

Language Typology and
Language Universals
Sprachtypologie und
sprachliche Universalien
La typologie des langues et
les universaux linguistiques

An International Handbook /
Ein internationales Handbuch / Manuel international

Edited by / Herausgegeben von / Edité par
Martin Haspelmath · Ekkehard König
Wulf Oesterreicher · Wolfgang Raible

Volume 1 / 1. Halbband / Tome 1

Offprint

Walter de Gruyter · Berlin · New York
2001

len modalen Kategorien (Palmer 1986, Bybee & Fleischman 1995), was fehlt, sind Übersichten über Gesamtheiten von modalen Subsystemen in verschiedenen Sprachen, die zum Beispiel den Eindruck präzisieren oder korrigieren könnten, dass die modalen Subsysteme in einigen Sprachen stärker volitionalitätsorientiert sind, in anderen Sprachen hingegen stärker normenorientiert, und dass das Deutsche den ersteren Typ repräsentiert, das Lateinische den letzteren.

6. Zitierte Literatur

- Barnes, J. 1984. „Evidentials in the Tuyuca verb“. *International Journal of American Linguistics* 50: 255–271.
- Bartsch, Renate. 1972. *Adverbialsemantik*. Frankfurt am Main: Athenäum.
- Bybee, Joan L. & Fleischman, Suzanne (eds.) 1995. *Modality in Grammar and Discourse*. Amsterdam: Benjamins.
- Bybee, Joan L. & Perkins, Revere & Pagliuca, William. 1994. *The Evolution of Grammar: Tense, Aspect and Modality in the Languages of the World*. Chicago: University of Chicago Press.
- Chafe, Wallace & Nichols, Johanna (eds.) 1986. *Evidentiality: The Linguistic Coding of Epistemology*. Norwood, NJ: Ablex.
- Chung, Sandra & Timberlake, Alan. 1985. „Tense, Aspect, and Mood“. In: Shopen, Timothy (ed.). *Language Typology and Syntactic Description. Vol. III: Grammatical Categories and the Lexicon*. Cambridge: Cambridge University Press, 202–258.
- Cole, Peter. 1982. *Imbabura Quechua*. Amsterdam: North Holland.
- Comrie, Bernard. 1987. „Conditionals: A Typology“. In: Traugott, Elisabeth C., et al. (eds.) *On Conditionals*. Cambridge: Cambridge University Press, 77–99.
- Comrie, Bernard & Smith, Norval. 1977. „Lingua Descriptive Studies: Questionnaire“. *Lingua* 42: 1–72.
- Deuchar, Margaret. 1984. *British Sign Language*. London: Routledge.
- Donaldson, Tamsin. 1980. *Ngiyambaa: the language of the Wangaaybuwan*. Cambridge: Cambridge University Press.
- Enç, Mürvet. 1996. „Tense and Modality“. In: Lapin, Shalom (ed.). *The Handbook of Contemporary Semantic Theory*. London: Blackwell.
- Givón, Talmy. 1990. *Syntax. A functional-typological introduction. Vol. II*. Amsterdam: Benjamins.
- Haan, Ferdinand de. 1997. *The interaction of modality and negation. A typological study*. New York: Garland.
- Hengeveld, Kees. 2000. „Mood and Modality“. In: Booij, Geert & Lehmann, Christian & Mugdan, Joachim (eds.). *Morphology: A Handbook of Inflection and Word Formation*. Berlin: de Gruyter.
- Jakobson, Roman. 1971. „Shifters, Verbal Categories, and the Russian Verb“. In: Jakobson, Roman. *Selected Writings II: Word and Language*. Den Haag: Mouton.
- Janzen, Terry & Shaffer, Barbara. 2001. „Gestures as the Substrate in the process of ASL Grammaticization“. In: Cormier, Kearsy, et al. (eds.). *The Effects of Modality on Language and Linguistic Theory*. Austin, TX: Texas Linguistic Forum.
- Kemmerling, Andreas. 1999. „Grice actions“. In: Cosenza, Giovanna (ed.). *Paul Grice's Heritage*. Turnhout: Brepols.
- Kratzer, Angelika. 1991. „Modality“. In: Stechow, Armin von & Wunderlich, Dieter (eds.). *Semantik/ Semantics. Ein internationales Handbuch der zeitgenössischen Forschung/An International Handbook of Contemporary Research*. Berlin: de Gruyter.
- Lang, Ewald. 1983. „Einstellungsausdrücke und ausgedrückte Einstellungen“. In: Ruzicka, Rudolf & Motsch, Wolfgang (Hgg.). *Untersuchungen zur Semantik*. Berlin: Akademie Verlag, 305–341.
- Lazard, Gilbert. 1999. „Mirativity, evidentiality, mediativity, or other?“. *Linguistic Typology* 3: 91–109.
- Lewis, G. L. 1967. *Turkish Grammar*. Oxford: Oxford University Press.
- Li, Charles N. & Thompson, Sandra A. 1984. „Mandarin“. In: Chisholm, William S., jr. (ed.). *Interrogativity*. Amsterdam: John Benjamins, 47–61.
- Löbner, Sebastian. 1990. *Wahr neben falsch: duale Operatoren als die Quantoren natürlicher Sprache*. Tübingen: Niemeyer.
- Lowe, I. 1972. „On the relation of the formal and the sememic matrices with illustrations from Nambiquara“. *Foundations of Language* 8, 360–390.
- Lyons, John. 1977. *Semantics*. 2 vols. Cambridge: Cambridge University Press.
- Montague, Richard. 1974. *Formal Philosophy*. New Haven: Yale University Press.
- Nedjalkov, Igor. 1997. *Evenki*. London: Routledge.
- Noonan, Michael. 1985. „Complementation“. In: Shopen, Timothy (ed.). *Language typology and syntactic description. Vol. II: Complex constructions*. Cambridge: Cambridge University Press, 42–140.
- Palmer, Frank Robert. 1986. *Mood and Modality*. Cambridge: Cambridge University Press.
- Payne, John R. 1985. „Negation“. In: Shopen, Timothy (ed.). *Language Typology and Syntactic Description. Vol. I: Clause structure*. Cambridge: Cambridge University Press 1985, 197–242.
- Payne, Thomas E. 1997. *Describing morphosyntax. A guide for field linguists*. Cambridge: Cambridge University Press.

- Portner, Paul. 1999. „The semantics of mood“. In: *Glott International* 4.1: 3–9.
- Quirk, Randolph & Greenbaum, Sidney & Leech, Geoffrey & Svartvik, Jan. 1985. *A Comprehensive Grammar of the English Language*. London: Longman.
- Rescher, Nicholas. 1968. *Topics in Philosophical Logic*. Dordrecht: Reidel.
- Sadock, Jerrold M. & Zwicky, Arnold M. 1985. „Speech Act Distinctions in Syntax“. In: Shopen, Timothy (ed.). *Language Typology and Syntactic Description. Vol. I: Clause structure*. Cambridge: Cambridge University Press, 155–196.
- Searle, John R. 1969. *Speech acts. An essay in the philosophy of language*. Cambridge: Cambridge University Press.
- Searle, John R. 1975. „A Taxonomy of Illocutionary Acts“. In: Gunderson, K. (ed.), *Language, Mind, and Knowledge*. Minneapolis: University of Minnesota Press. [Reprinted in: Searle 1979, 1–29, and under the more appropriate title „A Classification of Illocutionary Acts“ in: *Language in Society* 5 (1976), 1–23.]
- Searle, John R. 1979. *Expression and Meaning*. Cambridge: Cambridge University Press.
- Searle, John R. 1989. „How Performatives Work“. In: *Linguistics and Philosophy* 12, 535–558.
- Smeets, Ineke. 1989. *A Mapuche Grammar*. Doctoral Dissertation, Rijksuniversiteit te Leiden. Leiden.
- Talmy, Leonard. 1988. „Force Dynamics in Language and Cognition“. *Cognitive Science* 12: 49–100.
- Trask, Robert Lawrence. 1993. *A Dictionary of Grammatical Terms in Linguistics*. London: Routledge.
- van der Auwera, Johan & Plungian, Vladimir A. 1998. „Modality's semantic map“. *Linguistic Typology* 2: 79–124.
- Wittgenstein, Ludwig. 1960. *Schriften I: Philosophische Untersuchungen*. Frankfurt am Main: Suhrkamp.
- Zaefferer, Dietmar. 1983. „The Semantics of Non-Declaratives: Investigating German Exclamatories“. In: Bäuerle, Rainer, et al. (eds.). *Meaning, Use, and Interpretation of Language*. Berlin: de Gruyter, 466–490.
- Zaefferer, Dietmar. 1990. „On the Coding of Sentential Modality“. In: Bechert, Johannes & Bernini, Giuliano & Buridant, Claude (eds.). *Toward a Typology of European Languages. Proceedings of the Workshop held at Consiglio Nazionale delle Ricerche, Rome, 7–9 January 1988*. Berlin: Mouton de Gruyter, 215–237.
- Zaefferer, Dietmar. 1991. „Conditionals and Unconditionals: Cross-linguistic and Logical Aspects“. In: Zaefferer, Dietmar (ed.). *Semantic Universals and Universal Semantics*. Berlin: Foris 210–236.
- Zaefferer, Dietmar. 2001. „Deconstructing a classical classification: A typological look at Searle's concept of illocution type“. *Revue internationale de philosophie*.

