonological criterion to account for this distribution. There is, however, a morphological pattern, as shown by other verbs, where there is an irregular stem alternation, as with *wiedzieć* ‘know’:

	singular	plural
first	wiem	wiemy
second	wiesz	wiecie
third	wie	wiedzą

Table 5: Present tense of Polish *wiedzieć* ‘know’

Several verbs share this pattern (there is another pattern in which first singular and third plural are opposed to the other forms). These Polish data show that the distribution can follow a morphological pattern, rather than a phonological pattern. Like the French data, they support a morphological rather than a morphosyntactic account.

We must ask which types are required. The instances of possible phonological determination can also be treated as morphological patterns. However, the Polish example is morphological in nature and is not analysable in phonological terms. Hence we need to recognize morphological patterns as a possible determiner for the distribution of suppletive stems, and it is not clear whether we need also to recognize phonological determination. The key evidence would be a paradigm where the forms could be determined by some phonological generalization and not by morphological or morphosyntactic regularities. I have not been able to find such an example, and conclude tentatively that we do not require phonological determination of suppletion.⁶

Since we need to allow for morphological determination of patterns of suppletion, we should ask whether we need morphosyntactic determination. The examples given so far do not require this. However, the Norwegian data to be discussed in §4.2.2 do suggest that morphosyntactic determination must be recognized as a possible type.

3.1.4. Relevance for a database of suppletion

The canonical approach has again proved useful in setting up a database of suppletion. Like the other typological databases, this one, developed by Dunstan Brown, Marina Chumakina, Greville Corbett and Andrew Hippisley has been made freely available over the web (at <http://www.smg.surrey.ac.uk/>). It draws on a sample of over 30 genetically diverse languages. We have adopted a strict version of Mel’čuk’s definition; this means that many of the instances are relatively close to canonical ones and that claims made on the basis of the data included will be founded on clear-cut examples.

⁶ Andrew Carstairs-McCarthy points out (personal communication) that if, as here, affixes are not considered suppletive, then this makes it easier to maintain the claim that phonological determination of suppletive elements is not required.

3.2. Interactions with other morphological phenomena

When we have a clear notion each of the morphological phenomena from a canonical perspective, there is another major task, namely to investigate the possible interactions between these phenomena. Here, as illustration, it makes sense to investigate the interaction between suppletion and the other phenomenon we took as a case study, namely syncretism. We then look at suppletion interacting with the ‘opposite’ of syncretism.

3.2.1. *Syncretism*

When discussing interactions, we mean that the same cells in the paradigm are potentially affected by the two phenomena (and not just that a lexeme independently shows an instance of both phenomena). Interactions between syncretism and suppletion are rare, but there is one remarkable instance, found in the South Slavonic language Slovene, discussed by (Plank 1994, Corbett and Fraser 1997, and Evans, Brown and Corbett 2001: 215). In Slovene there are syncretisms affecting all nouns, which results in the identity of the genitive dual and genitive plural, and of the locative dual and locative plural. The unique noun *človek* ‘man, person’ has a suppletive stem for the plural, as distinct from singular and dual. Syncretism requires the stems to be identical while suppletion requires that they be distinct. The ‘outcome’ is given in Table 6.

	SINGULAR	DUAL	PLURAL
NOM	človek	človeka	ljude
ACC	človeka	človeka	ljude
GEN	človeka	ljude	ljude
DAT	človeku	človekoma	ljudem
INST	človekom	človekoma	ljudmi
LOC	človeku	ljudeh	ljudeh

Table 6: Slovene *človek* ‘man, person’ (Priestly 1993: 401)

One interpretation is that syncretism dominates suppletion. The data have also been proposed as evidence in favour of rules of referral, since the dual forms refer to the plural in an asymmetric way.

3.2.2. *Overdifferentiation*

We take overdifferentiation as the second example of a phenomenon interacting with suppletion, because it is in one sense the opposite of syncretism. Overdifferentiated lexemes are those which have an additional form in their paradigm, like English *be* (Bloomfield 1933: 223-224). A typical case of overdifferentiation is that of an element which distinguishes three gender values in a language where normally only two are distinguished. For instance, in Kolami, there are two genders, male human and other, for almost all agreement targets. However, the numerals ‘two’, ‘three’ and ‘four’ – and no other agreement targets – have additional forms for female human; these three numerals are said to be overdifferentiated (Emeneau 1955: 56, Corbett 1991: 168-169).

Suppletion can interact with overdifferentiation. Hans-Olav Enger has provided the following interesting data on his East Norwegian dialect (and a comparable situation

is found in some other dialects).⁷ The adjective ‘small’ has suppletive stems as follows:

Norwegian (East Norwegian dialect, Hans-Olav Enger p.c.)

