Journal of Universal Language 6
September 2005, 1-37

Numerals and their Position in Universal Grammar

Mark Donohue
Australian National University Singapore

Abstract

One of Greenberg's claims concerning the appearance of numerals as modifiers concerns the order of the elements, suggesting that post-nominal numerals are "favored in indefinite and approximate constructions". Examining a small but representative body of data from a variety of languages, and one of the common grammaticalisation targets of singular numerals, the indefinite article, we conclude that the generalisation does not have a universal status, and exists at best as a statistical universal.

Keywords: word order, numerals, Greenberg, lexical classes

1. Greenberg's Ideas on Numerals

Greenberg (1978) listed many universals concerning number systems (how to build complex numbers form smaller units), or the marking of number on nouns and verbs, but surprisingly few

universals dealing with the use of numerals in the phrase or the clause. The universal that I shall examine in this article is Greenberg's universal 44, concerning the position of the numeral in a phrase modifying a head (1978: 284). This universal reads as follows:

> 44. The order noun-numeral is favored in indefinite and approximate constructions.

Here we have a claim that, when a language allows both Numeral-Noun and Noun-Numeral as orders for the modification of a noun by a numeral, the order Numeral-Noun will be used for a more exact, definite meaning, and that the post-nominal use of a numeral is associated with a less exact number. It is the aim of this paper to show that there are well-attested counter-examples to this claim, spread over a number of widely-separated areas of the world's surface. While this is damaging for the proposed universal, the lack of consistency is perhaps understandable because of the confused position of numerals in any semantic 'map' of different predicate types. An additional goal of this paper is to present data from some languages of eastern Indonesia and Papua New Guinea, areas that are otherwise under-represented in discussions of typology and universals of language.

I shall start the exposition with an examination of the notion of 'numeral' as a valid syntactic category, cross-linguistically. While there are some lexical types that show consistent behaviour relative to other word classes in a given language, even though the morphosyntactic realization of that behaviour may vary from language to language, this is not the case for numerals. This lack of consistent relative behaviour shows that 'numeral' is not a coherent entity in terms of the linguistic semantic 'space' that we might claim the class to occupy.

Following this I present information from several unrelated languages in which both orders are possible, and in which the post-nominal use is clearly more exact. It is true that in some cases the post-nominal use of a numeral in these languages is arguably not functioning as a counting unit any more, but rather as a means of identifying one unit from a set of (perhaps disparate?) items, but importantly the fact remains that it is a numeral in a post-nominal position functioning to indicate a more definite and exact construction than the same numeral in a prenominal position. Finally, some languages are reported with a numeral functioning to indicate the number of tokens of the head noun in either pre- or post-nominal position, and with a more definite, exact meaning when occurring post-nominally.

2. The Nature of Numerals

Where do numerals fit in a schema of different semantic types of lexical concepts? Or, rather, *do* numerals fit in a schema of different semantic types of lexical concepts? As a starting point, consider the following table, from Croft (1991: 65), that sets out some morphosyntactic properties typical of different semantic 'prototypes'.

Table 1. Semantic Properties of Prototypical Lexical Classes

	Objects	Properties	Actions
Valency	0	1	≥1
Stativity	state	state	process
Persistence	persistent	persistent	transitory
Gradability	nongradable	gradable	nongradable

The rationale behind this table is that we can identify (at least) three different 'prototypes' of lexical items, which can be verified according to certain semantic properties they display. Note that these are *not* lexical categories, or word classes: a 'verb', for instance, might be either an action or a property, and a property might be coded in a particular language as an adjective, or a verb, or (less commonly) a nominal.

Determining lexical categories, according to Croft (1991), follows from an examination of the morphological coding of these different semantic prototypes in different discourse functions. Table 2, adapted from Croft (1991: 66ff), shows the application of this methodology to English data.