Dietmar Zaefferer,
Ludwig-Maximilians-Universität München
(Deutschland)

61. Number

1. Nominal and verbal number
2. Number values
3. Implicational claims (the Number Hierarchy)
4. The nominals involved
5. The expression of number
6. Conclusion
7. Special abbreviations
8. References

1. Nominal and verbal number

What type of category is number? The obvious answer, certainly for speakers of Indo-European languages, is that it is a nominal category, affecting primarily nouns and pronouns. The difference between *head* and *heads*, or the Russian equivalents *golova* and *golovy* is the number of heads involved.

Number may be shown by verbs too in English (and Russian, and many other languages):

- (1) *my friend speaks Russian*
- (2) *my friends speak Russian*

Though number is marked on the verb here as well as on the noun, the essential difference between (1) and (2) is, of course, the number of friends involved. The type of number we are dealing with can be seen particularly clearly in these examples:

- (3) *the sheep drinks from the stream*
- (4) *the sheep drink from the stream*

Though the form of the noun does not change, and the marker of number is on the

verb, it still indicates the number of sheep involved. (Example (4) cannot be used in English for the situation in which *one* sheep drinks several times.) In other words, we have nominal number which is expressed on the verb (usually, in English, in addition to being expressed on the noun). Number in English is obligatory (there is no form of a noun which allows us to avoid specifying the number, instead we have to indicate one or more than one), it is relevant to syntax, as the agreement facts show, and it is largely regular (examples like *friend* ~ *friends*, greatly outnumber those like *sheep* ~ *sheep* and *criterion* ~ *criteria*). Thus number is an inflectional category in English: *friend* (singular) and *friends* (plural) are forms of the same lexical item FRIEND.

There are many languages which are broadly comparable to English in this respect. But there are also many languages in which number is fundamentally different: in particular it may be not a nominal category but a verbal one. Moreover, it is often optional, rather than obligatory, and highly irregular, which suggest that it is then a derivational category. For an example we turn to Rapanui (the language of Easter Island, one of the Oceanic languages within Austronesian):

- (5) Rapanui (Veronica Du Feu, personal communication)

ruku
'dive'

- (6) *ruku ruku*
'go diving'

The form in (6) implies more than one dive, but not necessarily more than one diver. The event is in a sense plural and reduplication is used here to indicate verbal plurality. There are other possibilities for verbal number, the main one being that it may be concerned with the number of participants (thus several eating together may count as different from one eating alone). An important account is that of Durie (1986). Verbal number is found in many linguistic areas: it is particularly widespread in North America (Mithun 1988: 231). It is also found in the South Central Dravidian group of languages of southern India (Steever 1987) and in many languages of Africa (Brooks 1991), the Chadic group being particularly well documented (Newman 1990: 53–87). Verbal number may be restricted to relatively small numbers of verbs,

and it rarely shows more than a two-way distinction (one versus several). For typological work it is vital to be clear what type of number is being discussed. Since nominal number shows greater variation, we shall concentrate on that type. The material presented here in summary form is treated more extensively in Corbett (2000).

2. Number values

We now consider the possible values for nominal number. The common ones are singular and plural (as in English and Russian), but there are several more. First, we should note that sometimes number can be 'avoided' as it were, that is there are forms which are outside the number system. An example is the Cushitic language Bayso which at the last count had a few hundred speakers on Gidicho Island in Lake Abaya (southern Ethiopia) and on the western shore of the lake. In Bayso, nouns have a form which represents what we shall call 'general' meaning (the German tradition is to call such forms 'transnumeral'), that is, it is non-committal as to number (Corbett & Hayward 1987). *Lúban* 'lion' denotes a particular type of animal, but the use of this form does not commit the speaker to a number of lions – there could be one or more than that. Other forms are available for indicating reference specifically to one or to more than one lion when required, as we shall see in § 2.4.

- (7) Bayso (Dick Hayward, personal communications, Corbett & Hayward 1987)

lúban *foofe*
lion.GENERAL watched.1.SG
'I watched lion' (it could be one, or more than that.)

While it is rare to have separate general forms, there are very many languages which can express general meaning, but with a form shared with the singular. This more usual situation, with general identical to singular, can be illustrated from the Cushitic language Arbore. We find pairs like the following:

- (8) Arbore (Hayward 1984: 159–183)
- | | |
|--------------------------|--------------------------|
| <i>general</i> | <i>plural</i> |
| <i>kér</i> 'dog(s)' | <i>ker-ó</i> 'dogs' |
| <i>garlá</i> 'needle(s)' | <i>garlá-n</i> 'needles' |

It is important to stress that, though the morphology may appear comparable to English,

the semantics of the forms is quite different: *keró* guarantees more than one dog, while *kér* does not imply only one: it might be one, it might be more than that. (There are other, less frequent number pairings in Arbore.) Systems like this, in which number is not an obligatory category (and so is arguably not inflectional) are common in the world's languages.

From now on we shall assume that number is to be expressed, and consider the possible values of the category.

2.1. The plural

The simplest system, and a common one, has an opposition:

- (9) singular plural

2.2. The dual

The dual refers to two distinct real world entities. If a dual is added to our previous system, we have another common system:

- (10) singular dual plural

Examples can be found all over the world, for instance, in Upper Sorbian, an endangered West Slavonic language. Some of the forms are given in Table 61.1.

It is important to note that the introduction of the dual has an effect on the plural. More generally, a change in system gives the plural a different meaning; if the system is singular-dual-plural, the plural is for three or more real world entities, a point made by Saussure (1971: 161). The dual has long fascinated linguists, a notable early example being Humboldt; see Plank (1989) for discussion and references.

2.3. The trial

Just as the dual is for two, the trial is for referring to three distinct real world entities. Adding it to systems like those just discussed gives the following system of number values:

- (11) singular dual trial plural

Such a system is found in Larike, a Central Moluccan language with 8–10,000 speakers on the western tip of Ambon Island, Central Maluku, Indonesia. Central Moluccan forms part of the Central Malayo-Polynesian subgroup of Austronesian; the data are from Laidig & Laidig (1990). Larike distinguishes singular/dual/trial/plural in its free pronouns (though there are no third person pronouns for non-human referents):

- (12) Larike (Laidig & Laidig 1990)
Duma hima aridu naʔa
 house that 1.TRIAL.EXCL own.it
 'We three own that house.'

It also makes these distinctions in its various series of pronominal affixes:

- (13) *Kalu iridu-ta-ʔeu, au-na-wela*
 if 2.TRIAL-NEG-go 1.SG-IRR-go.home
 'If you three don't want to go, I'm going home.'

It is interesting to note that the dual and trial forms originate from the numerals 'two' and 'three', and that the plural comes historically from 'four'. Such developments are fairly common in Austronesian languages. There are also numerous instances of former trials becoming paucals. This is a potential hazard for the typologist: the term 'trial' is sometimes used according to the form of the inflections (derived historically from the numeral three), even when the forms are currently used for small groups including those greater than three (and so are paucals) and sometimes the term is used according to meaning (for genuine trials). This shows the need for typologists to be careful in the use of terms in this area. The Larike trial is a genuine trial:

"... it should be stated explicitly that Larike trials are true trial forms. In other words, they represent the quantity three, and are not used to refer to the more vague notion of several, as is a paucal or limited plural." (Laidig & Laidig 1990: 92)

Table 61.1: The dual in Upper Sorbian (Stone 1993)

singular	dual	plural
<i>ja</i> 'I'	<i>mój</i> 'we two'	<i>my</i> 'we'
<i>ty</i> 'you'	<i>wój</i> 'you two'	<i>wy</i> 'you (all)'
<i>hród</i> 'palace, castle'	<i>hrodaj</i> 'two palaces'	<i>hrody</i> 'palaces'
<i>džětam</i> '(I) work'	<i>džětamoj</i> '(we two) work'	<i>džětamy</i> '(we) work'

The Larike trial is 'facultative', a question to which we return in § 3.2 below. Ngan'git-yemerri (a Daly language with two dialects, Ngan'gikurunggurr and Ngan'giwumirri, and with 100 speakers, 300 miles SW of Darwin, Australia) has a trial, strictly for three (Nicholas Reid 1990: 118–119 and personal communication) as does Marrithiyel, another Daly family language (Green 1989: 136–139).

