

## Agreement: Canonical instances and the extent of the phenomenon<sup>1</sup>

Greville Corbett  
Surrey Morphology Group  
g.corbett@surrey.ac.uk

### 1. INTRODUCTION

We attempt to clarify some of the conceptual problems that characterize discussions of agreement. We first establish ‘canonical’ instances of agreement, by which we mean best, clearest, indisputable (according to the ‘canon’); such cases need not be common. Then we discuss weakenings of the criteria as a result of which some but not all linguists would accept a particular phenomenon as agreement. In doing so, we set in place some of the underpinnings for a typological database of agreement.<sup>2</sup> We may start from Steele’s definition:

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

We shall 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**. 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. If agreement is determined by factors which do not themselves mark agreement, then these are agreement **conditions**.

---

<sup>1</sup> The support of the ESRC (UK) under grant R000238228 is gratefully acknowledged. Many thanks to Nicholas Evans, Marianne Mithun, Matthew Baerman, Dunstan Brown, Carole Tiberius and Marina Tchoumakina for discussion. A version was given at the Third Mediterranean Meeting on Morphology, Universitat Pompeu Fabra, Barcelona, 20-23.9.2001; I am very grateful to all present for lively debate.

<sup>2</sup> Details can be found at <http://www.surrey.ac.uk/LIS/SMG/projects/agreement/>

As an instance of canonical agreement, consider agreement in gender in the Spanish noun phrase (data from Max Wheeler personal communication):

- (1) un cuadro hermos-o  
INDEF.SG.MASC picture beautiful-SG.MASC  
'a beautiful picture'
- (2) un-os cuadro-s hermos-os  
INDEF-PL.MASC picture-PL beautiful-PL.MASC  
'beautiful pictures'
- (3) un-a cortina hermos-a  
INDEF-SG.FEM curtain beautiful-SG.FEM  
'a beautiful curtain'
- (4) un-as cortina-s hermos-as  
INDEF-PL.FEM curtain-PL beautiful-PL.FEM  
'beautiful curtains'

Put briefly, the canonical features of these examples are as follows:

- the controller (the noun in these examples): is present, has overt features, and is consistent in the agreements it takes; the part of speech is not relevant (a vacuous criterion in this particular instance)
- the target (article or adjective): has bound expression of agreement, obligatory marking, which doubles the marking of the controller, the marking is regular, alliterative, productive; the target has a single controller; the part of speech is not relevant
- the domain: agreement is asymmetrical (the gender of the adjective depends on that of the noun), local, and the domain is one of multiple domains

- the features: lexical (in one instance), matching values, offering no choice in values
- conditions: there are no conditions

## 2. CANONICAL AGREEMENT

The different canonical aspects of agreement converge so that agreement in gender of the modifier with the noun in the noun phrase comes out as the canonical instance. The criteria do not conflict. Phenomena which extend the instances ‘outwards’ are discussed in turn, grouped under the five components of our account of agreement.<sup>3</sup>

### 2.1. Controllers

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

*C-1: controller present > controller absent*

Here ‘>’ indicates ‘more canonical than’. Compare these two similar examples:

Russian	Serbian/Croatian/Bosnian
(5) ty čitaš’	(6) čitaš
you read.2.SG	read.2.SG
‘you are reading’	‘you are reading’

In such sentences in Russian the controller is typically present, while in Serbian/Croatian/Bosnian it is typically absent (an instance of pro-drop). We treat as canonical the type seen in (5), which is sometimes called ‘grammatical agreement’ rather than ‘anaphoric agreement’ (Siewierska 1999; Bresnan, & Mchombo 1989, Bresnan 2001). An effect of this is that the canonical type is restricted to relatively few languages, since pro-drop is common.

---

<sup>3</sup> A more formal approach to some of these issues can be found in Avgustinova & Uszkoreit (forthcoming).

*C-2: controller has overt features > controller has covert features*

Compare these French examples:

(7) elle est content-e  
she be.3.SG happy-SG.FEM  
'she is happy'

(8) je suis content / content-e  
I be.1.SG happy-SG.MASC / happy-SG.FEM  
'I am happy (of a man/of a woman)'

In (7) the controller is overtly feminine: the pronoun *elle* 'she' contrasts with *il* 'he'.

In (8) there is no distinction in the controller for 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. An example where the number of features of the controller is lower than that of the target is found in Amele (Roberts 1987: 162, 201, 203 and personal communication, Corbett 2000: 136-137), where a subject noun phrase headed by a noun need not mark number (there are optional ways of doing so), but where the predicate verb does mark number (singular/dual/plural).

On the basis of these criteria discussed (and others to be considered below), a more general principle may be suggested (compare Moravcsik 1988: 90):

*Principle-I: Canonical agreement is redundant rather than informative*

In the French *elle est contente* 'she is happy', the feminine feature is available from the controller (C-2) and so the agreement is redundant. In *je suis content(e)* 'I am happy' the agreement adds information.

More generally, the amount of information in the controller position can be expanded almost indefinitely (*the tall man, who lives at 10 Railway Cuttings and who*

... *is leaving*). The amount of *controller* information which is available in the target position is restricted to a set of feature specifications. Hence the principle above. 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.