Table 2. Matching Semantic Types to Discourse Functions in English

	Reference	Modification	Predication
Objects	UNMARKED NOUNS	genitive, adjecti-valizations, PP's on nouns	predicate nominals
Properties	deadjectival nouns	UNMARKED ADJECTIVES	predicate adjectives
Actions	action nominals, complements, infinitives, gerunds	participles, relative clauses	UNMARKED VERBS

When we examine the cells of this table in a language like English, we find that the predictions are borne out: nouns must appear with *'s* when they modify other nouns (*dog**'s** house*), and with a copular verb when they are predicative (*That **is** a dog*). The same copular is use with predicative adjectives (*That dog **is** big*), which require the filler 'one' when they are used referentially (*a big **one***). A verb uses the same dummy NP head, and often requires a gerundive form (*runn**ing one***), which is used for modification in the

NP (alternative, a morphologically explicit nominal derivation is used: *runn**er***). The interested reader is referred to Croft 1991, especially pages 66-67, for further discussion of this approach to the determination of lexical categories. I shall simply note that an approach such as this produces an easily testable answer to the question of lexical category membership that is more 'rounded' than, for instance, simply examining one morphosyntactic property in one function (e.g., the use or non use of copular verbs when a non-verb is predicative). Croft has more recently proposed modifications to this schema; Baker (2003) presents an alternative, more formal, approach to lexical categories.

Of the three semantic prototypes listed here, numerals share some of the characteristics associated with prototypical 'properties', in having a valency of 1 and being stative, and some of those associated with prototypical 'actions', being only marginally persistent and not easily gradable. Numerals bear little, if any, resemblance to 'objects'. We have to look at the morphosyntactic behaviour of numerals in order to decide where they are assigned, language by language, in terms of syntactic categories. This should tell us something about their tendencies towards being morphosyntactically coded in ways similar to 'properties' (that is, in the lexical class 'adjective') or similarly to 'actions' (that is, in the lexical class 'verb').

2.1. Numerals as Verbs

Numerals frequently function as verbs if anything that modifies inside an NP is a verb. The morphology used when numerals appear in this discourse function, or (in the extreme) the inability to appear as an NP-internal modifier indicates this. Evidence can be found across a range of languages. I shall present a brief example from Koasati, a Muskogean language of the south-eastern United States (all data are from Kimball 1991). Koasati is an SOV language with a

complex agreement system on its verbs that more closely monitors semantic type than syntactic roles, and which has a switch-reference and case marking system that shows nominative-accusative contrasts. In Koasati (Kimball 1991: 358) numerals cannot appear as modifiers inside the NP, and if they appear in a clause they must be the main predicate of that clause. In example (1) we can see that in order to express the numeral '12' (though other numerals, both lower and higher, show the same behaviour) as a property modifying the object of 'see', Koasati speakers must encode a complex sentence with two clauses, separated by the switch reference marker *-n*. The use of this morpheme to join the clauses indicates that the two clauses do not share the same subject; the subject of 'be twelve' is interpreted as the object of 'see' in the second clause by a set of 'minimum-distance' discourse principles. The structure of (1a) is shown schematically in (1b).[1]

(1) Numeral Predicate in Koasati

a.	Ná:ni-ha	pokkó:l	awáh	tóklo-n	hí:ca-li-:s.
	man-PL	ten	and	two-SW	see-1SG.PAST
	'I just saw twelve men.'				

[1] I have in general followed the glossing conventions of the sources cited. The following abbreviations have been used in portmanteau forms: 1, 2, 3: first, second and third person. SG, DU, PL: singular, dual and plural number. III, VI, VII, VIII, X, XII: class numbers in Arapesh. The following additional abbreviations are used: A: absolutive, AN: animate, ART: article, AUX: auxiliary, CLF: classifier, DEF: definite, DIM: diminutive, E: ergative, FEM: feminine, GENR: generic tense, HOUSE: 'house' class, HUM: human, INAN: inanimate, INSTR: instrumental, LOC: locative, LOWER.AN: lower animate, M, MASC: masculine, N, NEUT: neuter, NOM: nominative, NP: noun phrase, OBJ: object, P: most patient-like argument in a clause, PA: epenthetic syllable, PAST: past tense, PF, PRES: present, PRF: perfective, POSS: possessive, R: realis, RC: relative clause, RED: reduplication, REL: relativiser, REL.PART: relative participle, SI: S,A infix, SUBJ: subject, SW: switch-reference marker (different subject).

b. [(the) men (were) twelve] and [I saw (them)]
SUBJ$_1$ SUBJ$_1$≠SUBJ$_2$ SUBJ$_2$