These then are languages with genuine trials, appropriate just when referring to three entities. There is a question as to whether there are also languages with quadrals (for reference to four entities). However, having raised the issue of paucals, we shall first continue the analysis of these, and only then return to the question of quadrals.

2.4. The paucal

The paucal is used to refer to a small number of distinct real world entities. It is similar to the English quantifier 'a few' in meaning, particularly in that there is no specific upper bound that can be put on its use. (Its lower bound, like that of the plural, will vary according to the system in which it is embedded.) Let us return to the Cushitic language Bayso. Besides the general number forms, there are also these:

- (14) Bayso (Dick Hayward, personal communications, Corbett & Hayward 1987)
lubán-titi foofe
 lion-SG watched.1SG
 'I watched a lion'
- (15) *luban-jaa foofe*
 lion-PAUCAL watched.1SG
 'I watched a few lions'
- (16) *luban-jool foofe*
 lion-PLURAL watched.1SG
 'I watched (a lot of) lions'

Bayso then has a paucal, with singular and plural, giving the following system (in addition to general number):

- (17) singular paucal plural

The paucal is used in Bayso for reference to a small number of individuals, from two to about six. Bayso has this system in nouns but not in pronouns: thus the system of number values in a given language can vary according to which part of the grammatical system is examined. This is an essential point for typologists: when we say that a language has a

particular number value, we need to be clear about its range of use – whether is it available for most nouns or found just with the personal pronouns, for example.

Systems with just a paucal in addition to singular and plural are rare. It is much more common to find it together with a dual, giving this system:

- (18) singular dual paucal plural

Here the meaning of the paucal changes to exclude two. This system is found, for instance in Yimas, a Lower Sepik language with 250 speakers in the Sepik Basin of Papua New Guinea. The paucal is found in the pronoun and in the pronominal affixes on the verb.

"The paucal expresses a set of a few; more than two and usually less than seven, but the exact number varies quite widely according to context. Prototypically, however, it refers to a class of three to five individuals, and is always restricted to humans." (Foley 1991: 216)

The restriction to humans is specific to Yimas, of course.

This system (with dual and paucal) is found widely in Oceanic languages, for instance in Paamese, spoken in Vanuatu. The factors governing the choice of paucal and plural in Paamese have been well described by Crowley (1982: 81): the lower the absolute size of the group, the more likely the paucal is to be used, the larger, the more likely the plural. But for groups in the middle (around six to twelve) then relative number becomes important: if the group is contrasted with some larger group, then the paucal is more likely, if contrasted with a smaller group, this will favour the plural.

2.5. The largest number systems

We now consider whether there are languages with the following system:

- (19) singular dual trial quadral plural

Such languages would have a quadral, a set of forms specifically for the quantity four. If such languages exist, they are rare. All the claims come from within the Austronesian family. A well-documented case is Sursurunga (Hutchisson 1986, and personal communications), which has some 4000 speakers in southern New Ireland. The forms labelled quadral are restricted to the personal pronouns, but are found with all of them, the first person (inclusive and exclusive), the second and the third:

Table 61.2: Emphatic pronouns in Sursurunga (Hutchisson 1986 and personal communications)

	singular	dual	"trial"	"quadral"	plural
1 exclusive	<i>iau</i>	<i>giur</i>	<i>gimtul</i>	<i>gimhat</i>	<i>gim</i>
1 inclusive	—	<i>gitar</i>	<i>gittul</i>	<i>githat</i>	<i>git</i>
2	<i>iáu</i>	<i>gaur</i>	<i>gamtul</i>	<i>gamhat</i>	<i>gam</i>
3	<i>-ilonlái</i>	<i>diar</i>	<i>ditul</i>	<i>dihat</i>	<i>di</i>

Here is an example of a quadral form in use (we retain the traditional label 'quadral' here although we are about to give reasons for replacing it):

- (20) Sursurunga (Hutchisson 1986 and personal communications)
gim-hat
 1.EXCL-QUADRAL
káwán
 maternal.uncle:nephew/niece
 'we four who are in an uncle-nephew/niece relationship'

(*á* is used to indicate schwa (*a*); this and other changes from the 1986 paper are based on personal communications from Don Hutchisson.) Besides being used of four, the quadral has two other uses. First, plural pronouns are never used with terms for dyads (kinship pairings like uncle-nephew/niece), and then the quadral is used instead for a minimum of four, and not just for exactly four (Hutchisson 1986: 10). The second additional use is in hortatory discourse; the speaker may use the first person inclusive quadral, suggesting joint action including the speaker, even though more than four persons are involved. These two special uses account for most instances of the quadral. If our terminology is based on meaning, the term 'quadral' is hardly appropriate, when in the majority of its uses the forms are not restricted to denoting foursomes. The forms would be better designated 'paucal'.

Let us consider the rest of the system in more detail (data and judgements from Don Hutchisson, personal communications). The dual is used quite strictly for two people (if there are two it must be used, and if it is used it indicates two). It is also used for the singular when the referent is in a taboo relationship to the speaker. This is a special use which does not alter the fact that its main use is as a regular dual. The trial will be used for three. But, it is also used for small groups, typically around three or four, and for

nuclear families of any size. It is therefore not strictly a trial, rather it could be labelled a paucal (an appropriate gloss would be 'a few'). We saw earlier that the trial frequently develops in this way. The quadral, as we have noted, is primarily used in hortatory discourse and with dyad terms; but otherwise it is used with larger groups, of four or more (an appropriate gloss would be 'several'). This too would qualify as a paucal; we therefore have two paucals, a paucal (traditionally trial) and a greater paucal (traditionally quadral).

The following example is particularly helpful for distinguishing the use of the two forms. It is from a letter to Don Hutchisson written in 1976:

- (21) *Iau lala hol pas gamhat*
 1.SG greatly think about 2.QUADRAL
kabin ngo iau lu mákái
 because that 1.SG HAB see
málálár gamtul minái i rum
 photo 2.TRIAL here in house
 'I am thinking about you [QUADRAL] all the time because I often see the picture of you [TRIAL] here in my house.'

The family consists of four members; the quadral is used first (perhaps to stress that all four are included), but then the writer moves to the trial, more normal usage for a small group. The entire family is intended in each case.

The plural, as we would expect, is for numbers of entities larger than what is covered by the quadral; however, there is no strict dividing line (certainly not at the number five). If we use semantic labels, as in the rest of this chapter, we should not call the forms trials and quadrals. Both have functions found with paucals elsewhere. We may therefore represent the system in Sursurunga like this:

- (22) singular – dual – paucal –
 greater paucal – plural

The system is no less interesting since it has a well documented five-valued number category. There are certainly other languages with five number values; we do not have such detailed information as for Sursurunga and there is no certain case of a quadral: it seems that in all cases the highest value in such systems can be used as a paucal. There are several false trails in the literature, that is, suggestions of other Austronesian languages with quadrals, which turn out in fact to have four number values not five. In such cases, the plural may have a form in which the numeral four can be reconstructed.

Besides the split in the paucal, we may also find a split in the plural, with 'greater plurals' of different types. For instance in Syrian Arabic some nouns have a plural *dabbānāt* 'flies' and an additional form, the greater plural, *dababīn* 'many flies' (Cowell 1964: 369). Greater plurals may imply an excessive number or else all possible instances of the referent. They are as yet poorly researched, except in a few languages.

3. Implicational claims (the Number Hierarchy)

According to Greenberg's universal 34: "No language has a trial number unless it has a dual. No language has a dual unless it has a plural" (Greenberg 1963). This claim appears fully justified. However, it is only a part of the overall typology of systems of number values. Some researchers give a Number Hierarchy, along these lines, suggesting it covers the possible number systems:

- (23) singular > plural > dual > trial

There are two problems with this hierarchy: first, it cannot be modified to include the different systems which have a paucal; and second, the patterns of what we shall call 'facultative' number are problematic.