*C-3: consistent controller > hybrid controller*

A consistent controller is one which controls a consistent agreement pattern. The notion ‘consistent agreement pattern’ is intuitively easy, but not so easy to define (for the details see Corbett 1991: 176-181). As a rough and ready characterization, a consistent agreement pattern is the set of agreements controlled by a typical regular controller. 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). A plural verb can be used with a singular noun phrase, to indicate reference to a person or persons besides the one indicated directly. We have an ‘associative’ construction, indicated not by a marker on the nominal, but by plural agreement:

- (9)    moj        brat                tam    tóža    žýl'-i<sup>4</sup>  
       my.SG   brother.SG    there   also   lived-PL  
       ‘my brother and his family also lived there’

Plural agreement is found in the verbal predicate, but not in the noun phrase, and so we have different agreements according to the target.

The possible patterns of agreement with hybrid controllers are tightly constrained by the Agreement Hierarchy (Corbett 1979, 1983, 1991: 225-260, 2000: 188-192; Cornish 1986: 203-211; Barlow 1991, 1992: 136-137; Kirby 1999: 92-96).

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

---

<sup>4</sup> Bogdanov's transcription has been transliterated here.

The idea is that given a domain, say subject-predicate agreement, in the canonical case we do not require further information on the part of speech. For instance, in Russian we do not need 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 & Hayward (1987), Corbett (2000: 181-183).

These two criteria (C-3 and C-4) fall under a second general principle:

*Principle II: Canonical agreement is syntactically simple*

## **2.2. Targets**

The largest number of criteria relates to the target. This makes sense, since it is the target which is the locus of agreement. The criteria tend to ‘cluster’, though as we shall see they can be teased apart in some languages. We begin with the nature of the expression of agreement on the target. In the most general terms, a major criterion is:

*C-5: bound > free*

We are concerned with the *expression* of agreement here. Some argue that unless the expression is bound to the target, then we are not dealing with agreement; others are more liberal. To discuss alternative stances on this, let us expand out the criterion:

*C-5': inflectional marking (affix) > clitic > free word*

The canonical expression of agreement is through inflectional morphology bound to the target. Some also treat certain uses of clitics as agreement. Halpern (1998: 105) discussing verbal clitics, states that: ‘... they are often assumed to be types of inflectional affixes themselves, perhaps simply agreement markers.’ But he adds that ‘...there are also several respects in which clitics are not like canonical agreement affixes.’ Clitic doubling constructions are particularly relevant, as in this example from Macedonian (Friedman 1993: 285):

- (10) kuče-to                                    ja                                    kasa                                    mačka-ta  
 dog-DEF.SG.NEUT    3.SG.FEM.ACC    bite-3.SG    cat-DEF.SG.FEM  
 ‘the dog bites the cat’

In (10) the clitic *ja* ‘doubles’ the noun phrase *mačka-ta* ‘the cat’; it is singular and feminine, like its controller. In such examples, where the object is definite, there must be a doubling clitic pronoun in Macedonian (Friedman 1993: 285). Clitics vary as to ‘how bound’ they are; verbal clitics are ‘more bound’ than second position clitics, and so are somewhat closer to being canonical agreement markers. For helpful analysis of clitics and their relation to agreement see Spencer (1991: 384-390); for a ‘liberal’ view of clitics as agreement see Lyons (1990).

We should now ask whether a free word can be an expression of agreement. It is important to be clear that we are looking at the expression of agreement, not just at a potential stem or host. A predicate verb is a common target: it acts as a stem (for inflectional marking) or a host (for a clitic), but is not itself the expression of agreement. (The distinction merges particularly easily with pronouns, where an anaphoric pronoun may function as such, and be a target for agreement, or may develop into a form which loses its anaphoric function and be considered, at least by some, to be entirely an expression of agreement: see Lehmann 1982: 234-241 for early discussion, and Siewierska 1999).

The most convincing examples of free words as the expression of agreement are found in Daly languages of north Australia. For instance, Ngan’gityemerri, a Daly language comprising two dialects, Ngan’gikurunggurr and Ngan’giwumirri, with 100 speakers, 300 miles SW of Darwin, Australia, has arguably 15 genders. Of these, six genders have optional free-form generics/classifiers (Reid 1997: 177):

- (11) (syiri)      magulfu                                    (syiri)      marrgu  
 STRIKE    cylindrical.fighting.stick    STRIKE    new  
 ‘a new cylindrical fighting stick’

*Syiri* is the free-form generic for weapon-like objects which have a striking type of contact. In its first use in (11) it is analogous to a classifier. Its second use is more like an agreement marker. The repetition of this free form in the noun phrase is, according to one's point of view, an example of agreement with a free word as the expression of agreement, or else a phenomenon on the edge of agreement. There is strong evidence that such free-form generics develop into agreement markers, in a system in which the generics are still feeding the gender system (Reid 1997: 211-222).