(12)	en
ART.M.SG.INDF
‘a little boy’	lit-en
little-M.SG.INDF	gutt
boy(M)[SG.INDF]
(13)	den
ART.M/F.SG.DEF
‘the little boy’	vesle
little.SG.DEF	gutt-en
boy(M)-SG.DEF
(14)	ei
ART.F.SG.INDF
‘a little girl’	lit-a
little-F.SG.INDF	jent-e
girl(F)-SG.INDF
(15)	den
ART.M/F.SG.DEF
‘the little girl’	vesle
little.SG.DEF	jent-a
girl(F)-SG.DEF
(16)	et
ART.N.SG.INDF
‘a little child’	lit-e
little-N.SG.INDF	barn
child(N)[SG.INDF]
(17)	det
ART.N.SG.DEF
‘the little child’	vesle
little.SG.DEF	barn-et
child(N)-SG.DEF

The form *vesle* is only found in the singular definite. In the plural, irrespective of definiteness and irrespective of gender, another suppletive form namely *små* is used. While there are three genders, as the articles show, a normal adjective in the singular distinguishes masculine and feminine on the one hand, from neuter on the other. For example, *tjukk* ‘thick, fat’ (masculine and feminine singular indefinite) versus *tjukt* (neuter singular indefinite). Thus *lit-* is overdifferentiated, distinguishing masculine (*lit-en*), feminine (*lit-a*) and neuter (*lit-e*) indefinite singular (Enger and Kristoffersen 2000: 104).

It has a further surprise. A typical adjective does not distinguish definite plural from definite singular, for instance *tjukk-e* serves for both. In the lexeme we are interested in, *vesle* is the definite singular, but in the plural *små* is used. The lexeme is thus further overdifferentiated, distinguishing singular from plural within the definite part of the paradigm. In the positive, it has five forms where the typical adjective has three. Thus suppletion interacts with overdifferentiation, in that the suppletive form may introduce values not normally distinguished.

We now have the data to take further the discussion in §4.1.3 concerning the distribution of suppletive stems within a paradigm. We saw there that there is

⁷ In the dialect the forms given are obligatory. Other Norwegians I have asked find them acceptable, but consider the use of *vesle* optional.

evidence for distribution according to a morphological pattern. Now the East Norwegian data provide evidence for morphosyntactic determination. The stems *lit-*, *vesle* and *små* are distributed according to the features of number and definiteness. The distribution does not follow a morphological pattern – adjectives do not normally make all these distinctions. Thus we have evidence for a lexeme which requires morphosyntactic distribution of its suppletive stems.

We have sketched a part of a typology of suppletion. Our aim here was not to lay out a full typology, since that is still work in progress, but rather to show how the canonical approach in typology applies well even to ‘extreme’ phenomena such as suppletion.

4. Conclusion

We have looked at an approach to defining ‘possible human language’ and have given most attention to possible words. We have proposed a ‘canonical’ approach in typology, and have illustrated it with examples from syntax and morphology. We have seen how we can build theoretical spaces and then look for distributions within them. This approach works well even with extreme phenomena, as illustrated by our partial typology of suppletion. The examples which arise from the interaction of suppletion with other morphological phenomena will require us to extend our notion of what is a possible word.

References

- Aronoff, Mark. 1994. *Morphology by itself: Stems and inflectional classes* (Linguistic Inquiry Monographs 22). Cambridge MA: The M.I.T. Press.
- Aski, Janice. 1995. “Verbal suppletion: an analysis of Italian, French and Spanish *to go*.” *Linguistics* 33.403–432.
- Baerman, Matthew; Brown, Dunstan and Corbett, Greville G. 2002a. *The Surrey Syncretisms Database*. Available at: <http://www.smg.surrey.ac.uk/>
- Baerman, Matthew; Brown, Dunstan and Corbett, Greville G. 2002b. “Case syncretism in and out of Indo-European.” In: Mary Andronis, Christopher Ball, Heidi Elston and Sylvain Neuvel (eds) *CLS 37: The Main Session. Papers from the 37th Meeting of the Chicago Linguistic Society, Vol. 1*, 15–28. Chicago: Chicago Linguistic Society.
- Baerman, Matthew; Brown, Dunstan and Corbett, Greville G. forthcoming. *Syncretism*. Cambridge: Cambridge University Press.
- Barlow, Michael 1991. “The Agreement Hierarchy and grammatical theory.” In: Laurel A. Sutton, Christopher Johnson and Ruth Shields (eds) *Proceedings of the Seventeenth Annual Meeting of the Berkeley Linguistics Society, February 15–18, 1991: General Session and Parasession on the Grammar of Event Structure*, 30–40. Berkeley: Berkeley Linguistics Society, University of California.
- Barlow, Michael 1992. *A Situated Theory of Agreement*. New York: Garland.
- Bloomfield, Leonard 1933. *Language*. New York: Holt, Rinehart and Winston.
- Bogdanov, V. N. 1968. “Osobyj slučaj dialektного согласования skazuemogo s podležščim po smyslu i kategorija predstavitel’nosti.” *Naučnye doklady vysšej školy: filologičeskie nauki* no. 4. 68–75.
- Bresnan, Joan. 2001. *Lexical-functional syntax*. Oxford: Blackwell.
- Bresnan, Joan and Mchombo, Sam. 1989. “Topic, pronoun, and agreement in Chichewa.” *Language* 63.741–782.
- Brown, Dunstan. 1995. *rusverbs.dtr* [a fragment of the verbal system of Russian]. Available at: <ftp://ftp.cogs.sussex.ac.uk/pub/nlp/DATR/dtrfiles/rusverbs.dtr>