3.1. Possible systems of number values

The first problem with the Number Hierarchy given above is that it does not account for systems which include a paucal. A modified hierarchy has been proposed:

- (24) singular > plural > dual >
paucal/trial

This would account for systems like the following:

- (25) singular plural (Russian)
singular dual plural (Upper Sorbian)
singular dual trial plural (Laríke)
singular dual paucal plural (Yimas)

However, it does not allow for systems which include the paucal in a different combination:

- (26) singular paucal plural (Bayso)

We must allow for the paucal to be an option at more than one point, which makes it clear that no straightforward hierarchy will be adequate. To make progress here we need to draw a distinction between 'determinate' and 'indeterminate' number values. These terms are to differentiate situations where, given the knowledge of real world which the speaker has, we can determine that only one form is appropriate (determinate number) from those where we cannot (indeterminate). Thus, in a language with an obligatory dual, this would be an instance of determinate number, since to refer to two distinct entities only the dual is appropriate. The determinate numbers are basically the numerical ones, (singular, dual, trial, plural). Use of determinate number values is agreed across speakers (different speakers agree that, say, the dual is appropriate for referring to two referents), it remains constant for the same speaker across different occasions, and it does not vary according to the referent (thus elephant-DUAL refers to two elephants just as ant-DUAL refers to two ants).

The indeterminate number values are the paucal, greater paucal and the greater plural. These may vary across speakers (there is no clear dividing line between paucal and plural, for one speaker across occasions), and can vary according to the referent (elephant-PAUCAL may refer to fewer real world entities than ant-PAUCAL). While the determinate numbers can be defined in terms of numerals, the indeterminates correspond to other quantifiers: 'a few', 'many', 'all'.

How then are number systems constrained? First, a language may take any number of the determinate number values, in the order given (i.e. in accord with the old number hierarchy). However, this should be seen as adopting a series of binary choices, and choice after the plural should be seen as removing a part from the range of the plural and hence dividing the plural. This gives the possibilities that are shown in Figure 61.1.

We have chosen to arrange the branches in this figure with the values for larger numbers

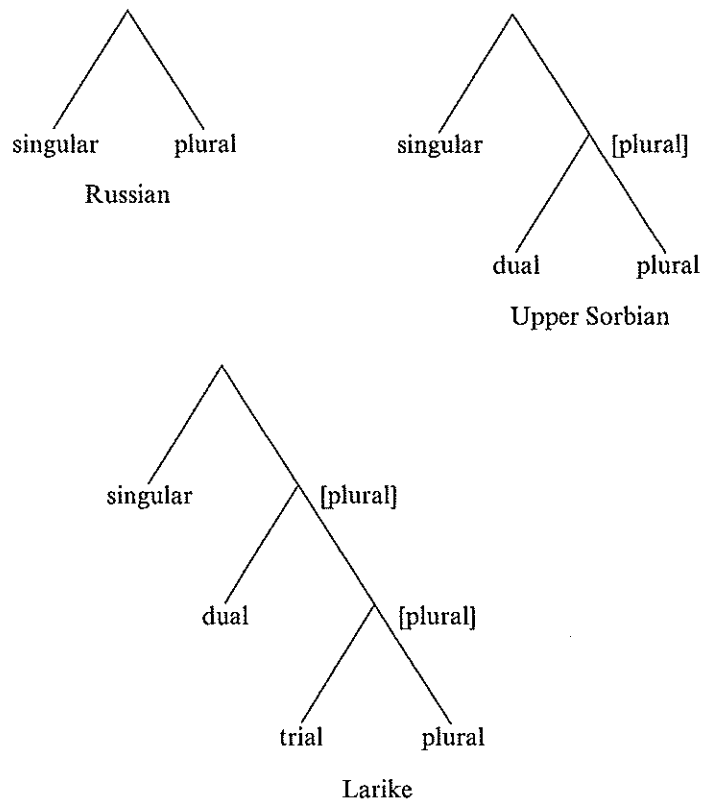


Figure 61.1: Illustration of possible number systems

of entities to the right. '[plural]' indicates what the value would be at that point if no further choices were made; this will be relevant when we consider facultative number. In addition to the determinate number values, languages may further divide up the plural space by taking an indeterminate number value. Most commonly, only one is selected. Some of the possibilities which result are shown in Figure 61.2.

While it is more common for just one indeterminate value to be selected, as in Bayso and in Mokilese (Harrison 1976: 88–89), two is also possible. The Mele-Fila system is based on material from Ross Clark (personal communications); it has a 'constructed system' in the sense discussed in § 5.3 below, cf. Figure 61.3.

Mele-Fila is perhaps the less surprising, in that it takes two indeterminate values of different types. Sursurunga has two paucals. It is tempting to try to add further constraints in order to bring the systems permitted into closer match with those so far recorded. This would be premature since we are still short

of data on the larger systems; it is to be hoped that highlighting these examples will encourage others to report on large number systems with indeterminate values included.

Besides making synchronic predictions, typology also makes diachronic predictions in that languages move from one possible system to another. Thus, a language with singular – dual – trial – plural may lose the trial, since the resulting system is allowed by the typology, but it could not lose the dual without first losing the trial. The drift from trial to paucal is easy to understand, since they occur in similar configurations.

We have set out a typology of possible number systems. As we shall see shortly, this same typology imposes further constraints on the number system.

3.2. Facultative number

We have considered how number systems vary according to how many number values they have, that is, how many different numbers of real world entities may be referred to by different means. But they may also differ

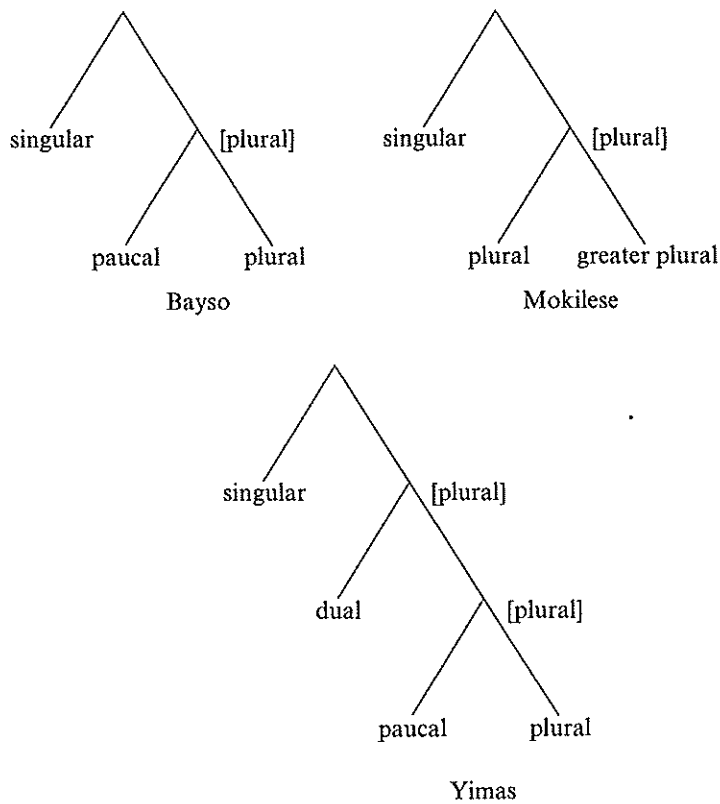


Figure 61.2: Possible number systems including an indeterminate value

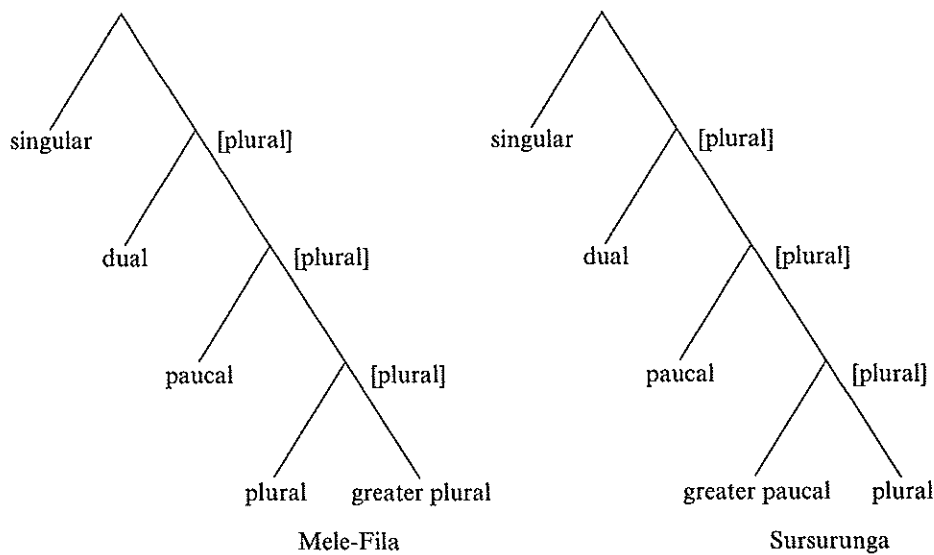


Figure 61.3: Number systems with two indeterminate values

in a more subtle way, according to whether the use of particular values is obligatory or 'facultative' (Greenberg 1966: 28). For instance, in Ngan'gityemerri (as noted in § 2.3)

there is singular, dual, trial and plural. The dual must be used to refer to two entities, the plural must be used for four and more. For three entities, the trial is used when the fact

of there being three is salient (for example, at the first mention in discourse) but otherwise the plural is used for three. Recall that the trial is strictly for three, and is not a paucal (Reid 1990: 118–119 and personal communication).