*C-6: obligatory > optional*

Canonical agreement is marked obligatorily; optional marking is less canonical. This criterion is linked to the previous one (since inflectional marking is usually obligatory), but the two can vary independently. We find optional inflectional marking of agreement, if rarely, while less canonical types of marking are more likely than inflectional marking to be optional. An example is again Ngan'gityemerri (Reid 1997). Of the 15 genders, nine are distinguished by the agreement markers found on agreement targets, such as adjectives :

- (12) a-syensyerrgimi                      a=tyentyenmuy  
ANIMATE-white.rock.wallaby ANIMATE-tame  
'a tame white rock wallaby' ('=' indicates clitics and '-' affixes)

Reid argues that the marker on the head noun is a prefix, while that on the agreement target is a proclitic, on the basis of stress and assimilation processes (1997: 212-215). The important point for us is that these agreement markers on targets are optional (1997: 168). As noted earlier, like the proclitic agreement markers, the generics/classifiers are optional.

Another instance of optional agreement is the agreement of possessive forms in Chukchee, which are formed from nouns by suffixation (Skorik 1961: 240-241; Spencer ms. §6.4.3). When functioning as an attributive, such possessives can agree in



number with the head noun, but they do so only rarely (they are more likely to take the plural marker when the noun is dropped, Skorik 1961: 233). It seems that agreement is conditioned by there being greater focus on the attributive than on the noun. These instances of optionality of agreement are less canonical than, for example, the Spanish examples (1)-(4) where agreement is obligatory.

Further examples are frequent in clitic doubling. The closely related South Slavonic languages Macedonian and Bulgarian both have clitic doubling for objects (as well as more canonical subject agreement). We have seen that Macedonian clitic doubling is obligatory under certain circumstances. Bulgarian also has clitic doubling but it is ‘generally optional’ (Scatton 1993: 234), though there are circumstances in which it is required. Thus, if clitic doubling is included as a type of agreement, the type found in Macedonian is closer to canonical agreement than that of Bulgarian.

Let us now concentrate on the morphology of the agreement marking. There are three relevant criteria here, which we consider in turn.

*C-7: regular > suppletive*

The canonical marking is by regular inflectional morphology (affixation). Perhaps surprisingly, we also find instances of agreement expressed by suppletion.<sup>5</sup> This is found in Norwegian (Tore Nessel, personal communication):

(13)	en	liten	øl	(14)	to	små	øl
	one/a	small.MASC.SG	beer.SG		two	small.PL	beer.SG/PL
	‘one small beer’				‘two small beers’		

Here number agreement is expressed through suppletion; for the majority of adjectives, of course, it is expressed regularly.<sup>6</sup>

---

<sup>5</sup> I am grateful to Nigel Vincent for alerting me to this.

<sup>6</sup> This criterion is logically independent of the others, which is worth bearing in mind below when, in the discussion of domains, we ask whether the English pronouns *he/she/it* show agreement with their

*C-8: alliterative > opaque*

This criterion is related to the last but differs from it. Consider this example from Swahili (Welmers 1973: 171):

- (15) ki-kapu ki-kubwa ki-moja ki-lianguka  
7-basket 7-large 7-one 7-fell  
'one large basket fell'

Two points about this type of agreement system deserve attention. First, the agreement marker on the target is identical to a formant of the controller. This might seem to repeat the second criterion, concerning overt marking. However, criterion C-2 is concerned with the *fact* of overt marking; such marking may or may not be alliterative. The essential point about alliterative agreement is the phonological identity (compare Aronoff 1998). The second point is that the same agreement marker is used for different agreement targets (the adjective, numeral and verb). Both of these characteristics may be found to a greater or lesser degree (see Corbett 1991: 117-119 for discussion).

*C-9: productive > sporadic*

The canonical situation is for each potential target of a given type to show agreement. Thus in Russian, every verb shows agreement in number. Compare this with the Nakh-Daghestanian languages Chechen and Ingush where only around 30% of the verbs show agreement (Bickel & Nichols forthcoming §1). An extreme case of sporadic agreement is found in Kuwaa, a Kru language (a group within Niger-

---

antecedent. A probable reason why many say this cannot be agreement, almost without thinking, is that the pronouns would then show suppletive expression of gender and number. When considering the domain question we should compare English with other languages where the suppletion issue can be factored out, for instance in Russian where the third person pronoun *on/ona/ono/oni* is not suppletive (at least in the nominative).

Kordofanian) where only one adjective retains agreement in number (Marchese 1988: 335, personal communication from R. Thompson).

These various criteria can be seen as aspects of a single 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 now go on to three criteria which concern the target from a wider perspective (and fall under the principle of syntactic simplicity). The first relates back to the earlier discussion of doubling:

*C-10: doubling > independent*

The term comes from clitic doubling, but we use it here for all types of target. The intuition is that a target shows more canonical agreement if the agreement occurs independently of the presence or absence of the controller. That is, it occurs when the controller is present, rather than only appearing when the controller is absent (cf. the discussion of Chukchee above). Naturally this criterion relates to the controller criterion C-1 “*controller present > controller absent*”. We need two criteria, in order to generalize both over types of controller and over types of target. Thus agreement markers which appear irrespective of the type of controller (and therefore ‘double’ it when present) are more canonical than those whose appearance or absence depends on the type of controller.

