

Minor number and the plurality split

Greville G. Corbett

When plurality is grammatically relevant for some pronouns and nouns but not others, the division will be in accord with the Smith-Stark Hierarchy, as shown by Smith-Stark (1974). Several languages have an additional number, a 'minor number', which involves a relatively small proportion of the nouns of the language. Data on these minor numbers (paucal, dual, collective, general, mass) are taken from Avar, Budugh, Italian, Hebrew, Maltese, Fula and Spanish dialects. The data show clearly that these minor numbers are not subject to the Smith-Stark Hierarchy. However, data from Maltese and from Spanish dialects show that minor numbers are subject to the Agreement Hierarchy.

1. Introduction¹

Smith-Stark (1974) demonstrated that when plurality 'splits a language', that is to say where it is grammatically relevant for some pronouns and nouns but not others, this split will be in accord with an Animacy or Topicality Hierarchy, which we shall refer to here as the Smith-Stark Hierarchy. However, going beyond a simple singular-plural distinction, we find that several languages have in addition a 'minor number', that is to say a number which involves a relatively small proportion of the nouns of the language. The question then arises as to whether these minor numbers are also subject to the Smith-Stark Hierarchy.

We first set out Smith-Stark's claim briefly (section 2). Then we investigate several instances of minor number, some of which are of wider significance for a general typology of number (so extending the typology in Corbett 1992). We look first at languages where there is only morphological evidence (section 3), then at those where there is syntactic evidence too (section 4), and we establish that minor numbers are not subject to the Smith-Stark Hierarchy. Maltese turns out to be of particular interest since it has two minor numbers. We then look at the remnants left by a minor number, after it has ceased to function as a number (section 5). We go on to consider the constraints on minor number, and the question of minor categories more generally (section 6) before drawing our conclusions (section 7).

2. The plurality split

Smith-Stark considered that plurality 'splits' a language if 'it is a significant opposition for certain categories but irrelevant for others' (1974:657). The type of evidence he produced concerned marking of the noun phrase for number (usually by marking on the noun itself) and agreement in number (mainly verbal agreement but with some instances of agreement within the noun phrase). He claims, for instance, that if in Georgian the subject is plural and denotes an animate the verb will be plural, if it denotes an inanimate then the verb will be singular. Thus Georgian nouns are split, and the division is between animates and inanimates. Various languages make the split at different points, according to a hierarchy which can be presented as follows:

speaker > addressee > kin > rational > human > animate > inanimate
 (1st person (2nd person
 pronouns) pronouns)

This is clearly akin to what in other publications has been termed the Animacy Hierarchy. Smith-Stark provides a good deal of data to support his claim, and notes some problematic cases too.²

Smith-Stark considers numbers apart from the plural only briefly; he expects the dual to pattern as the plural, but notes a potential problem (1974:669n6). It is therefore worth investigating other numbers, and in particular, what we have termed 'minor numbers' – those which involve a relatively small proportion of the nouns of a given language. We consider first those where marking on the noun is involved, that is to say the cases where the evidence is morphological, and next move on to agreement, the instances where the evidence is syntactic.

3. Minor number: morphological evidence

In this section we examine minor numbers from different language families where the evidence for the additional number comes from marking on the noun itself.

3.1. The paucal in Avar

The North-East Caucasian language Avar has singular and plural. In addition, for a limited number of nouns, there is a paucal opposed to the plural. The paucal is used when the number of referents is restricted ('a few'), the unrestricted plural for larger numbers ('many'). (Sulejmanov

1985, from whom the data are taken, calls them 'restricted plural' and 'unrestricted plural'.) Examples include the following:

| singular | paucal | plural | gloss |
|----------|----------|----------|-------------------|
| nus | nús-al | nus-ábi | 'daughter-in-law' |
| boróq | boróq-al | bórq-al | 'snake' |
| t'ut' | t'út'-al | t'ut'-ál | 'fly' |
| kután | kután-al | kútn-al | 'plough' |
| bel | bél-al | búl-dul | 'spade' |
| žul | žul-al | žul-ál | 'brush' |

Sometimes, as in the last example, the difference between paucal and plural is marked only by the position of the stress. Sulejmanov says only that this three-way opposition is available for a restricted group of nouns; he lists 89 which have the contrast. Of these, only one is a kin term (*nus* 'daughter-in-law') and a further eight denote non-human animates. It appears then the nouns which have a distinct paucal form are not in general high on the Smith-Stark Hierarchy.