Koasati represents an extreme example of a language that only allows numerals to be coded as the predicates of the clauses in which they occur. In other languages we can observe a cline in terms of the degree to which a numeral must be coded as a verbal predicate. Table 3 shows the morphosyntactic distribution of numerals in different Austronesian languages, arranged approximately from east to west. In all cases there are (at least) some property concepts which show less verb-like behaviour than do numerals (Data from Mosel & Hovdhaugen 1992, Bauer et al 1993, Donohue 1999a, Sneddon 1996). In Samoan, a Polynesian language from the central Pacific, 'numerals form a sub-class of verbs', and, like more prototypical 'action' verbs, can be found either as predicates or as modifiers, though the use of a numeral as a modifier in an NP is rare, and requires that the numeral appear in a relative clause, just as do verbs. Recall that in Koasati the verb-modifying-noun option that Samoan displays was not allowed; numerals can only be coded as predicative verbs.

Table 3. Behaviour of Numerals in Austronesian Languages

	Samoan >	Maori	> Tukang Besi	> Indonesian
Predicative	as verb	as verb	as verb	as verb?
Modification	in RC	(special)	as 'adjective'	as 'noun'

The appearance of numerals in NPs in relative clauses in Samoan is grammatical, but dispreferred. Mosel and Hovdhaugen (1992: 115) state that 'the most common way to refer to a certain number of specific items is to employ the cardinal numeral as the predicate of an independent clause', as in (2). Example (3) shows a numeral modifying inside an NP; note that the other modifiers in this NP, the possessor and the adjective *laiti* 'small' do not require coding in a

relative clause to function as modifiers (Mosel & Hovdhaugen 1992: 115, 319).

(2) E tolu tusi [RC na maua].
GENR three letter PAST get
'(She/he) got three letters.'
(That is, 'the letters that she/he got are three')

(3) nai o=u tei
DIM.PL POSS=1SG small.sibling
laiti e toa=lua.
little.PL GENR HUM=two
'my two little sisters'

Numerals in NPs in Maori uses a construction that is not replicated for any other semantic types, and is similarly dispreferred in favour of predicative numerals, rather than modificational ones. In Tukang Besi the behaviour of numerals is complex, and will be discussed in more detail in 2.3.2; a numeral shows the coding properties of an adjective when it modifies inside an NP, and shows the morphosyntax of a verb when it is predicative.

In these languages there is clearly no difference between numerals and action concepts when predicative, in which position they all function as verbs. When modifying inside the NP the numerals of Samoan show all the trappings of verbs, appearing in relative clauses. In Tukang Besi and Indonesian the modifying numeral is not a verb, but does show unusual properties, indicating that there are special constraints on this property concept compared to others. Both Tukang Besi and Indonesian shall be discussed in greater detail later in this article.

2.2. Numerals as the Opposite of Verbs

In addition to very verb-like numerals, we also find case in which, even though some other property concepts show verb-like behaviour when modificational, numerals escape this requirement. The following sections shall present some case studies.

2.2.1. Pengo

Pengo is a Dravidian language from Central India, and has numerals that are less verb-like than other modifiers in an NP. (The examples have been drawn from Burrow and Bhattacharya 1970: 93, 48, 54 and 49). In (4) we see the relative participle form used with a modificational verb, and in (5) the same morphology used for a modificational property concept. In (6), on the other hand, the demonstrative simply shows number and gender agreement, and not the relative participal, and this same strategy is the one used when a numeral is found modifying the noun, as in (7).

(4) Modificational Action within an NP
ēnd-ni himṇaŋ
dance-REL.PART children
'dancing children'

(5) Modificational Property within an NP
ṛinj-ni puy
white-REL.PART flower
'white flower.'

(6) Modificational Deictic within an NP
adel kogle
that.SG.FEM woman
'that woman'

(7) Modificational Numeral within an NP
rinḍ-ek kogle-k
two-FEM.PL women-PL
'two women'

Compare (5) with (8), which shows 'white' in a predicative function (Burrow and Bhattacharya 1970: 48). Here we the property concept behaves morphosyntactically as a verb when it is predicative, as well as when modificational, indicating clearly that 'white' is lexically a verb, and not a member of a separate adjective class.

(8) Predicative Property Behaving as a Verb
Puy ṛinj-nat
flower white-PRES.3NEUT-MASC
'(The) flower is white.'