- Brown, Dunstan. 2001. "Constructing a typological database for inflectional morphology: the SMG database for syncretism." In: Steven Bird, Peter Buneman and Mark Liberman (eds) *Proceedings of the IRCS Workshop on Linguistic Databases*. Philadelphia: Institute for Research in Cognitive Science, University of Pennsylvania. 56-64. [Available at: <http://www ldc.upenn.edu/annotation/database/proceedings.html>.]
- Brown, Dunstan; Chumakina, Marina; Corbett, Greville G. and Hippisley, Andrew. 2004. *The Surrey Suppletion Database*. Available at: <http://www.smg.surrey.ac.uk/>
- Carstairs, Andrew. 1988. "Some implications of phonologically conditioned suppletion." In: Geert Booij and Jaap van Marle (eds) *Yearbook of Morphology 1988*, 67-94. Dordrecht: Kluwer.
- Carstairs-McCarthy, Andrew. 1994. "Suppletion." In: R.E. Asher (ed.) 1994. *Encyclopedia of Language and Linguistics*: 8, 4410-4411. Oxford: Pergamon.
- Chumakina, Marina; Hippisley, Andrew and Corbett, Greville G. forthcoming. "Istoričeskie izmenenija v ruskoj leksike: slučai čeredujuščegosja suppletivizma". To appear in *Russian Linguistics*.
- Comrie, Bernard. 2003. "When agreement gets trigger-happy." In: Dunstan Brown, Greville G. Corbett and Carole Tiberius (eds) *Agreement: A Typological Perspective* (special number of *Transactions of the Philological Society* 101), 313-337. Oxford: Blackwell.
- Corbett, Greville G. 1979. "The Agreement Hierarchy." *Journal of Linguistics* 15.203-224.
- Corbett, Greville G. 1991. *Gender*. Cambridge: Cambridge University Press.
- Corbett, Greville G. 1999. "Defectiveness, syncretism, suppletion, 'deponency': four dimensions for a typology of inflectional systems". Guest lecture at The Second Mediterranean Meeting on Morphology, September 1999, Malta.
- Corbett, Greville G. 2000. *Number*. Cambridge: Cambridge University Press.
- Corbett, Greville G. 2003a. "Agreement: The range of the phenomenon and the principles of the Surrey Database of Agreement." In: Dunstan Brown, Greville G. Corbett and Carole Tiberius (eds) *Agreement: A Typological Perspective* (special number of *Transactions of the Philological Society* 101), 155-202. Oxford: Blackwell.
- Corbett, Greville G. 2003b. "A typology of suppletion and the notion 'Possible Word'. Paper read at the "Workshop on "Paradigm Irregularities", University of Manchester, 10-11 April 2003.
- Corbett, Greville G. 2003c. "Agreement: Canonical instances and the extent of the phenomenon." In: Geert Booij, Janet DeCesaris, Angela Ralli and Sergio Scalise (eds) *Topics in Morphology: Selected papers from the Third Mediterranean Morphology Meeting (Barcelona, September 20-22, 2001)*, 109-128. Barcelona: Universitat Pompeu Fabra. Available at: <http://www.surrey.ac.uk/LIS/SMG/projects/agreement/>
- Corbett, Greville G.; Baerman, Matthew and Brown, Dunstan. 2002. "Domains of syncretism: a demonstration of the autonomy of morphology." In: Mary Andronis, Christopher Ball, Heidi Elston and Sylvain Neuvel (eds) *CLS 37: The Panels. Papers from the 37th Meeting of the Chicago Linguistic Society*, Vol. 2, 385-398. Chicago: Chicago Linguistic Society.
- Corbett, Greville G. and Fraser, Norman. 1997. "Vyšislitel'naja lingvistika i tipologija." *Vestnik MGU: Serija 9: Filologija* no. 2.122-140.
- Corbett, Greville G. and Hayward, Richard J. 1987. "Gender and number in Bayso." *Lingua* 73.1-28.
- Corbett, Greville G.; Hippisley, Andrew; Brown, Dunstan and Marriott, Paul. 2001. "Frequency, regularity and the paradigm: a perspective from Russian on a complex relation." In: Joan Bybee and Paul Hopper (eds) *Frequency and the Emergence of Linguistic Structure*, 201-226. Amsterdam: John Benjamins.
- Cornish, Francis 1986. *Anaphoric relations in English and French: A Discourse Perspective*. London: Croom Helm.