3.2. The collective plural in Budugh

In a second North-East Caucasian language, Budugh, just five nouns distinguish two plural forms, as follows (Kibrik 1992:15):³

| collective plural | ordinary plural |
|-----------------------------------|---|
| t'il-iber 'fingers (of one hand)' | t'il-imber 'fingers (of several hands)' |
| šül-über 'eyes (of one person)' | šül-ümer 'eyes (of several people)' |
| ibr-imer 'ears (of one person)' | ibr-imber 'ears (of several people)' |
| č'er-iber 'hair (of one person)' | č'er-imber 'hair (of several people)' |
| čärX-imer 'wheels (of one car)' | čärX-imber 'wheels (in general)' |

It can be seen that the collective form, when available, is used for referents which in some sense belong together. All the examples denote inanimates. Again we are dealing with a minor number which is not constrained by the Smith-Stark Hierarchy.

3.3. The collective plural in Italian

In Italian too we find the remnants of a collective plural (I am grateful to Marina Chini and Davide Ricca for the data). This arose in the transition from the three-gender system of Latin to the two-gender system of Italian: some neuter nouns preserved their old plural in *-a* as

a collective, and gained a new regular plural in *-i* as well. The opposition was productive for a time, and new nouns which had not been neuter were added to the group (like the Latin masculine noun *digitus* 'finger', Italian *dito*, Vincent (1988a:44; 1988b:289)). Now the opposition is no longer productive and only a very few nouns maintain the distinction, for example:

| | | |
|----------------------|--------------------------|----------------------------|
| <i>singular</i> | <i>collective plural</i> | <i>distributive plural</i> |
| l'osso 'bone' | le ossa | gli ossi |
| il lenzuolo 'sheet' | le lenzuola | i lenzuoli |
| (masculine singular) | (feminine plural) | (masculine plural) |

All the nouns of this type are masculine in the singular (the variations in the form of the definite article, like *l' / il* are not relevant for our purpose). Forms like *le ossa* 'bones' are, however, feminine in the plural. *Le ossa* is used of bones which belong together, a particular person's bones – a skeleton. *Gli ossi* is for bones which do not belong together: it would be used, for instance, when buying bones for a dog. Similarly *le lenzuola* would indicate a pair of sheets *i lenzuoli* several individual sheets.

The situation is in flux, and speakers differ in their assessment of particular forms. Nevertheless the different ways in which such a collective plural may be lost can be identified.⁴ One obvious possibility is for the collective plural to disappear, leaving the noun with the normal singular/plural opposition of the vast majority of nouns. A similar end result can come about from the specialization of the collective form. Thus *il muro* 'wall' has the regular plural *i muri* 'walls'. The old collective form *le mura* also exists, but it tends to be used for 'town walls'; this is moving towards being a related but separate lexical item; however, its use as a collective of *muro* is not yet lost. Another possibility is for the collective to be reinterpreted as a singular (*-a* is a common singular ending, the unmarked one for feminine nouns). From the triplet *il frutto* 'fruit', *le frutta*, *i frutti*, the original collective *frutta* may still be used as such in high style, but it has also been reinterpreted as a feminine singular: *la frutta*.

The second main type of development is for what was the ordinary plural to be lost, leaving the original collective plural as the only plural form. This appears to be happening with *il dito* 'finger'. 'Fingers' would normally be *le dita*: this can be used for a set of fingers or more than one set of fingers. There are few occasions on which it is necessary to refer to fingers not in sets. But in talking of, say, little fingers as opposed to index fingers some speakers would accept the normal plural *diti mignoli*

'little fingers'. Some nouns have reached the point at which only the originally collective plural is available. Thus *l'uovo* 'egg', *le uova* 'eggs' (**gli uovi*), and *il miglio* 'miles', *le miglia* 'miles' (**i migli*). From the point of view of number, these latter nouns no longer show any special characteristics: they have singular and plural like any typical noun. However, their special path of development means that they are unusual from the point of view of gender, being masculine in the singular and feminine in the plural. They form an inqorate gender (Corbett 1991:170-5). What is of greatest importance for the present argument, however, is that the examples quoted as having three forms available in Modern Italian denote inanimates.

3.4. *The dual in Modern Hebrew*

Smith-Stark's concern is the plural, but as mentioned earlier he considers whether the hierarchy would also constrain the dual (1974: 669n6): 'Although I am not addressing that problem here, my first guess would be that the dual will also split along the same hierarchy as the plural. Bob Hoberman informs me however that such is not the case in Hebrew.'

It is the case that there are languages which have a dual which parallels the plural, and splits the pronouns and nouns in a similar way (Slovene appears to be such a language). We need to find a language in which plural and dual split the nouns differently, to see whether the dual is also constrained by the Smith-Stark Hierarchy. Modern Hebrew is such a language and, as we shall see, the dual is not constrained by the Smith-Stark Hierarchy. (The fact that the dual does not pattern like the plural in several languages was noted by Forchheimer (1953:17-19). For a much fuller survey see Plank (1989:296-312)).