*C-11: target agrees with a single controller > agrees with more than one controller*

Canonically, a target has a single controller, as in examples (1)-(4). Sometimes the target may mark agreement more than once, in fact it may mark it up to four times as in Archi (Corbett 1991: 108, based on Kibrik 1977: 128-130, 320). What is less canonical is for a single target to agree with more than one different controller. An example of this is found in associative/possessive constructions in some Bantu languages. We shall take examples from Shona (Welmers 1973: 178):

- (16) *imbwa* *na-vana* *v-a-dz-o*  
dogs.10 and-young.2 2-ASSOCIATIVE-10-their  
'the dogs and their pups'

The last item shows agreement with both nouns. The head noun *imbwa* 'dogs' belongs to gender 9/10, 10 being the plural form, and the corresponding agreement marker is *-dz-*; the associated noun *vana* 'children, young' is gender 1/2 and takes the agreement marker *v-*.

*C-12: target's part of speech irrelevant > relevant (given the domain)*

The intuition here is that it is more canonical to be able to specify targets at a high level, as a general part of a domain, rather than having to make additional stipulations for sub-types. Thus we treat it as canonical to specify, for instance, that attributive modifiers agree with their head noun. Thus when we discussed Swahili (15), we noted that attributive adjective and numeral both agree. Being able to give a rule for attributive modifiers in general is a more canonical situation than that in a language where one would have to specify that certain types of attributive modifier agree while some do not. This differs from the 'productive-sporadic' criterion (C-9) in that the latter operates within a part of speech (do all adjectives behave alike?), while the current criterion (C-12) compares across parts of speech (do all targets of a particular type behave alike, irrespective of part of speech?)

### 2.3. Domains

There are few criteria here, but they are substantial. The use of the terms 'controller' and 'target' suggests that agreement is an asymmetric relation. We might treat this as a defining characteristic, or we may see it as a property of canonical agreement. If two items match for the same external reason, this is not canonical agreement. If one is in a particular form because of the properties of the first, then this is potentially canonical agreement. An analogy may be helpful. If houses numbers 10 and 12 are both white because it has snowed on both, this is not canonical agreement. If Mrs

White paints number 10 white and Mrs Green in number 12 does likewise, that is, potentially, canonical agreement.

*C-13: asymmetric > symmetric*

This criterion fits with the view that syntax is typically about hierarchical/dependency relations. If canonical agreement is typically an asymmetric relation, that leads to the problem of ‘agreement in case’. Take this common Russian phrase:

- (17) s drug-oj storon-y  
 from other-SG.FEM.GEN side-SG.GEN  
 ‘on the other hand’ (literally ‘from the other side’)

*Storon-a* ‘side’ is feminine; the agreement of the adjective in gender and number is comparable to that in our original examples (1)-(4). The adjective and noun are also in the same case. The question is whether this is also an instance of agreement.

If one adopts a view of syntax based on the notion of constituency, then examples like this, the traditional instances of ‘agreement in case’ within the noun phrase, are not canonical agreement. Case matching of features within the noun phrase results from government by the same element, in this instance the preposition *s* ‘from’, which governs the genitive. The argument holds for *Suffixaufnahme* phenomena (Plank 1995). *S drugoj storony* (!), however, for those who accept a dependency view of syntax, then the opposite conclusion follows. If the noun is the head of the phrase and the adjective depends on it, and both show case, then we would have agreement in case (as in Mel’čuk 1993: 329, 337).<sup>7</sup> We conclude that canonical agreement is asymmetric. Furthermore, which instances count as asymmetric, and therefore potentially canonical, depends on other assumptions about syntax.

Next we consider the possible structures within domains:

---

<sup>7</sup> For the practical business of providing a database for the linguistic community, we would not wish to prejudge what users’ view of syntax should be, and so data on case agreement will be included.

*C-14: local > non-local*

The question as to whether agreement is a local phenomenon is rarely asked. Those who have treated agreement as a prime focus of study, for instance Moravcsik (1978: 334) and Lehmann (1982: 211), typically assume that the feature values of anaphoric pronouns are determined by agreement mechanisms. Within the discussion of agreement they cite examples, such as *The man is in the room. He is old.* (Moravcsik 1978: 337). Those for whom agreement is a less central concern often assume that it is a local phenomenon, and so exclude examples like that just cited. The only extended discussion of the issue of which we are aware is found in Barlow (1991, 1992: 134-52) who concludes that there are no good grounds for distinguishing between agreement and antecedent-anaphor relations. This conclusion is confirmed in, for instance, Pollard & Sag (1994: 60-99) and Siewierska (1999: 225). The main types of evidence are discussed in Corbett (forthcoming a). ‘Agreement’ is used by different linguists to cover feature matching in a range of domains, from within the noun phrase to antecedent-anaphor relations. Some limit it, more or less drastically. If we are to draw a boundary, then we need to be clear whether this is based on evidence from agreement itself (which would be hard to justify), or whether the boundary is being drawn as a result of other considerations within the syntactic model adopted. If such a boundary is proposed, then we should ask whether it claims to handle the distribution of syntactic versus semantic agreement. Such a claim is unlikely to be well founded, since evidence from the Agreement Hierarchy shows that there is no one point at which agreement phenomena can be neatly divided into two. Rather there are several domains for agreement, related in hierarchical fashion. Since some reject non-local agreement, we can reasonably conclude that local agreement (within the clause) is more canonical than non-local (beyond the clause).