- Emeneau, M. B. 1955. *Kolami: A Dravidian Language* (University of California Publications in Linguistics vol. 12). Berkeley and Los Angeles: University of California Press.
- Enger, Hans-Olav and Kristoffersen, Kristian E. 2000. *Innføring i norsk grammatikk: Morfologi og syntaks*. Oslo: Landlaget for Norskundervisning / Cappelen Akademisk Forlag.
- Evans, Nicholas. 2003. "Typologies of agreement: some problems from Kayardild." In: Dunstan Brown, Greville G. Corbett and Carole Tiberius (eds) *Agreement: A Typological Perspective* (special number of *Transactions of the Philological Society* 101), 203-234. Oxford: Blackwell.
- Evans, Nicholas; Brown, Dunstan and Corbett, Greville G. 2001. "Dalabon pronominal prefixes and the typology of syncretism: a Network Morphology analysis." In: Geert Booij and Jaap van Marle (eds) *Yearbook of Morphology 2000*, 187-231. Dordrecht: Kluwer.
- Greenberg, Joseph H. 1959. "A quantitative approach to the morphological typology of language." *International Journal of American Linguistics* 26.178-194.
- Hayward, Richard J. 1979. "Bayso revisited: some preliminary linguistic observations - II." *Bulletin of the School of Oriental and African Studies, University of London* 42.101-32.
- Hippisley, Andrew; Chumakina, Marina; Corbett, Greville G. and Brown, Dunstan. 2004. "Constraints on suppletion." *Studies in Language* 28.387-418.
- Itkin, I. B. 2002. "K tipologii suppletivnyx paradigm". Paper read at the Third Winter Typological School, Moscow District, 29 January - 6 February 2002.
- Kirby, Simon 1999. *Function, Selection and Innateness: the Emergence of Language Universals*. Oxford: Oxford University Press.
- Lönngren, Lennart 1993. *Častotnyj slovar' sovremennogo russkogo jazyka*. (=Acta Universitatis Upsaliensis, Studia Slavica Upsaliensis 33). Uppsala.
- Maiden, Martin. 1992. "Irregularity as a determinant of linguistic change." *Journal of Linguistics* 28.285-312.
- Maiden, Martin. 2002. "Sound change, morphemic structure and the rise of suppletion in the Romance verb". Paper presented at the *Surrey Linguistic Circle*, University of Surrey.
- Matthews, Peter. 1981. "Present stem alternations in Italian." In: Christian Rohrer (ed.), *Logos semantikos: Studia linguistica in honorem Eugenio Coseriu*, 1921-1981, vol. IV, 57-65. Berlin: de Gruyter.
- Matthews, P. H. 1997. *The Concise Oxford Dictionary of Linguistics*. Oxford: Oxford University Press.
- Mel'čuk, Igor. 1994. "Suppletion: toward a logical analysis of the concept." *Studies in Language* 18.2.339-410.
- Mel'čuk, Igor. forthcoming. Chapter 8 of *Aspects of the Theory of Morphology*. Berlin: Mouton-de Gruyter. [Revised version of Mel'čuk 1994.]
- Mithun, Marianne. 2003. "Pronouns and agreement: The information status of pronominal affixes." In: Dunstan Brown, Greville G. Corbett and Carole Tiberius (eds) *Agreement: A Typological Perspective* (special number of *Transactions of the Philological Society* 101), 235-278. Oxford: Blackwell.
- Moravcsik, Edith A. 1988. "Agreement and markedness." In: Michael Barlow and Charles A. Ferguson (eds) *Agreement in Natural Language: Approaches, Theories, Descriptions*, 89-106. Stanford: Center for the Study of Language and Information.
- Nübling, Damaris. 1998. "Zur Funktionalität von Suppletion." *Germanistische Linguistik* 141-142.77-101.
- Polinsky, Maria. 2003. "Non-canonical agreement is canonical." In: Dunstan Brown, Greville G. Corbett and Carole Tiberius (eds) *Agreement: A Typological Perspective* (special number of *Transactions of the Philological Society* 101), 279-312. Oxford: Blackwell.
- Plank, Frans 1994. "Homonymy vs. suppletion: A riddle (and how it happens to be solved in ...)" *Agreement gender number genitive & (EUROTYP Working Papers VII/23)* 81-86. Konstanz: University of Konstanz.