In Modern Hebrew the dual is restricted in the number of nouns for which it is normally available. Tobin (1988; 1990:100-50) gives a substantial list with over 100 items, though some of them appear to be fixed idiomatic expressions. There is a second restriction on the dual by comparison to the plural: it is found only in noun morphology, and not in the verb, as the following examples show (David Gil p.c.):

| | | | |
|-----|--------------|-------------------------------|---------|
| (1) | ha-yom | ʕavar | mahar |
| | DEF-day | pass.PAST.3.SG.MASC | quickly |
| | | 'the day passed quickly' | |
| (2) | ha-yom-ayim | ʕavru | mahar |
| | DEF-day-DUAL | pass.PAST.3.PL | quickly |
| | | 'the two days passed quickly' | |

- (3) ha-yam-im ɣavru maher
 DEF-day-PL pass.PAST.3.PL quickly
 'the days passed quickly'

For agreement with a controller headed by a noun in the dual, the plural is used, as in (2). (Note that in (3), plurality is marked both by form of the stem and by the ending.)

If we return now to the type of noun involved, we find various types of measure (including measures of time), for instance *ħodšayim* 'two months', many paired body parts (*yadayim* 'hands'), paired items of clothing (*magafayim* 'boots') and certain items perceived as pairs (these are *dualia tantum*, for instance *ofanayim* 'bicycle'). Note, however, that within these semantic groupings, not all the nouns have a dual form. Almost all nouns which have a dual denote inanimates, but there is also the archaic *parvayim* 'two oxen'.

3.5. The dual in Maltese

Maltese too has the dual as a minor number, and one which is on the verge of being lost. Fenech (this volume) lists just over 30 nouns which have a dual (as distinct from the plural: other nouns preserve dual morphology but retain it in place of the plural).⁵ He groups them into expressions of time, number, old Maltese weights and measures (including coins), and some food items and miscellaneous familiar objects. Of these the dual is obligatory for seven nouns only – all denoting time and number.

Thus the Hebrew and Maltese data show that is not just less usual numbers, such as collective plurals, which fail to conform to the Smith-Stark Hierarchy. When the dual is a minor number, as in Modern Hebrew and in Maltese, it too fails to conform. Plank suggests (1989:309) that "If in any language some nouns are eligible for dual marking while others are not (or less readily), the criterion is whether or not they denote natural pairs". However, there are languages in which the 'duality split' is aligned with the plurality split, and so Plank adds the following condition (1989:310): "unless the criterion is the same as that determining the eligibility of nouns for other number differentiations". While this combined claim may hold for languages in which the dual is a major number, it clearly does not hold for Modern Hebrew or Maltese. In neither language is the criterion the same for the dual as for other numbers (and so the second part of the criterion cannot apply). But neither is it a matter of natural pairs – in neither language is it the case that the nouns which have a dual are equivalent to all and only the nouns denoting natural pairs.⁶

3.6. General number in Fula (Fouta Jalon dialect)

In the Fouta Jalon dialect of Fula, which has over two million speakers in Guinea (and over 100 000 more in Sierra Leone and Senegal, there is a minor number which is not found in other dialects of Fula. Fula has an exceptionally large gender (noun class) system reflected in various agreeing elements; nouns also typically carry a marker which indicates gender and number: *caa-ngol/can-di* 'river/s' (-*ngol* shows singular number and membership of a particular gender; -*di* is one of the plural markers - the relations between the singular and plural markers are complex, and need not detain us here). Another example would be *gabb-ii/gabb-i* 'hippopotamus/es' (Koval' 1979:39; Evans 1994). In most instances, when a noun is used the speaker is required to indicate singular or plural number, since one or other marker is required. But some nouns have a third form, which does not have such a marker (Koval' 1979:11; and personal communication):

| general | singular | plural |
|-------------------|--------------------|------------------------------|
| toti 'toad(s)' | totii-ru 'toad' | totii-ji 'toads' |
| nyaari 'cat(s)' | nyaarii-ru 'cat' | nyaarii-ji 'cats' |
| gerto 'hen(s)' | gerto-gal 'hen' | gertoode 'hens' ⁵ |
| boofoo 'egg(s)' | woofoo-nde 'egg' | boofoo-de 'eggs' |
| biini 'bottle(s)' | biinii-ri 'bottle' | biinii-ji 'bottles' |

The forms which have no suffix express general meaning, as discussed in Corbett (1992), that is they are used when number is irrelevant, for instance (Koval' 1979:11):

- (4) ko biini tun waawi marde beere
 PARTICLE bottle only can-PERFECTIVE preserve beer
 'only a bottle/bottles can preserve beer'.