The patterns of morphosyntax used in NP-internal modification are summarised in table 4. 'Actions' at the left of the table refers to the coding of those lexemes that are typically encoded as verbs, 'actions'; when they modify nominals in Pengo they employ a particular relative clause construction. Properties (which might be expected to be coded as an 'adjective' cross-linguistically) use this same strategy, but additionally mark the gender of the head noun. By contrast, demonstratives code gender, but do not use the relative clause construction. Numerals and demonstratives show less verb-like behaviour than do property concepts such as colour terms.

Table 4. Verb-like Behaviour in Modification in Pengo

	Actions	Properties	Dem	Numeral
CLASSIFIER				
GENDER				
RELATIVE CLAUSE				

2.2.2. Australian Languages: Wangkajunga

Languages such as those covering most of Australia are not a counter-example to this argument that, among the different classes of modifiers within the NP, numerals show morphosyntactic behaviour that is more similar to verbs than that shown by other types. In most languages of Australia almost everything that can occur modifying a noun is a nominal, and is treated in much the same way as nominals. Some examples from Wangkajunga, the northernmost variety of the Western Desert language, spoken near Fitzroy Crossing, in Western Australia (Jones 2002), illustrate this point. In (9) (Jones 2002: 134) we can see the occurrence of the ergative case *-lu* on both the head noun, *tuju*, and the numeral. The fact that the second-position clitic *-pula* occurs following both *tuju* and *kujarra* is evidence that the two words do form an NP, since the entire NP counts as a single unit in the clause, and not as two independent 'nominal' elements.

(9) Tuju-lu kujarra-lu=pula parnaparnti
woman-ERG two-ERG=3DL.S goanna
pung-in-pa.
hit-PRES-PA
'Two women are killing a goanna.'

A numeral may stand alone as an NP, as with any other nominal. This is shown in (10). In (10) *kujarra* is the only element of the NP

other than the derived, modifying, *pitikurlu*. In (11) we can see *kujarra* occurring as a modifying element within the subject NP (Jones 2002: 75, 270).

(10) Ya-nu-pula piti-kurlu kujarra.
go-PAST-3DL.S coolamon-HAVE two
'The two with coolamons went.'

(11) Warta-ngka pirrpilya-ngka jii-ngka-pula
tree-LOC branch-LOC DEM-LOC-3DL.S
turru kujarra nyin-in-pa.
Bird two sit-PRES-PA
'Two birds are sitting on a branch of that tree.'

When predicative a numeral is similarly treated in the same way as any other nominal. This lack of differentiation between numerals and other elements of an NP, including nouns themselves, means that these languages do not enter into a typology of different behaviour within an NP.

2.2.3. Arapesh

Arapesh is a Torricelli language from northern Papua New Guinea. The following examples have been drawn from Fortune (1942: 55, 57, 56, 54, 94, 89), but are completely compatible with more recent work such as Conrad & Wogiga (1991). We can see the morpheme *i*, glossed as 'relativiser', is common in a range of modificational constructions, including both clear relatives clauses with action concepts, such as that seen in (12), as well as modification by property concepts such 'red' in (13). Other examples show that colours (and other property concepts), unlike action concepts, do not require the relativiser, and so property-

concept modification must be considered as a separate construction type from action-concept modification. This can be seen in the example in (14), in which *i* does not occur.

(12) jawei b-i
bad VI.PL-Pro
[RC sigaliuh wha-b-ah i]
white.cockatoo.XII.SG XII.SG.SUBJ-VI.PL.OBJ-ate REL
'The (coconut)s that the white cockatoo ate are bad.'
(glossing literally, 'They are bad [the __ which white cockatoos ate them]')

(13) araman n-auh-i
man.VII.SG VII.SG.SUBJ-red-REL
'red man'

(14) araman bagara-ni
man.VII.SG white-VII.SG
'white man'

Genitive constructions employ the same morpheme *i*, but with affixation to index the class and number features of the possessor.

(15) mbul i-g berag
pig.X.PL REL-III.SG head.III.SG
'(the) pig's head'

Finally modification by a demonstrative or a numeral does not employ the relativiser, but simply requires that the modifier agree in class and number with the noun it modifies, as shown in (16) and (17).

(16) aguhuda eguh̰
these.X.PL fish.X.PL
'these fish'

(17) eɲeɲ kwodzar
one.VIII foot.rope.VIII.SG
'one footrope'

The different types of morphosyntactic behaviour observed when different semantic types of modifiers are found inside an NP are summarised in table 5; the two different types of property concepts have been coded in two separate columns. Here it is again clear that numerals features on the opposite end of the continuum from verbs, as embodied in action concepts.