Under this criterion we concentrate on the syntactic position of anaphoric pronouns. However, their morphology can also vary, and in part independently of their syntax. Thus anaphoric pronouns can be morphologically free or bound (the

latter often being termed ‘pronominal affixes’ or ‘incorporated pronouns’). Pronominal affixes are less canonical in terms of their domain than say subject-verb agreement, since they are part of a non-local domain, but are more canonical than free pronouns in being morphologically bound (see Evans 1999, Mithun 1999: 189-192, 202-203 for examples and discussion).

A second type of non-local agreement is ‘long-distance’ agreement. This was originally called ‘transparent agreement’ (*prozračnoe soglasovanie*), starting from the notion of a predicate being transparent, used by Kibrik in Bergel’son, Zaliznjak & Kibrik (1982: 49). It is well attested in Nakh-Daghestanian languages, and we take data from Tsez (Polinsky & Comrie 1999: 116-117). Example (18) is the expected/canonical construction, where a sentential complement is treated as the controller of agreement, and so the agreement is in the default gender, gender IV (genders are given in Roman numerals). The experiencer argument stands in the dative, hence *eni-r* ‘mother-DAT’. (As in Polinsky & Comrie, for clarity the important agreement marker is in bold, and the complement in square brackets.)

- (18) *eni-r*            [uΩ-å    magalu        b-åc’-ru-~i]  
          mother-DAT   boy-ERG   bread.ABS.III   III-eat-PST\_PRT-NMLZ  
                          **r-iy-xo**.  
                          IV-know-PRS

‘The mother knows that the boy ate bread.’

Surprisingly, in Tsez a matrix verb can agree with a nominal in the absolutive inside the complement. In (19), the matrix verb has gender III agreement, agreeing with *magalu* ‘bread’, an absolutive phrase within the sentential complement:

- (19) eni-r            [uΩ-å      magalu            b-åc'-ru-~i]  
mother-DAT   boy-ERG   bread.ABS.III   III-eat-PST\_PRT-NMLZ  
                  b-iy-xo.  
                  III-know-PRS  
                  ‘The mother knows that the boy ate bread.’

In Tsez long-distance agreement is found ‘when the referent of the absolutive noun phrase is the main internal topic of the embedded clause’ (Polinsky & Comrie 1999: 122). While we find similar constructions elsewhere, the conditions differ from language to language. The important thing here is the nature of the domain in (19). The controller and target are more distant syntactically than we would expect, hence the name ‘long distance agreement’. Given any ‘choice’ of this type, then the more local domain is more canonical, hence long distance agreement is non-canonical.

Finally in this section we shift from looking at individual relations to looking at the system:

*C-15: domain is one of set > single domain*

In canonical instances, a given domain will be one of a set of domains (agreement with a given controller may be expressed by different targets). Thus if we take Russian subject-verb agreement, this is one domain of several (attributive modifier agreeing with head noun, relative with antecedent ...). This is a more canonical situation than that in a language where, say, there is only one agreement domain. This criterion links back to the notion of redundancy.

#### **2.4. Features**

Here we find three criteria, one relating to features as a whole, and two relating to their values.

*C-16: feature is lexical > non-lexical*

Agreement in gender (where lexical) is considered the canonical case. The reason is that the target could not be marked with the feature independently, if it is lexical, and so this links to the directionality of agreement. Another way of stating this criterion is



that features which are based at least in part on formal assignment are more canonical for agreement than features where assignment is semantically based. (This criterion falls under the principle of redundancy.) An interesting consequence concerns anaphoric pronouns; the fact that in many languages these can match according to the feature of lexical gender suggests they belong within agreement.

*C-17: features have matching values > non-matching values*

Some would claim that the definition of agreement must refer to the matching of values. However, once a construction is identified as involving agreement, because there is a matching of features, we would not want to rule out the analogous instances where the features do not match. Specifically, since English subject and predicate verb regularly have matching features, we have to address examples like this one:

(20) the committee have decided

We cannot simply say that *committee* is plural, since we do not find *\*these committee*. We need to invoke a notion of semantic agreement for such cases and, as noted earlier, the distribution of semantic agreement is tightly constrained by the Agreement Hierarchy. From the point of view of the feature values, we can say that examples like (20) are less canonical instances of agreement than those where the feature values match (*the committee has decided*); for further discussion see Corbett (2000: 188-191). Semantic agreement involving gender or number is relatively common. Examples with person are rarer, but are found, as in this Spanish type (G. Martínez Sierra, quoted by Harmer & Norton 1957: 270):