Various nouns are able to show general number; as our list indicates, those denoting animals are well represented, and the list could be extended to include nouns denoting humans (Koval' 1979:22). Barrie Evans (personal communication) gives interesting statistics on the availability of general forms; his database includes 180 items with a general form (11.5% of all nominals in the database). They are spread across about half the genders (noun classes); of the human class, around 30% have general forms; thus 70% do not, yet many nouns lower on the hierarchy do have general forms. This distribution is clearly not in conformity with the Smith-Stark Hierarchy. In addition, there is the interesting restriction that the form without the suffix must have at

least two syllables. Compare the following examples (Koval' 1979:12, 22):

- (5) nyaari pefay
 cat(s) scratch
 'a cat scratches/cats scratch'

Here the unsuffixed form, expressing general number meaning, is used (the singular, as noted above, would be *nyaarii-ru* and the plural *nyaarii-ji*). This is not possible in the next example:

- (6) pucc-u latay
 horse.SG kick
 'a horse kicks/horses kick'

Here the unsuffixed form would consist of a single syllable, and so the singular must be used. A further limitation on the use of forms with no agreement is that they are usually restricted to contexts in which no agreement is required (in exceptional cases where agreement is found, it is singular, according to Koval' (1979:12-13); however, Evans (1994) claims that both singular and plural agreements are possible. Note that in this dialect of Fula, number is not marked on the verb, thus in (5) and (6) the verb gives no information on number. Thus Fula does not conform to the Smith-Stark Hierarchy in terms of the nouns which have a third number form (general number).

4. Minor number: morphological and syntactic evidence

We now turn to languages where there is syntactic evidence (agreement) for minor number. As we shall see, in both languages involved there is also morphological evidence.

4.1. Mass number in the dialects of northwest Spain

In various Spanish dialects, we find a distinct mass number. We take examples from the Lena dialect (spoken in the Asturias region of northwestern Spain), closely following Hualde (1992). Hualde in turn makes extensive use of Neira (1955; 1982). For further discussion and sources see Ojeda (1989), who gives extensive data to back his theoretical account of the possible systems (the data are also available in Ojeda (1992a)), Penny (1970), and Hall (1968), who also considers comparable data from Central-South Italian dialects, and Harris (1992:82-4).

In Standard Spanish we find a straightforward singular-plural opposition of number; furthermore nouns are divided into two genders, masculine and feminine. In the Lena dialect, by contrast, some nouns have three forms: in addition to *pílu* 'hair' (singular) and *pélos* 'hair' (plural) we also find *pélo* 'hair' (mass). Which nouns have this third form? There appear to be two restrictions. The first is semantic, the noun must have a possible mass interpretation. The second restriction is morphological: only masculine nouns ending in *-u* have a distinct mass form. Thus the masculine noun *kafé* 'coffee' can make no such distinction, nor can the feminine noun *boróna* 'cornbread'.

In terms of noun morphology, then, this is another example of a minor number whose distribution is not constrained by the Smith-Stark Hierarchy. However, the data from the Lena dialect are much more interesting than the purely morphological data suggest. It turns out that the opposition between count and mass is indicated for other types of noun too, but by agreement. Consider the following (Hualde 1992:108):

- (7) la maéra tába sék-o
 the wood was dry-MASS
 'the wood (mass) was dry'

- (8) la maéra tába sék-a
 the wood was dry-FEM.SG
 'the (piece of) wood was dry'.

Thus though this particular noun may not itself show a distinct form, the count-mass distinction may be indicated by agreement. Adjectives, for instance, may have a separate form. Besides nígru 'black (MASC SG)', and *négra* 'black (FEM SG)', there is a form *négro* 'black (MASS)' which can be used equally of masculine and feminine nouns:

- (9) el kafé négr-o
 the coffee black-MASS
 'the black coffee'

- (10) la boróna négr-o
 the cornbread black-MASS
 'the black cornbread'.

This minor number is particularly interesting. Recall that in the other languages which had a minor number we found a sub-division within the plural. But these examples from Spanish dialects show that we are dealing with a distinction which is not within the plural number: the articles *el* and *la* are singular.

The article deserves further attention: in (9) and (10) we have masculine and feminine singular articles respectively, even though the adjective shows the mass form. Similarly in example (7) the article was feminine singular, despite the presence of an adjective in the mass form.⁸ The question then is where gender agreement is possible with mass nouns. According to Hualde (1992:109) the mass feature "will determine agreement with certain elements (adjectives outside the noun phrase or to the right of the noun in the noun phrase and clitics) and the gender feature will control the agreement of other elements (determiners and pronominal adjectives)". It appears then that gender agreement with mass nouns is found only within the noun phrase. This is fully in accord with the Agreement Hierarchy (Corbett 1979), which constrains the distribution of semantic and syntactic agreement. The Agreement Hierarchy distinguishes four types of agreement target:

attributive < predicate < relative pronoun < personal pronoun

Possible agreement patterns are constrained as follows:

As we move rightwards along the hierarchy, the likelihood of semantic agreement will increase monotonically (that is, with no intervening decrease).