Consider now the systems with singular – dual – plural. The use of the dual may be obligatory, as in Sanskrit, or it may be facultative, as in the South Slavonic language Slovene. Here we do not find the same degree of choice as with the Ngan'gityemerri trial, but the important point is that the dual is not obligatory in the way that the plural is in Slovene:

"[...] in non-pronominal noun phrases with, for example, body parts that come in pairs like 'eyes' and 'feet', dual forms tend to be used only when the quantifiers 'two' or 'both' are explicitly stated in the context, and are replaced by the plural when this quantifier is unstated, even if a pair of referents are obviously implicit [...]" (Priestly 1993: 440–441)

Priestly gives the following example:

- (27) Slovene (Priestly 1993: 441)
nóge me bolijo
 foot.PL 1.SG.ACC hurt.PL
 'my feet hurt'

It is assumed that two feet are referred to, and the dual is not required in this example. Nominals express number obligatorily in Slovene; however, for referring to two entities, the use of the dual is not obligatory. Just as the plural is different in English (no dual) and Sanskrit (with dual), so it is different in Sanskrit (with an obligatory dual) and Slovene (with a facultative dual). A plural in Slovene may be for reference to just two real world entities.

Let us now consider how facultative number relates to the Number Hierarchy, repeated here for convenience:

- (28) singular > plural > dual > trial

If we have a system in which the dual is facultative, then in its place the less marked number, the plural, is used. It appears that the hierarchy is making useful predictions, based on markedness. Unfortunately this is only apparent here. Consider again Ngan'gityemerri: it has a trial which is facultative and so we would predict that the less marked dual could be used in its place. But of course this is not the case, the plural is used. This is what is expected if the system is viewed as a set of binary choices (see Figure 61.4).

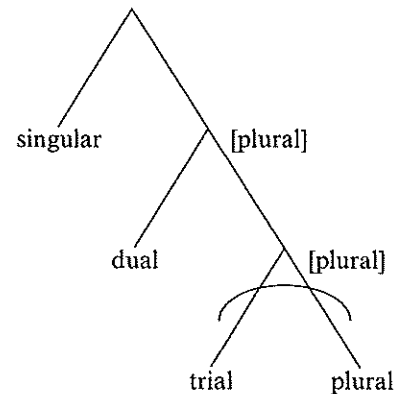


Figure 61.4: The facultative trial of Ngan'gityemerri

The point is that the last choice is facultative. If it is removed, as by the arc in figure 61.4, then Ngan'gityemerri has another possible system, singular – dual – plural, and the plural covers the area otherwise covered by trial and plural.

In Slovene, the situation is as in Figure 61.5:

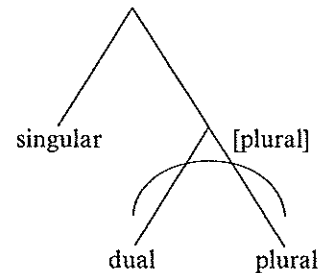


Figure 61.5: The facultative dual of Slovene

If the dual – plural choice is not taken up, then the system reverts to a straightforward singular – plural system. It is tempting to suggest that facultative number can only affect the 'last choice' of number, as in the examples so far. However, there are languages which show the situation is rather more interesting.

Let us consider just Larike, which we considered earlier as an example of a language with a genuine trial. Unlike Ngan'gityemerri, it is not only the trial which is facultative, the dual is as well:

"The Larike plural forms may also be used when referring to quantities of two or three. Thus, in spite of the fact that duals and trials are used to specifically denote twos and threes, plural forms can still be used with the meaning of two or more. In these situations, the choice of whether to use plural versus dual or trial forms depends upon the

speaker's desire to specify or focus upon the number of the referent nouns. Although the plural forms are probably most frequently used (even when referring to twos and threes), duals and trials are also quite common, and are often heard in routine conversations as well as in more formal language contexts." (Laidig & Laidig 1990: 93)

We represent this system in Figure 61.6.

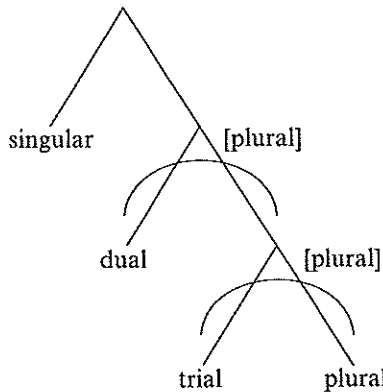


Figure 61.6: The facultative numbers of Larike

We cannot restrict facultative number values to 'the last choice'; rather we must say that if there is facultative number it must involve 'the last choice'. It may involve other numbers, working up from the last choice. Thus it may affect the dual-plural choice in Larike, because it also affects the trial-plural choice. But there could not be, we claim, a language with a facultative dual and an obligatory trial or paucal.

The existence of facultative numbers show how careful the typologist must be: Sanskrit and Slovene both have a dual, but they are rather different duals.

4. The nominals involved

So far we have concentrated on number values. Now we change tack and ask which nominals may be involved in the number system of a language, examining first the basic singular-plural opposition. We might say that we wish to establish the patterns of 'count' or 'countable' nouns in different languages. Unfortunately the terminology here has become rather confused. We shall call 'count' nouns those which are distinguishable for number, rather than those which may be counted. In part we are simplifying by looking at nominals, since, as Allan (1980) has argued, countability is a characteristic of noun phrases and not of nouns. However, as he points out:

"Even though countability is characteristic of NP's, not of nouns, it is nonetheless a fact that nouns do show countability preference – insofar as some nouns more often occur in countable NP's, others in uncountable NP's, and still others seem to occur quite freely in both." (Allan 1980: 566).

Thus, strictly speaking, we are investigating the countability preferences of nominals (we are as interested in pronouns as in nouns). There is considerably more variety in the world's languages than we might have expected. Consider the following Warrgamay example (Queensland, Australia):

- (29) Warrgamay (Dixon 1980: 266–268)
yibi-yibi gulmburu-ngu
 child-REDUP.ABS woman-ERG
wurrbi-bajun-du buudi-l-gani-y
 big-VERY-ERG take-CONT-UNM
malan-gu
 river-ALL
 'The very big woman/women is/are
 taking the children to the creek'

This example indicates that a noun can be marked for number in Warrgamay, as in *yibi-yibi* 'child', but this is not required; forms like *gulmburu-ngu* 'woman' are quite normal; in fact Dixon (1980: 267) says that a noun in this language 'is not normally specified for number' and suggests that this is the typical situation in Australia (1980: 22). Note especially that the verb in (29) does not determine number either (*-l-* and *-gani-* together indicate continuative and *-y-* indicates unmarked tense, hence the gloss 'is/are taking', Dixon 1980: 268). To check on the pronouns we turn to Dixon (1981: 39–40) The first and second persons, singular, dual and plural, and the third dual and plural are 'strictly specified for number' and are available only for reference to humans (and occasionally tame dogs). The form filling the third singular slot can range over all persons and all numbers (it can have non-human as well as human reference) but its 'unmarked sense' (1981: 40) is third person singular.