(21) ¡ Qué desgraciadas somos las mujeres !  
 how unfortunate-PL.FEM be.1.PL DEF.PL.FEM women  
 ‘how unfortunate we women are’

If we accept that semantic agreement is non-canonical, then we should include here instances of resolution rules, which specify the feature values of targets when the controller consists of conjoined noun phrases (Corbett forthcoming b), as in this example from Slovene (Priestly 1993: 433):

- (22) Milka in njeno tele sta bil-a zunaj  
Milka.FEM and her calf.NEUT AUX.DU been-DU.MASC outside  
'Milka and her calf were outside'

Here we have a feminine singular and a neuter singular conjoined; the verb is dual and masculine. Clearly the features do not match. It is the resolution rules which specify these particular feature values (for details see Corbett 1991: 261-306). The fact that such instances are taken to be non-canonical fits with the discussion in Corbett (forthcoming b) where the peripheral nature of such rules is demonstrated.

The general effect of this criterion is to claim that syntactic (formal) agreement is more canonical than semantic agreement. This has the interesting effect that unification is an adequate mechanism for formalizing canonical instances of agreement. This demonstrates well that this criterion falls under the principle of syntactic simplicity. The criterion is also consonant with the 'redundancy' principle. And if we extend the notion somewhat, we might say that what Bickel (2000) calls 'integrative' systems of agreement are more canonical than 'associative' systems.

*C-18: no choice of feature value > choice of value*

In sentences such as the following, English allows no choice of form:

- (23) The five applicants arrive tomorrow.

Similarly in Hungarian, predicate agreement with numeral phrases does not allow an option. The form, however, differs from that of English (Edith Moravcsik and Peter Barta, personal communications):

- (24) Két lány beszélget  
 two girl.SG chat.SG  
 ‘Two girls are chatting.’

The Russian situation is more complex; for details see Corbett (1998) and references there. Let us take just one type: here both possibilities are fully acceptable:

- (25) vošl-o / vošl-i pjat' devušek  
 came.in-SG.NEUT / came.in-PL five.NOM girl.PL.GEN  
 ‘five girls came in’

The choice is affected by the animacy of the subject, and its position relative to the predicate. For a set of the quantifiers, all relevant examples in a corpus of texts from the 19th and 20th centuries were counted (details in Corbett 1983: 150-153); there were 235 relevant examples, of which 54% showed singular agreement and 46% plural agreement. The situation found in English and Hungarian (no choice of feature value) is considered more canonical than that found in Russian (choice of value).

This criterion links to the last, but is distinct from it. While choices typically involve semantic agreement, semantic agreement may or may not involve an agreement choice for a particular target. For instance, in the example

- (26) this man and woman have travelled all day to meet you

The use of *have*, the result of the resolution rule, is an instance of semantic agreement, but is obligatory (at least for some speakers).

## 2.5. Conditions

Conditions deserve further investigation. However, the main criterion is straightforward and intuitive:

*C-19: no conditions > conditions*

In Hungarian (Edith Moravcsik, personal communication) if the subject consists of conjoined inanimate singulars, the verb will be in the singular:

- (27) A könyv és a kommentár megérkezett / \*megérkezt-ek.  
ART book and ART commentary arrived.SG / arrived-PL  
'The book and the commentary arrived.'

However, if the conjuncts denote animates, the verb may take the singular or the plural, with the plural preferred:

- (28) John és Jill megérkezt-ek / megérkezett.  
John and Jill arrived-PL arrived.SG  
'John and Jill arrived.'

Thus the possibility of number resolution, resulting in the plural, is dependent on a condition, namely that the conjuncts denote animates. This is less canonical than a construction in which there is no such condition.

Agreement conditions are particularly prevalent when agreement is non-canonical in some other way. In the example from Hungarian, the condition interacts with an agreement choice, itself a non-canonical feature.

### **3. THREE GENERAL PRINCIPLES**

Three general principles were introduced earlier, and deserve brief discussion here.

*Principle-I: Canonical agreement is redundant rather than informative*

Several separate criteria (numbers 1, 2, 16, 17 and 18, and secondarily number 15) converge on this principle. It may well be that it is this principle which leads to canonical agreement being relatively rare among the world's languages.

As a partial restating of this principle, we might add that the greater the reliance on formal properties the more canonical. This view of it is best seen by

imagining its opposite. If we had fully semantic agreement, then it would hardly exist as a distinct phenomenon, since all the forms could be predicted directly from semantics; the matching effect would arise from controller and target corresponding simply by virtue of having a common semantic source. It is in the converse case - agreement in gender in instances where the gender is not assigned by a semantic rule - that we most evidently require special rules of agreement.

*Principle II: Canonical agreement is syntactically simple*

This principle is a generalization of criteria numbers 3, 4, 10, 11, 12, 13, 14, 15, 17, 18, 19 and partially of number 6. It is reflected in criteria relating to each aspect of agreement (controller, target and so on).

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

Several separate criteria converge on this principle, namely numbers 5, 6, 7, 8 and 9. Note that they are all in the target group. There are different views as to which target types are legitimately considered to be a part of agreement, but surely no-one would exclude the type of targets with canonical inflectional morphology found in Russian or German from an account of agreement.