In the case of the Lena dialect, mass agreement represents semantic agreement and gender agreement (of mass nouns) is an instance of syntactic agreement. This latter is found, it appears, only within the noun phrase (attributive agreement).

However, there is a further restriction in that even within the noun phrase semantic (mass) agreement is found, as in examples (9) and (10). We should determine whether the split is determined by the category of the agreeing element or by its position. It turns out, in fact, that the elements within the noun phrase which precede the noun take syntactic agreement, those which follow take semantic agreement. This can be demonstrated by examples with an adjective preceding the noun:

- (11) bwén-a šente
 good-FEM.SG people
 'good people'
- (12) šente bwén-o
 people good-MASS
 'good people'.

Where within the noun phrase the adjective precedes the noun, it

takes the syntactic (gender) form, where it follows it takes the semantic (mass) form. This too follows a general pattern. Semantic agreement is more likely in targets which follow the controller than in those which precede. Where there is a split within a position on the Agreement Hierarchy (the attributive position in this instance) then the distribution of forms can be determined by word order (Corbett 1979:218-20).

Returning to our main theme of minor number, we see that the mass number in Spanish dialects like the Lena dialect is not constrained by the Smith-Stark Hierarchy. In terms of noun marking it is semantically and morphologically constrained. In terms of agreement, the morphological constraint is removed. But the semantic restriction is quite clear: animates may indeed be involved (as in example (12)) but the Smith-Stark Hierarchy is clearly not what determines which nouns can have mass agreement and which not.

4.2. The collective in Maltese

In Maltese, a typical noun has singular and plural, as in the case of *raġel* 'man', *raġiel* 'men'. As we have already seen, some nouns also have a dual. There is a second minor number in Maltese, often called the collective (see Sutcliffe 1936:30-1, 36-7;⁹ Aquilina 1965:71-3; Borg 1981:15-16, 106-9; and especially Mifsud, this volume). This collective form is available for a sizable minority of nouns; nouns denoting fruits and vegetables are well represented; also included are some smaller animates, particularly insects (but never persons). The category still gains new members, through borrowings. The mixture of animate and inanimate nouns involved clearly does not conform to the Smith-Stark Hierarchy. There is a further restriction in that the collective form typically ends in a consonant, and its singular is formed by the affixation of *-a*. Thus we have collective *dubbien* 'flies', singular *dubbiena* 'fly'.¹⁰

How does such a case differ from a normal singular and plural pair? In Maltese the numerals 2-10 normally take a noun in the plural, while those from 11 upwards (like the numeral 1) take a singular. The singular *dubbiena* 'fly' occurs as expected with the numerals 11 and upwards, but the collective may not appear with the numerals 2-10. Nouns which have a collective have an additional form, the 'determinate plural' (*dubbiniet*), which is used for this purpose, as in (14) below. This resulting morphological pattern of three forms distinguishes nouns with a collective.¹¹

A further distinguishing feature is that the collective is not a straightforward plural. Depending on the 'number preference' of the noun (see Gil, this volume) it may have different interpretations. In the case of *dubbien*, the natural interpretation is plural 'flies', but the use of

the form can cover instances, if rarely, where only a single fly is involved.¹²

So far we have established that the collective is a minor number in Maltese in that a minority of nouns have a third number form rather than the basic singular and plural forms, and that the distribution of items which have the collective is not constrained by the Smith-Stark Hierarchy. These nouns are also distinguished by the agreements they take. Collective nouns, including those whose number preference is plural (as with *dubbien* 'flies'), usually take singular agreement (see Fabri 1993:86-8). The following data and acceptability judgements are from Manwel Mifsud (there appears to be wide variation in this area). Consider first the contrast between the agreements with the three forms of the noun:

Singular

- (13) Dik id-dubbiena l-kbira dahlet mit-tieqa
 that.F.SG the-fly the-large.F.SG entered.F.SG from.the-window
 'That large fly came in through the window.'

Determinate plural

- (14) Dawk il-hames dubbiniet kbar¹³ dahlu mit-tieqa
 those.PL the-five flies large.PL entered.PL from.the-window
 'Those five large flies came in through the window.'

Collective

- (15) Dak id-dubbien il-kbir dahal mit-tieqa
 that.M.SG the-flies the-large.M.SG entered.M.SG from.the-window
 'Those large flies came in through the window.'

With singular nouns and with those in the determinate plural form (with a numeral 2-10) the agreements are just as expected. But with the collective nouns (15) we find masculine singular agreements both in attributive position and in the predicate. According to Manwel Mifsud, this is the normal agreement; however, some speakers would use the plural (*dawk id-dubbien il-kbar dahlu*) as a result of the influence of English. Let us now look at other types of agreement, starting with the relative clause:

- (16) Dak id-dubbien il-kbir li hareg mill-bieb
 that.M.SG the-flies the-large.M.SG that went.out.M.SG
 mill-bieb through.the-door
 'Those large flies which went out through the door entered through the window.'