Thus the word for 'woman' is not normally specified for number, while in English it must be. Yet the first and second persons are. Could there be a language in which the word for 'woman' specified number but the first person pronoun did not? It seems not. It was known for some while that the patterns of nominals involved in number distinctions was related to animacy; this observation was taken up and developed by Smith-Stark (1974), who proposed this version of the Animacy Hierarchy:

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

Figure 61.7: The Smith-Stark (Animacy) Hierarchy

Smith-Stark suggested 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 in Georgian if 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. In Kalkatungu, a language of western Queensland with no known remaining full speakers (Barry Blake, personal communication), pronouns (free and bound) and demonstratives distinguish singular, dual and plural (Blake 1979: 31–32, 34–37). There is a dual and a plural marker for nouns; both are "common with kinship nouns", are part of the number system of demonstratives, but are "rarely used" with other nominals (1979: 80–81). And according to Masica (1991: 225–226) in Bengali number is obligatory for pronouns; other plural suffixes are optional.

The hierarchy presented by Smith-Stark is clearly akin to what in other publications has been termed the Animacy Hierarchy or the Topicality Hierarchy. He provides a good deal of data to support his claim, and notes some problematic cases too. Smith-Stark's article (1974) was a major step forward in our understanding of number systems; on the other hand, it is rather confusing in places, and a lot of the relevant data are missing (that remains the case – it has not been followed up as well as it deserved).

It is worth considering the nouns which are off the bottom of the scale, those which do not enter into number oppositions. In English they typically pattern with the singulars, thus *honesty* has the form of a singular and takes singular agreements. This is not the only possible pattern: In Manam (Lichtenberk 1983: 269), mass nouns are treated as plural (unless they refer to a single quantity):

(30) *day di-éno*
 water 3.PL-exist
 'there is water (available)'

In various Bantu languages we typically find that some mass nouns are singular and some plural.

The reference to Manam mass nouns being plural unless they refer to a single quantity (when they are singular) recalls English, where mass nouns can also be recategorized as count nouns. There are two motivations. For portions, as in *a coffee and two beers please* and for types, as in *they had two wines at dinner*.

Finally in this section we should note that a goal of the typological investigation of number is to integrate the typology of values (§ 2) with the typology of nominals involved. This is quite a challenge. The point is that Smith-Stark considered only plurals, suggesting that other values, such as the dual, would behave in the same way. If for example, a language has singular, dual and plural, he assumed that the nominals with a dual would be the same as those with a plural. This situation is found, but it is far from being the only possibility. For instance, in Modern Hebrew and in Maltese, there are some nouns with a dual; they are relatively few in comparison with those with a plural, and they are certainly not those at the top of the Smith-Stark Hierarchy. These duals have been analysed as 'minor numbers'; they do not pattern according to the Smith-Stark Hierarchy, but they are counter-examples of a narrowly definable type (Corbett 2000: 95–101). Then there are other apparent number values, which appear to run counter to the hierarchy. For example, associatives, like Central Alaskan Yup'ik *cuna-nku-t* (CHUNA-ASSOCIATIVE-PL) 'Chuna and his family/friends'. These apparent exceptions can be shown to involve an additional category and so are outside the scope of the constraints discussed here (Corbett & Mithun 1996). A discussion of these complications is beyond the scope of this chapter.

5. The expression of number

Having discussed the possible number values, and the nominals involved in the number

system, we now turn to the question of how number is expressed. There are various means available.

5.1. Number words

Some languages have special 'number words', just for the purpose of indicating number. Thus in Tagalog, virtually any constituent can be pluralized by the word *mga* [maŋa], perhaps best characterized as a clitic (David Gil, personal communication):

- (31) *mga bahay*
PL house
'houses'
- (32) *mga tubig*
PL water
'cups/units of water'
- (33) *mga Marcos*
PL Marcos
'Marcoses'
- (34) *mga ma-puti*
PL STAT-white
'white ones'

Further examples of number words can be found in Dryer (1989). Diachronically, number words are a potential source of number morphology.

5.2. Syntax

In § 1 we saw that nominal number may be marked by agreement. It is worth noting that nominals whose number marking and agreements differ will be more regular with regard to agreement than for nominal marking. Thus English *sheep* is exceptional as far as the Smith-Stark Hierarchy is concerned, if we consider its morphology. In English, animates and most inanimates distinguish number, and so we would expect *sheep* to distinguish number, as *goat* does. If we look at its agreements it is fully regular, however:

- (35) *this sheep is grazing*
(cf. *this goat is grazing*)
- (36) *these sheep are grazing*
(cf. *these goats are grazing*)

5.3. Morphology

Number morphology varies from relative simplicity in some languages to great complexity in others. Even in languages where the main patterns are straightforward, we often find isolated examples of more interesting types.

We start from the notion of 'base' (or 'basic inflectional stem'). The base of a lexical item is the form which cannot be further reduced as far as inflectional categories are concerned. English *play* is the base for the forms *play-s* and *play-ed*. Let us consider a language which has at least two numbers, singular and plural. What are the possible relations between the number forms and the base for a given lexical item (or group of lexical items)? Let us start from a maximally general model, shown in Figure 61.8.

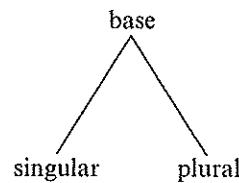


Figure 61.8: Possibilities for number marking

How can the singular and plural forms differ from the base? First they may differ in inflection. Or they can vary from the base through stem formation. These two devices, inflection and stem formation may occur separately or together. The fourth logical possibility is that neither inflection nor stem formation is employed. If this means that the singular form, plural form and basic stem are all identical, then clearly number is simply not marked morphologically for the items in question (as in the case of English *sheep*).

Having considered the possibilities in an abstract way, let us now consider the options in more detail, allowing for different stems:

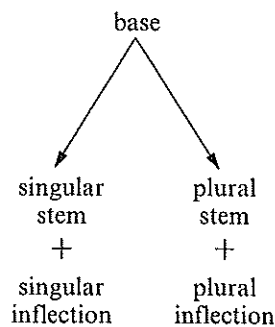


Figure 61.9: Possible stems and inflections

We will now look at examples of number marking, considering in particular whether all the elements identified in the diagram are distinct or not in particular examples. It is important to note that different patterns often coexist within a single language; if an

example is given from a particular language this does not mean that the pattern is the norm for that language.

If we start with the relations between the base and the stems, the first logical possibility is that all are distinct. This possibility can be illustrated by the irregular Russian noun, *xozjain* 'landlord'. The base is *xozja(j)*-, the singular stem is *xozja-in-* and the plural stem is *xozja-ev-*. Both stems allow the normal addition of endings. The extreme type of difference is found in cases of suppletion, where there are different stems which are not related by any regular or irregular type of stem formation; their relation is purely semantic. An example is Russian *čelovek* 'person', plural *ljud-i* 'people'. Note that we are indeed dealing with stems here: *čelovek* 'person', takes normal singular inflections, and *ljud-i* 'people' takes plural inflections.

It is unusual for the root and the singular and plural stems all to be distinct, in Russian and more generally. Often we find that the base and the singular stem are identical, as in this diagram:

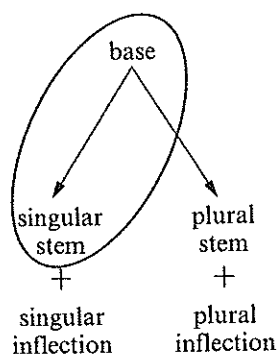


Figure 61.10: Singular stem equals the base

This pattern too can be illustrated from Russian. The noun *krylo* 'wing' has the base *kryl-*, to which the singular endings are added directly (*kryl-o*, *kryl-a*, *kryl-u* and so on). The plural stem is *kryl'j-* (the ' marks palatalization of the preceding consonant), as in the nominative plural *kryl'ja*. Why should we say that there is a distinct plural stem here, rather than that the nominative plural ending is palatalization plus *-ja*? The point is that *-a* is a regular nominative plural ending, found on hundreds of nouns which do not have a separate plural stem. The plural endings for the remaining five cases of Russian are also found on other nouns; we would be missing an obvious generalization if we claimed there

were special endings right through the plural paradigm while in fact nouns like *krylo* 'wing' differ from other nouns only in having a different stem for the plural.

The next possibility is that the plural stem should be the same as the base:

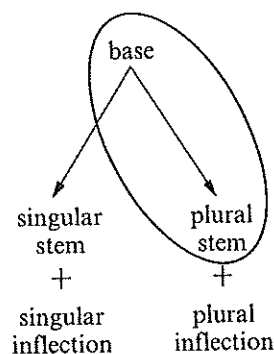


Figure 61.11: Plural stem equals the base

Again the pattern is found in Russian. The noun *bolgarin* 'a Bulgarian' has the base *bolgar-*, and the plural stem is identical, as in forms like the nominative plural *bolgar-y*. The singular stem differs, and is *bolgarin-*. Several nouns denoting nationalities and other social groupings behave in this way.