There is a certain irony here. We claim that when the expression of agreement is closest to canonical inflection, it is there that we find canonical agreement. Furthermore Booij, who divides inflectional morphology into two parts, inherent and contextual (with agreement markers and structural case markers being good instances of the latter) calls contextual inflection ‘the prototypical case of inflection’ (1996: 14). And yet agreement poses substantial problems for theories of morphology.

Two observations are offered towards understanding this apparent conundrum. On the one hand, agreement is a major interface problem between morphology and syntax, and hence appears difficult when viewed from the heartland of either component. And second, the external pressure on agreement targets from the general

syntactic rules of agreement inhibits specifically lexical effects (such as suppletion), and thus helps to maintain canonical inflection for the expression of agreement.

#### 4. CONCLUSION

We have attempted to clarify some of the conceptual problems that characterize this area. We have seen how different properties cluster, which makes it particularly important that we specify which properties are the basis for our analytical decisions. We noted the position of morphology at the core of the canonical instances of agreement. Having seen the gradient nature of many of the properties, 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. For practical purposes, specifically the construction of a typological database, we do need to draw boundaries. Here the results of this paper are of use since we need to be clear where and why such a line is drawn, so that users of the database can be aware how the data relate to their own conceptions and analyses of the area.

#### References

- Aronoff, Mark. 1998. Gender agreement as morphology. In: Geert Booij, Angela Ralli & Sergio Scalise (eds) *Proceedings of the First Mediterranean Conference of Morphology*, 7-18. Patras: University of Patras.
- Avgustinova, Tania & Uszkoreit, Hans. forthcoming. Towards a typology of agreement phenomena. To appear in: William Griffin (ed.) *“The Role of Agreement in Natural Language”*: *Proceedings of the 2001 Texas Linguistic Society Conference, Austin, Texas, 2-4 March 2001*.
- Barlow, Michael. 1991. The Agreement Hierarchy and grammatical theory. In: Laurel A. Sutton, Christopher Johnson & Ruth Shields (eds) *Proceedings of the Seventeenth Annual Meeting of the Berkeley Linguistics Society, February 15-18, 1991*, 30-40. Berkeley: Berkeley Linguistics Society, University of California.
- Barlow, Michael. 1992. *A Situated Theory of Agreement*. New York: Garland.

- Bergel'son, Mira B., Zaliznjak, Anna A. & Kibrik, Aleksandr E. 1982. Konstrukcii s sentencial'nym aktanktom v tabasaranskom jazyke. In: *Tabasaranskije ètjudy: Materialy Dagestanskoj èkspedicii 1979* (Publikacii otdelenija strukturnoj i prikladnoj lingvistiki 15, series editor V. A. Zvegincev), 44-65. Moscow: Izdatel'stvo Moskovskogo universiteta.
- Bickel, Balthasar. 2000. On the syntax of agreement in Tibeto-Burman. *Studies in Language* 24.583-609.
- Bickel, Balthasar & Nichols, Johanna. forthcoming. Inflectional morphology. To appear in: Timothy Shopen (ed.) *Language Typology and Syntactic Description*, second edition. Cambridge: Cambridge University Press.
- Bogdanov, V. N. 1968. Osobyj slučaj dialektnogo soglasovanija skazuemogo s podležaščim po smyslu i kategorija predstavitel'nosti. *Naučnye doklady vyššej školy: filologičeskie nauki* no. 4. 68-75.
- Booij, Geert. 1996. Inherent versus contextual inflection and the split morphology hypothesis. In: Geert Booij & Jaap van Marle (eds) *Yearbook of Morphology 1995*, 1-15. Dordrecht: Kluwer.
- Bresnan, Joan & Mchombo, Sam. 1989. Topic, Pronoun, and Agreement in Chichewa. *Language* 63.741-782.
- Bresnan, Joan. 2001. *Lexical-functional syntax*. Oxford: Blackwell.
- Corbett, Greville G. 1979. The Agreement Hierarchy. *Journal of Linguistics* 15.203-224.
- Corbett, Greville G. 1983. *Hierarchies, Targets and Controllers: Agreement Patterns in Slavic*. London: Croom Helm.
- Corbett, Greville G. 1991. *Gender*. Cambridge: Cambridge University Press.
- Corbett, Greville G. 1998. Agreement in Slavic. WWW document prepared for the workshop "Morphosyntax in Slavic" available at: <http://www.indiana.edu/~slavconf/linguistics/index.html> [January 1998]: to appear in the *Proceedings*, ed. George Fowler, Bloomington, Indiana: Slavica.
- Corbett, Greville G. 2000. *Number*. Cambridge: Cambridge University Press.