The relative marker *li* does not itself show number, but the verb does, and in this instance it is singular. Plural agreements are also possible:

- (17) Dawk id-dubbien il-kbar li hareg mill-bieb
 those.PL the-flies the-large.PL that went.out.PL through.the-door
 dahlu mit-tieqa
 entered.PL from.the-window
 'Those large flies which went out through the door came through the window.'

According to Manwel Mifsud, in villages where Semitic Maltese is stronger (those further from the main centres), (17) might be noticed as odd, but for Maltese speakers who are bilingual with English, (17) would be quite normal. This leaves the question of the form of pronouns. We can continue the examples further, bearing in mind that subject pronouns are normally dropped:

- (18) Dak id-dubbien il-kbir li hareg
 that.M.SG the-flies he-large.M.SG that went.out.M.SG
 mill-bieb dahal mit-tieqa
 through.the-door entered.M.SG from.the-window
 u ghadu jtir fil-kamra
 and still.SG fly.SG in.the-room
 'Those large flies which went out through the door came in through the window and are still flying in the room.'

Għad 'still' carries a singular subject clitic *-u*, and the following verb is also singular. A plural variant is also possible:

- (19) Dawk id-dubbien il-kbar li hareg mill-bieb
 those.PL the-flies the-large.PL that went.out.PL through.the-door
 dahlu mit-tieqa u ghadhom itiru¹⁴ fil-kamra
 entered.PL from.the-window and still.PL fly.PL in.the-room
 'Those large flies which went out through the door came in through the window and are still flying in the room.'

The use of the plural in the subject clitic (and hence its agreeing verb) was felt to be better than in the other agreement positions (attributive modifier, predicate and relative pronoun). What then if there is a sentence break? Possible continuations include:

- (20) Huwa ahdar.
 3RD.M.SG green.M.SG
 'They are (literally 'it is') green.'

- (21) Huma hodór.
3RD.PL green.PL

'They are green.'

While Manwel Mifsud states that the singular is possible, as in (20), he believes that the plural (21) is more likely even as a continuation to (18); as a continuation to (19) only the plural is possible.

To sum up: the collective in Maltese is not subject to the Smith-Stark Hierarchy either in terms of the nouns which have the extra number form or in terms of the resulting alternative agreement possibilities. The agreement choice which results with those collectives which have a natural plural reference is, however, constrained by the Agreement Hierarchy. The plural - the semantic agreement form - is possible for some speakers in all positions on the hierarchy but is favoured in only one, the personal pronoun, which is of course the rightmost position on the hierarchy.

5. *The remnants of a minor number: the count form in Serbo-Croat*

The majority of Slavonic languages have lost the dual number, but it has left curious traces in morphology and syntax. Typically the effect is found with the numerals 2, 3 and 4 (and a small number of items like the word for 'both'). However, as we shall see, we are no longer dealing with minor number. The point is that there are morphological effects with a small number of nouns in some languages, and more extensive syntactic effects in several languages, but these forms have lost all semantic significance. That is to say, the special forms cannot be used except with the numerals which determine them.

For the morphological effects let us consider Russian. In this language phrases consisting of a numeral 2, 3 or 4 plus noun (when the phrase is in the nominative case) require the noun to stand in the genitive singular. This is the result of the loss of the dual number (for some types of noun the dual and the genitive singular were identical). Just a few nouns retain a special form, differentiated by stress. For instance, we find *dva čásá* 'two hours, two o'clock', while the normal genitive singular of *čas* is *časá*. The number of nouns involved is extremely small: *rjad* 'row', *šar* 'sphere', *šag* 'step'. Evidently these are not high on the Smith-Stark Hierarchy. The important point, though, is that the special forms like *časá* can only be used with the numeral - they have no independent semantic value. The other nouns in Russian use the genitive singular, and again this is totally determined by the presence of the numeral.