A final relation of base to stems is that all are identical, diagrammed as follows:

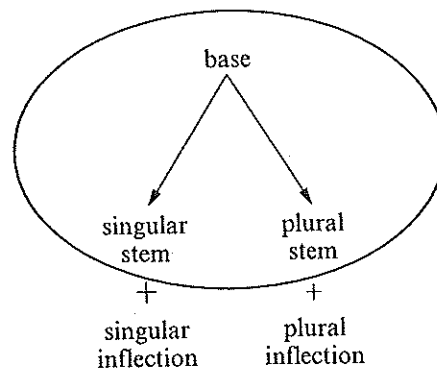


Figure 61.12: Both stems equal the base

This situation is extremely common. Again in Russian we find many nouns like that for 'newspaper', which has the basic stem *gazel-*. The (nominative) singular is *gazel-a* and the (nominative) plural is *gazel-y*. Here, stem formation has no role, and the entire burden of signalling a difference in number is carried by the inflections (endings in this instance).

We move on to look for identities elsewhere in the model. There is a further, ini-

tially rather surprising type of identity, shown in the next diagram:

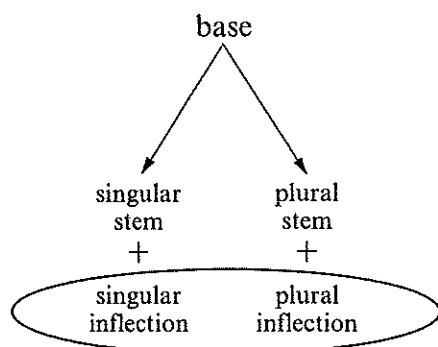


Figure 61.13: Inflections not sensitive to number

This pattern suggests that the inflections used for singular and plural could be identical. This situation regularly occurs in Daghestanian languages for the majority of the large numbers of cases they distinguish (often just the absolutive is an exception). The Akhvakh noun *nido* 'forehead' shows a clear case of identical endings: we take absolutive and ergative endings to illustrate the point:

- (37) Akhvakh (Kibrik 1991)
- | | | |
|------------|-------------------|----------------------|
| | singular | plural |
| absolutive | <i>nido</i> | <i>nido-di</i> |
| ergative | <i>nido-la-de</i> | <i>nido-di-le-de</i> |

In this example the base is *nido*, and the singular stem is identical to it. The plural stem is *nido-di-*. The absolutive case, in singular and plural, has no ending. In both numbers there is an oblique stem, distinct from the basic singular or plural stem; in the singular it is formed with *-la*, and in the plural with *-le-*. The various oblique case endings are added to this stem; in our example the ergative case is given, and the appropriate ending is *-de*. As with the absolutive, the ending is the same for singular and plural. The point is that information about number is signalled by the differences in the stems: *-di-* indicates plurality for this noun, *-la-* shows singular oblique, and *-le-* plural oblique. Thus, a form like *nido-di-le-de* indicates plurality twice. The endings have no role in the number system, their function is to mark the case of the noun (the case system is extensive). This identity of form of endings in the singular and plural is quite general in Daghestanian languages. It is to be distinguished from occasional syncretisms of form involving small numbers of nouns in languages where the coincidence of form is not systematic.

There is one final pattern of identity, noted earlier, which we should consider again, that in which both stems are identical to the base, and where the stems are identical to the forms with endings (that is, there are no endings). This means that the noun is indeclinable – number is not marked morphologically. There are numerous examples of this situation, both of languages where number is not marked morphologically on particular word classes (English adjectives, for example) or not marked morphologically at all. But it may be found for a subset of a word class within a system where number is usually marked morphologically. Thus in Russian, the majority of nouns distinguish two numbers but some, especially foreign borrowings do not. For example, *taksi* 'taxi' may denote one or more taxis (the ambiguity will often be removed by elements showing agreement in number).

Two special types of number marking deserve a mention. The first is 'inverse' number, where the marker for singularity for some nouns is used to mark plurality for others. This is found, for example in Kiowa (Watkins 1984). The other is 'constructed' numbers. Constructed numbers appear where there is a mismatch between number marking of different elements. Consider the following data from Hopi:

- (38) Hopi (Hale 1993: 19)
- | |
|-------------------------|
| <i>Pam wari</i> |
| that.SG run.PFV.SG/DUAL |
| 'he/she ran' |
- (39) *Puma wari*
- | |
|------------------------------|
| that.DUAL/PL run.PFV.SG/DUAL |
| 'they (two) ran' |
- (40) *Puma yu'tu*
- | |
|-------------------------|
| ihat.DUAL/PL run.PFV.PL |
| 'they (plural) ran' |

The pronominal forms on their own make only a two-way distinction, as does the verb. Put together, however, we have a singular-dual-plural system, 'constructed' from the two parts. It must be stressed, however, that this is only a part of the system: animate nouns in Hopi have a straightforward singular-dual-plural system.

This section has attempted to describe the typological space for number marking. But this is just a start: we need to be clear about which forms are the typical ones for given languages, and which occur only sporadically. And then we can investigate the pat-

terns of marking; thus, for instance, we tend to find greater irregularities (and greater use of stem alternations) for the items higher on the the Smith-Stark hierarchy. For pronouns it is common to find suppletion. And yet this patterning is overridden by items such as *geese* and *teeth*, whose distribution in texts between singular and plural is dramatically different from the average (see Tiersma 1982).

6. Conclusion

The category of number remains a challenge for typologists. We need to know more on the relation between nominal and verbal number (§ 1). We now have a fair understanding of the possible number values (sections 2 and 3) and a reasonable idea of the patterns of the types of nominal involved in number systems (§ 4). However, there is a good deal to be done for these two parts of the typology to be integrated. Similarly, while we have a reasonable inventory of the means of number marking (§ 5), we do not know the possible ways in which these means of marking may be distributed over types of nominal. Number, this apparently simple category, is far from being understood.

Acknowledgement

This entry is based on research funded in part by the ESRC (grant R000222419) and the British Academy. This support is gratefully acknowledged.

7. Special abbreviations

ALL	allative
CONT	continuative
HAB	habitual
IRR	irrealis
REDUP	reduplication
STAT	stative
UNM	unmarked

8. References

- Allan, Keith. 1980. "Nouns and countability". *Language* 56: 541–567.
- Blake, Barry J. 1979. *A Kalkatungu grammar* (Pacific Linguistics, series B, no. 57). Canberra: Department of Linguistics, Research School of Pacific Studies, Australian National University.

Brooks, Bryan. 1991. "Pluractional verbs in African languages". *Afrikanistische Arbeitspapiere* (Cologne) 28: 157–168.

Corbett, Greville G. 2000. *Number*. Cambridge: Cambridge University Press.

Corbett, Greville G. & Hayward, Richard J. 1987. "Gender and number in Bayso". *Lingua* 73: 1–28.

Corbett, Greville G. & Mithun, Marianne 1996. "Associative forms in a typology of number systems: evidence from Yup'ik". *Journal of Linguistics* 32: 1–17.

Cowell, Mark W. 1964. *A reference grammar of Syrian Arabic (based on the dialect of Damascus)* (Arabic series 7). Washington, DC: Georgetown University Press.

Crowley, Terry. 1982. *The Paamese language of Vanuatu* (Pacific Linguistics, series B, no. 87). Canberra: Department of Linguistics, Research School of Pacific Studies, Australian National University.

Dixon, Robert M. W. 1980. *The languages of Australia*. Cambridge: Cambridge University Press.

Dixon, Robert M. W. 1981. "Wargamay". In: Robert M. W. Dixon & Barry J. Blake (eds) *The handbook of Australian languages: II: Wargamay, the Mpakwathi dialect of Anguthimri, Watjarri, Margany and Gunya, Tasmanian*, Amsterdam: Benjamins, 1–144.

Dryer, Matthew. 1989. "Plural words". *Linguistics* 27: 865–95.

Durie, Mark. 1986. "The grammaticization of number as a verbal category". *Berkeley Linguistics Society* 12: 355–70.

Foley, William A. 1991. *The Yimas language of New Guinea*. Stanford: Stanford University Press.

Green, Ian. 1989. *Marrithiyel, a language of the Daly River region of Australia's Northern Territory*. Unpublished PhD thesis, Australian National University.

Greenberg, Joseph H. 1963. "Some universals of grammar with particular reference to the order of meaningful elements". In: Joseph H. Greenberg (ed.) *Universals of language*, Cambridge, Massachusetts: MIT Press, 73–113.