- Priestly, T. M. S. 1993. "Slovene." In: Bernard Comrie and Greville G. Corbett (eds), *The Slavonic Languages*, 388-451. London: Routledge.
- Rothstein, Robert A. 1993. "Polish." In: Bernard Comrie and Greville G. Corbett (eds), *The Slavonic Languages*, 686-758. London: Routledge.
- Siewierska, Anna 1999. "From anaphoric pronoun to grammatical agreement marker: why objects don't make it." In: Greville G. Corbett (ed.) *Agreement* (Special issue of *Folia Linguistica* XXXIII/2), 225-251.
- Spencer, Andrew 1991. *Morphological Theory: An Introduction to Word Structure in Generative Grammar*. Oxford: Blackwells.
- Steele, Susan. 1978. "Word order variation: a typological study." In: Joseph H. Greenberg, Charles A. Ferguson and Edith A. Moravcsik (eds) *Universals of Human Language: IV: Syntax*, 585-623. Stanford: Stanford University Press.
- Stump, Gregory T. 2001. *Inflectional Morphology: A Theory of Paradigm Structure*. Cambridge: Cambridge University Press.
- Stump, Gregory T. 2002. "Morphological and syntactic paradigms: Arguments for a theory of paradigm linkage." In: Geert Booij and Jaap van Marle (eds) *Yearbook of Morphology 2001*, 147-180. Dordrecht: Kluwer.
- Stump, Gregory T. 2003. "A theory of heteroclite inflectional paradigms." Ms. University of Kentucky.
- Tiberius, Carole; Brown, Dunstan and Corbett, Greville G. 2002. "A typological database of agreement." In: Manuel González Rodríguez and Carmen Paz Suárez Araujo (eds), *LREC2002: Third International Conference on Language Resources and Evaluation: Proceedings: VI*, 1843-1846. Paris: European Language Resources Association.
- Tiberius, Carole; Brown, Dunstan; Corbett, Greville G. and Barron, Julia. 2002. *The Surrey Database of Agreement*. Available at: <http://www.smg.surrey.ac.uk/>
- Trask, R. L. 1997. *A Student's Dictionary of Language and Linguistics*. London: Arnold.
- Vincent, Nigel. 1988. "Italian." In: Martin Harris and Nigel Vincent (eds) *The Romance Languages*, 279-313. London: Croom Helm/Routledge.

Appendix: Criteria for Canonical Agreement

Canonical agreement is outlined in §2 above and full details can be found in Corbett (2003c). The criteria are listed under the five components of our account of agreement, namely controllers, targets, domains, features and conditions. Comrie (2003) provides an important addition (Criterion 12).

1. *Controllers*

- C-1: controller present > controller absent
- C-2: controller has overt features > controller has covert features
- C-3: consistent controller > hybrid controller
- C-4: controller's part of speech irrelevant > relevant (given the domain)

2. *Targets*

- C-5: bound > free
- Criterion 5 can be expanded out as:
 - C-5': inflectional marking (affix) > clitic > free word
- C-6: obligatory > optional
- C-7: regular > suppletive
- C-8: alliterative > opaque
- C-9: productive > sporadic
- C-10: doubling > independent
- C-11: target agrees with a single controller > agrees with more than one controller
- C-12: target has no choice of controller > target has choice of controller
(is 'trigger happy', Comrie 2003))
- C-13: target's part of speech irrelevant > relevant (given the domain)

3. *Domains*

- C-14: asymmetric > symmetric
- C-15: local > non-local
- C-16: domain is one of set > single domain

4. *Features*

- C-17: feature is lexical > non-lexical
- C-18: features have matching values > non-matching values
- C-19: no choice of feature value > choice of value

5. *Conditions*

- C-20: no conditions > conditions

These criteria fall under three general principles, which are given in §1.3 and §1.4.