There are also interesting agreement effects in some instances, which deserve our attention. For these we will consider the South Slavonic language Serbo-Croat as an illustration, and concentrate on masculine nouns, which are the most interesting in this regard. As in Russian, in phrases with the numerals 2, 3 and 4, masculine nouns require a special form, a survival of the dual number which is synchronically a genitive singular (it is variously called the 'count form', 'dual' and 'paucal'). Moreover, attributive modifiers must take the ending -a; for some (such as *ovaj* 'this', in (23) below) this form cannot be a genitive, and it has been argued that it should be analysed synchronically as a neuter plural (Corbett 1983:13-14, 89-92):

- (22) dva dobra čoveka
two good.COUNT man.GEN.SG=COUNT
'two good men'

No matter how this form is analysed, it is an unexpected agreement form, resulting from the presence of the numeral. Note that all countable masculine nouns are involved: the unusual effect is thus a syntactic one, and the number of nouns involved is significant. In the predicate the count form is found, but so too is the masculine plural form:

- (23) ova dva čoveka su dobra /dobri
these.COUNT two man are good.COUNT /good.M.PL
'these two men are good'

The count form *dobra* represents syntactic agreement. The masculine plural *dobri* (which is the same as would be found with an ordinary masculine plural noun with no overt quantifier) represents semantic agreement. The relative pronoun is also found in both forms:

- (24) dva čoveka koja /koji ...
two man who.COUNT /who.M.PL
'two men who ...'

The personal pronoun must take the masculine plural form *oni* (**ona* is unacceptable). Thus we have syntactic agreement in attributive position, both types of agreement of the predicate and relative pronoun and only semantic agreement of the personal pronoun. This unexpected 'remnant' agreement, in competition with the masculine plural, is thus distributed in accordance with the constraint of the Agreement Hierarchy. We can go further, in that there are figures for the relative frequency of the two forms in the positions where there is an option. These are derived from Sand (1971:55-6, 63) and presented in Table 1:

Table 1. Percentage distribution of count and plural forms in Serbo-Croat.

| | attributive | predicate | relative pronoun | personal pronoun |
|--|-------------|---------------|----------------------------|------------------|
| percentage showing plural (semantic) agreement | 0 | 18 (N=376) | 62 (N=32) ¹⁵ | 100 |

Table 1 shows a monotonic increase¹⁶ in the likelihood of agreement forms with greater semantic justification.

6. Discussion

The main issue concerns minor number and its relation to the Smith-Stark Hierarchy and Agreement Hierarchy (section 6.1). However, we should also consider more carefully what can be called a minor number, and how the notion relates to other minor categories (section 6.2).

6.1. Constraints on minor numbers

It is clear that minor numbers are not subject to the Smith-Stark Hierarchy; we have seen this in each of the languages investigated, whether we looked at morphological¹⁷ or syntactic evidence. And if anything the counter-examples were more serious than indicated, because typically the minor numbers are found with nouns only, so that the personal pronouns are also counter-examples. The strength of the counter-evidence is surprising, given the extent of the evidence Smith-Stark brings in support of the 'plurality split'. I suggest that Smith-Stark's hierarchy is concerned not specifically with the plural number (leading naturally to investigations of the 'duality split', 'paucal split' and so on) but rather with the number category as a whole. The Smith-Stark Hierarchy determines the extent to which a particular language incorporates number - as shown by the range of nouns which are involved in number, by differentiating the basic number opposition of singular versus plural. Given a language with large numbers of nouns involved in the major number opposition (that is, nouns which are countable), then small additional 'patches' in the number system are possible. These may form a semantic group (such as paired body parts) or they may be more idiosyncratic in nature.

In the absence of an animacy constraint for these minor numbers,

it is understandable that in the main nouns of low animacy are involved. For instance, in the case of Fula general number allows the speaker to avoid choosing number, and instances where this is desirable are likely to include instances with nouns low on the animacy scale. And collectives of the Maltese type are more likely to be appropriate with referents of low individuation.

The minor numbers we investigated involved, in the main, small lexical clusters. Often the nouns in question had the normal number opposition, plus some additional possibility. In the case of Maltese, there were two minor numbers coexisting, but not interacting with each other.

In most instances the evidence consisted solely of marking on the noun; in two cases there was number marking and agreement. Given the small semantic clusters of nouns involved, it is not surprising that marking on the noun itself is the main indicator. There is insufficient evidence to claim that marking of a minor number on the noun is a necessary condition for the indication of a minor number by agreement. (Our sample is small, and the Spanish dialects show mass agreement with nouns which do not mark mass number as well as with nouns which do: it is not excluded that another language could mark mass number solely by agreement.)

For those languages which mark minor number by agreement, it is intriguing to ask why, since they are not subject to the Smith-Stark Hierarchy, they are subject to the Agreement Hierarchy. Furthermore, when minor numbers decay to the point of having no semantic significance, they may leave effects in morphology and syntax - again not subject to the Smith-Stark Hierarchy but subject to the Agreement Hierarchy. The Agreement Hierarchy is concerned with the interaction of the choice of syntactic or semantic agreement with different agreement targets. (The possibility of choice arises from lexical or syntactic peculiarities; the specifics of minor number are just one of the causes of a potential choice between syntactic and semantic agreement - others involve gender for instance.) As we move rightwards along the hierarchy, so agreement is determined decreasingly by lexical or syntactic peculiarities and increasingly by semantics. While minor numbers can apply to relatively small and even idiosyncratic groups of nouns, they cannot evade the constraint of the Agreement Hierarchy, for which they are subsumed under the relevant major number (normally the plural).