Greenberg, Joseph H. 1966. *Language universals: with special reference to feature hierarchies*. The Hague: Mouton.

Hale, Kenneth. 1993. "On the human value of local languages". In: André Crochetière, Jean-Claude Boulanger & Conrad Ouelon (eds.) *Proceedings of the XVth International Congress of Linguists, Québec, Université Laval 9–14 August 1992: Endangered Languages: Volume I*, Sainte-Foy: Les Presses de l'Université Laval, 17–31.

Harrison, Sheldon P. (with the assistance of Salich Y. Albert) 1976. *Mokilese Reference Grammar*. Honolulu: University Press of Hawaii.

Hayward, Dick [= R. J. Hayward]. 1984. *The Arbore language: a first investigation: including a vocab-*

- ulary. (Cushitic Language Studies 2). Hamburg: Buske.
- Hutchisson, Don. 1986. "Sursurunga pronouns and the special uses of quadral number". In: Ursula Wiesemann (ed.) *Pronominal systems* (Continuum 5), Tübingen: Narr. 217–55.
- Kibrik, Alexandr E. 1991. "Organising principles for nominal paradigms in Daghestanian languages: comparative and typological observations". In: Frans Plank (ed.) *Paradigms: the economy of inflection* (Empirical Approaches to Language Typology 9), Berlin: Mouton de Gruyter, 255–74.
- Laidig, Wyn D. & Laidig, Carol J. 1990. "Larike pronouns: duals and trials in a Central Moluccan language". *Oceanic Linguistics* 29: 87–109.
- Lichtenberk, Frantisek. 1983. *A grammar of Manam* (Oceanic Linguistics Special Publication 18). Honolulu: University of Hawaii Press.
- Masica, Colin P. 1991. *The Indo-Aryan languages*. Cambridge: Cambridge University Press.
- Mithun, Marianne. 1988. "Lexical categories and the evolution of number marking". In: Michael Hammond & Michael Noonan (eds.) *Theoretical morphology: approaches in modern linguistics*. San Diego: Academic Press, 211–34.
- Newman, Paul. 1990. *Nominal and verbal plurality in Chadic* (Publications in African Languages and Linguistics 12). Dordrecht: Foris.
- Plank, Frans. 1989. "On Humboldt on the dual". In: Roberta Corrigan, Fred Eckman & Michael Noonan (eds.) *Linguistic categorization* (Amsterdam Studies in the Theory and History of Linguistic Science 61). Amsterdam: Benjamins, 293–333.
- Priestly, T. M. S. 1993. "Slovene". In: Bernard Comrie & Greville G. Corbett (eds.), *The Slavonic languages*. London: Routledge, 388–451.
- Reid, Nicholas J. 1990. *Ngan'gityemerri: A Language of the Daly River region, Northern Territory of Australia*. Unpublished PhD thesis, Australian National University, Canberra.
- Saussure, Ferdinand de. 1971. *Cours de linguistique générale* (publié par Charles Bally et Albert Sechehaye avec la collaboration de Albert Riedlinger). Paris: Payot. [Third edition, first edition 1916.]
- Smith-Stark, T. Cedric. 1974. "The plurality split". *Chicago Linguistic Society* 10: 657–71.
- Steever, Sanford B. 1987. "The roots of the plural action verb in the Dravidian languages". *Journal of the American Oriental Society* 107: 581–604.
- Stone, Gerald. 1993. "Sorbian (Upper and Lower)". In: Bernard Comrie & Greville G. Corbett (eds.), *The Slavonic languages*. London: Routledge, 593–685.
- Tiersma, Peter M. 1982. "Local and general markedness". *Language* 58: 832–49.
- Watkins, Laurel J. 1984. *A grammar of Kiowa*. Lincoln and London: University of Nebraska Press.

Greville G. Corbett, University of Surrey
(United Kingdom)

62. Articles

1. Two approaches to the typology of articles
2. Identifying articles
3. Articles derived from demonstratives
4. Articles from other sources
5. Co-occurrence with other determiners
6. Special abbreviations
7. References

1. Two approaches to the typology of articles

The typology of articles has been approached in essentially two ways. The approach adopted by Grasserie (1896) and Krámský (1972) begins with a definition of definiteness and then identifies and classifies various grammatical phenomena which express definiteness in natural languages. The resulting typology is not limited to articles but includes other grammatical phenomena such as word order and verbal agreement. Semantically,

such a typology is based on the simple bilateral opposition 'definite' vs. 'indefinite', which is assumed to be universal. Consequently, the typological parameters are exclusively formal. Krámský (1972), for example, subdivides his list of definiteness expressing phenomena according to the following two parameters: the morphological shape of the definiteness markers (independent word, clitic, affix, non-segmental phenomena (order, stress)) and the number of definiteness markers (definite and indefinite, only definite, only indefinite).

The second approach, pioneered by Greenberg (1978), is based on the observation that articles are historically derived from other elements such as demonstratives and numerals. That is, articles are viewed as stages in the adnominal grammaticisation of these elements, the typology of articles then

being concerned with identifying different paths of grammaticisation in which articles occur and classifying the different stages of their development on the basis of semantic and formal criteria.

Probably the best-known grammaticisation path for articles is Greenberg's somewhat rudimentary grammaticisation path for articles derived from demonstratives:

- (1) DEM → DEF.ART → SPEC.ART → NOUN
MARKER

This path may be used briefly to preview some major features of the grammaticisation approach to the typology of articles and to highlight some salient differences with regard to the definiteness-based approach.

Most importantly, from a grammaticisation point of view, definiteness, though undeniably of central importance to the grammar and typology of articles, is only one of a number of meanings that can be conveyed by articles. Thus, for example, specific articles indicate specificity rather than definiteness. Definiteness and specificity are only two salient (and crosslinguistically well attested) stages on the grammaticisation path in (1). Given a sufficiently fine-grained set of criteria, many more stages could be distinguished. Thus, on the one hand, the grammaticisation approach provides a more detailed typological grid for articles and includes a variety of phenomena not considered in the definiteness-based approach (further details in § 3 and 4). On the other hand, the grammaticisation approach excludes some of the phenomena dealt with in the definiteness-based approach, in particular those grammatical structures which appear to be sensitive to definiteness but are not part of a grammaticisation path for articles (such as word order, case markers or verbal agreement).

Furthermore, the grammaticisation approach highlights the problem of distinguishing demonstratives from articles (and various types of articles from each other). In order to establish a grammaticisation path for articles, explicit and operational criteria have to be provided for determining that a given element is no longer a demonstrative but a definite article. This problem is addressed in § 2.

It should be noted that both approaches converge in that they both address the issues of classifying definite and indefinite articles according to their different morphosyntactic shapes (see in particular § 3.2) and according

to the paradigmatic relations they enter into (see §§ 4.1 and 5).

The present overview of the typology of articles basically follows the grammaticisation approach as sketched in Greenberg (1978) and further elaborated in Himmelmann (1997) and Laury (1997). It is primarily concerned with articles derived from demonstratives, the most common and widespread type in natural languages.

2. Identifying articles

This section presents some formal (morphosyntactic) and semantic criteria for distinguishing articles from related elements, in particular the elements they are historically derived from, such as demonstratives and numerals (some of these criteria are also relevant to distinguishing different types of articles).

Two essential, though not sufficient formal criteria are as follows. First, articles are grammatical elements which occur only in nominal expressions. Second, their position within such expressions is fixed (i. e. they occur either to the left or to the right of a noun but not alternatively on either side). At first sight, these may seem to be rather trivial criteria. But note that they distinguish articles from, for example, demonstratives in many languages. Demonstratives (more precisely: deictic elements) may appear in at least four different syntactic functions, i. e. as pronouns, as adverbials, as predicates, and as adnominal modifiers. In the first three of these functions they are not in construction with a noun (cf. Fillmore 1982: 47 f., Diessel 1999, Himmelmann 1997: 125–130). Furthermore, in many languages (e. g. Swahili, Latin, many Australian languages) demonstratives may occur on either side of a noun when used adnominally (cf. Dryer 1989). Thus, an element which can occur by itself, or one the position of which in nominal expressions is not fixed is, in general, not an article.

A further useful heuristic for identifying articles is obligatoriness in grammatically definable contexts. Such contexts include superlatives (*the greatest singer*), nouns taking complement clauses (*the fact that they lost the game*), and the fact that count nouns cannot be used in core argument positions without a marker for definiteness or specificity.

High textual frequency, often invoked as an argument for claiming that a particular el-