- Corbett, Greville G. forthcoming a. Agreement: Terms and boundaries. To appear in: William Griffin (ed.) *“The Role of Agreement in Natural Language”*: *Proceedings of the 2001 Texas Linguistic Society Conference, Austin, Texas, 2-4 March 2001*.
- Corbett, Greville G. forthcoming b. Types of typology, illustrated from gender systems. To appear in: Frans Plank (ed.) *Noun Phrase Structure in the Languages of Europe* (EUROTYP volume 7). Berlin: Mouton de Gruyter.
- Corbett, Greville G. & Hayward, Richard J. 1987. Gender and number in Bayso. *Lingua* 73.1-28.
- Cornish, Francis 1986. *Anaphoric relations in English and French: A Discourse Perspective*. London: Croom Helm.
- Evans, Nicholas 1999. Why argument affixes in polysynthetic languages are not pronouns: evidence from Bininj Gun-wok. *Sprachtypologie und Universalienforschung* 52.255-281.
- Friedman, Victor A. 1993. Macedonian. In: Bernard Comrie & Greville G. Corbett (eds) *The Slavonic Languages*, 249-305. London: Routledge.
- Halpern, Aaron L. 1998. Clitics. In: Andrew Spencer & Arnold Zwicky (eds) *The Handbook of Morphology*, 101-122. Oxford: Blackwell.
- Harmer, L. C. & Norton, F. J. 1957. *A Manual of Modern Spanish*. London: University Tutorial Press. [2nd edition, 1st edition 1935]
- 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.
- Kibrik, Aleksandr E. 1977. *Opyt strukturnogo opisanija arčinskogo jazyka: II: Taksonomičeskaja grammatika*. (Publikacii otdelenija strukturnoj i prikladnoj lingvistiki 12). Moscow: Izdatel'stvo Moskovskogo universiteta.
- Kirby, Simon 1999. *Function, Selection and Innateness: the Emergence of Language Universals*. Oxford: Oxford University Press.



- Lehmann, Christian 1982. Universal and typological aspects of agreement. In: Hansjakob Seiler & Franz J. Stachowiak (eds) *Apprehension: Das sprachliche Erfassen von Gegenständen: II: Die Techniken und ihr Zusammenhang in Einzelsprachen*, 201-67. Tübingen: Narr.
- Lyons, Christopher 1990. An agreement approach to clitic doubling. *Transactions of the Philological Society* 88.1-57.
- Marchese, Lynell. 1988. Noun classes and agreement systems in Kru: a historical approach. In: Michael Barlow & Charles A. Ferguson (eds) *Agreement in Natural Language: Approaches, Theories, Descriptions*, 323-341. Stanford: CSLI.
- Mel'čuk, Igor 1993. Agreement, government, congruence. *Linguisticae Investigationes* 17.307-372.
- Mithun, Marianne. 1999. *The Languages of Native North America*. Cambridge: Cambridge University Press.
- Moravcsik, Edith A. 1978. Agreement. In: Joseph H. Greenberg, Charles A. Ferguson & Edith A. Moravcsik (eds) *Universals of Human Language: IV: Syntax*, 331-374. Stanford: Stanford University Press.
- Moravcsik, Edith A. 1988. Agreement and markedness. In: Michael Barlow & Charles A. Ferguson (eds) *Agreement in Natural Language: Approaches, Theories, Descriptions*, 89-106. Stanford: CSLI.
- Plank, Frans (ed.) 1995. *Double Case: Agreement by Suffixaufnahme*. New York: Oxford University Press.
- Polinsky, Maria & Comrie, Bernard. 1999. Agreement in Tsez. In: Greville G. Corbett (ed.) *Agreement* (Special issue of *Folia Linguistica* XXXIII/2), 109-130.
- Pollard, Carl & Sag, Ivan A. 1994. *Head-driven Phrase Structure Grammar*. Chicago: University of Chicago Press.
- Priestly, T. M. S. 1993. Slovene. In: Bernard Comrie & Greville G. Corbett (eds) *The Slavonic Languages*, 388- 51. London: Routledge.

- Reid, Nicholas. 1997. Class and classifier in Ngan'gityemmeri. In: Mark Harvey & Nicholas Reid (eds) *Nominal Classification in Aboriginal Australia* (Studies in language companion series 37), 165-228. Amsterdam: John Benjamins.
- Roberts, John R. 1987. *Amele*. London: Croom Helm.
- Scatton, Ernest A. 1993. Bulgarian. In: Bernard Comrie & Greville G. Corbett (eds) *The Slavonic Languages*, 188-248. 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.
- Skorik, Petr Ja. 1961. *Grammatika čukotskogo jazyka: I: Fonetika i morfologija imennyx častej reči*. Moscow: Izd. Akademii Nauk SSSR.
- Spencer, Andrew. 1991. *Morphological Theory: An Introduction to Word Structure in Generative Grammar*. Oxford: Blackwells.
- Spencer, Andrew ms. The typology of morphosyntax (LG615). Unpublished Course Notes, University of Essex. <http://privatewww.essex.ac.uk/~spena/>
- Steele, Susan. 1978. Word order variation: a typological study In: Joseph H. Greenberg, Charles A. Ferguson & Edith A. Moravcsik (eds) *Universals of Human Language: IV: Syntax*, 585-623. Stanford: Stanford University Press.
- Welmers, William E. 1973. *African Language Structures*. Berkeley: University of California Press.