6.2. Minor number and other minor categories

Earlier we defined a minor number as one which involves a relatively small proportion of the nouns of the language. The difficulty with

this definition is that a language with a single number opposition which involved relatively few nouns would fall under this definition as having a minor number. Yet such languages typically provide good evidence in favour of Smith-Stark's claim: typically the nouns are high on the Smith-Stark Hierarchy. Intuitively we want to say that there cannot be minor number without there first being major number (a plurality split). We therefore modify our definition to say that a minor number is one which involves a proportion of the nouns of a given language which is relatively small by comparison to those involved in the major number(s). (The criteria for minor numbers are further tightened in Corbett & Mithun 1996). Since the same type of number (such as dual) may be a major or a minor number, it is important that language descriptions make it clear which is intended: 'Modern Hebrew has the dual' is true, but potentially misleading.

Similar oppositions may be found in other categories. The problem of minor (controller) genders - those which have few nouns in them - is discussed in Corbett (1991: 159-60, 170-5). It is argued that there can be 'genuine' genders having few members, but which must be recognised as true genders because of the unique agreement forms the nouns take. Conversely if the nouns take only gender agreement forms required by nouns in the major genders (though in an unusual combination) then they can be marked as lexical exceptions. In this case the more restricted notion of 'incurate gender' is invoked. The comparison with minor number is clear: when only a small number of lexical items is involved in a possible category, we are required to take the analysis further. There are instances when a number or gender has few members, but functions normally (thus if a language has few nouns which show singular and plural number, we have a normal category of plural). Equally there can be idiosyncratic minor numbers and genders which are not subject to the normal constraints.

In the category of case too we find comparable problems: thus Russian has six major cases, and two further cases restricted to a small subset of nouns and available in restricted syntactic contexts. Unlike the other cases, these two are marked on nouns only (and a noun of the appropriate declensional type may have either or both of these minor cases).

7. Conclusion

Minor numbers may vary widely in their semantic type, and in the number of nouns involved (from a handful to a sizable minority of the noun inventory). Minor numbers are not subject to the Smith-Stark

Hierarchy; they are, however, subject to the Agreement Hierarchy. Given the distinctive behaviour of minor number, it is important that linguistic analyses make clear the type of category involved since typological claims will differ in their validity according to whether major or minor number is involved. And more generally, it cannot be assumed that typological claims which are valid for major categories will be similarly valid for corresponding minor categories.

Address of the Author:

Department of Linguistic and International Studies, University of Surrey,
Guildford, Surrey, GU25XH, United Kingdom; email: g.corbett@surrey.ac.uk

Notes

- ¹ I am very grateful to Manwel Mifsud and to Ray Fabri for their help with the Maltese data. For several most helpful comments and references I would like to thank Dunstan Brown, Barrie Evans, Norman Fraser, David Gil, Jim Hurford, Edith Moravcsik, Almerindo Ojeda, Frans Plank, Mike Thacker, Nigel Vincent and Dieter Wunderlich; errors are mine. This is a revised version of a paper which appeared in BORG A. & PLANK F. (eds), *The Maltese noun phrase meets typology* (European Science Foundation Programme in Language Typology: Theme 7, Noun Phrase Structure: Working Paper 25). The support of the ESF and of the British Academy is gratefully acknowledged.
- ² As Cedric Smith-Stark has pointed out (personal communication) the hierarchy is prefigured in Forchheimer (1953:12-13). For other precursors see Plank (1987: 181).
- ³ For background information on plural formatives in Budugh and related languages see Ibragimov (1974), who claims that there was formerly a more general distinction of restricted plural or dual versus unrestricted plural.
- ⁴ A rich source of historical data is Santangelo (1981). For a wealth of textual examples see Brunet (1978:30-76), and for agreement problems see Brunet (1978:95-6). See also Rocchetti (1968). For a detailed account of the semantics of the different plurals see Ojeda (1995).
- ⁵ This is termed the 'pseudo-dual' by Blanc (1970), who surveys the status of the dual across the Arabic dialects.
- ⁶ Nor, of course, are the nouns which have a dual all nouns *except* those denoting natural pairs.
- ⁷ The symbol *d* indicates a preglottalized *d*.
- ⁸ See Ojeda (1989) for numerous examples of mass agreement from a range of dialects.
- ⁹ Sutcliffe is perhaps best read in the light of the comments by Borg (1981: 106-9).
- ¹⁰ Clearly the collective in Maltese is very different from the types of collective plural discussed earlier; as Gil (this volume) points out, the term 'collective' is used in a confusingly large number of different ways. For an account of the formal semantics of number in Arabic see Ojeda (1992b).