Table 5. Verb-like Behaviour in Attribution in Arapesh

	Actions	Prop. I	Prop. II	Nouns	Dem, Numeral
CLASS					
REL PRONOUN					
SUBJECT					

2.2.4. One

If we examine another Torricelli language, this time One, the westernmost language of the family (Arapesh is found in the east of the Torricelli ranges), we find another pattern with respect to the treatment of numerals, but one in which the numerals are clearly on the nominal end of the continuum. There are no morphological differences between modificational and predicative uses of different words, and so presenting data on the modificational uses will be

fully revealing.

When a verb is used, it takes not only a prefix that marks person and number of the subject of that verb, but also (for most verbs) a suffix that indicates plurality of the absolutive argument in the clause.

(18) a. meli n-uka-lo apa
children 3PL-tie-PL rope
'the children tying the rope'

b. mala y-uka apa
child 2/3SG-tie rope
'a child tying the rope'

Adjectives do not take an agreement prefix, but can take the absolutive suffix. Not all adjectives mark agreement with suffixes, as seen in (19) and (20).

(19) a. meli tafla-lo
children cold-pl
'the cold children'

b. mala tafla
child cold
'the cold child'

(20) a. meli napo
children big
'the cold children'

b. mala napo
child big
'the cold child'

Matching the no-agreement pattern seen with adjectives such as *napo*, we find that demonstratives, numerals, and possessive modification are all used without any agreement with the head. Only the genitive construction used with possession shows any distinction, and this is not obligatory for prenominal possession (see Donohue 2000 for further details of NP structure in One).

(21) a. meli nu
children that
'those children'

b. mala nu
child that
'that child'

(22) a. meli plana
children two
'two children'

b. mala ara
child one
'one child'

(23) a. meli i-enu
children 1SG-GEN
'my children'

b. mala i-enu
child 1SG-GEN
'my child'

The patterns seen in (18)-(23) are summarised in table 6. As in Arapesh, numerals are found at the opposite end of the continuum from verbs, though at least (compared to Arapesh) they are not less verby than nouns.

Table 6. Numerals as the Opposite of Verbs in One

	verbs	some verbs, some properties	other properties, numerals, demonstratives, nouns
NUMBER			
PERSON			

2.2.5. Olo

Olo is the third Torricelli language to be examined, the eastern neighbour of One but in a separate subgroup of Torricelli. In example (18), from McGregor and McGregor (1982: 53), we can see that some property concepts, such as *afo* 'fat', show the same inflectional requirements as do verbal actions like the predicative *ello* 'jump up'. Others, such as *namën* 'green', take no special marking when modificational, and numerals agree with the head noun only in gender, not in person and number. Again we have a case of a language in which the numerals show less verb-like behaviour than many other modifiers, though they do for the first time in this short survey of Torricelli languages show a pattern that is more 'verby' than are nouns and some properties. (Demonstratives

have not been exemplified in this example, but show gender and number agreement without person, as shown in table 7.)

(24) sie wëngkes t-afo-tei namën
grasshopper two.M 3DU.M-fat-very green
t-ello
3DU.M-jump.up
'Two very fat green grasshoppers jump up.'

Table 7. Verb-like Behaviour in Attribution in Olo

	verb, (some properties)	Dem	Numeral	other properties, nouns
GENDER				
NUMBER				
PERSON				

2.2.6. The Non-verbiness of Numerals

We have seen that not only are numerals sometimes nominal (Australia), and sometimes extremely verb-like (Austronesian, North America), and that different languages have different ways of sorting the position of numerals in semantic space. We cannot even posit a cline, something that could be stated as an implicational universal of the form "if any properties are treated as verbs morphosyntactically, then numerals are treated as verb." Numerals simply do not fit into 'semantic space' in any consistent way, with respect to other sorts of semantic modification types.

2.3. Influence from and to Other Modifier Types

The morphosyntax of different modificational constructions may influence the patterns found with numeral modification. Some of the

clearest examples are found in languages in which the genitive construction, prototypically used with possession, has extended to include other modificational constructions. It has been widely demonstrated that genitive morphosyntax can be found in many parts of the grammar, not simply their semantic 'core', possessive constructions. Partitive constructions usually involve a genitive construction ('two of the books', compare with 'covers of the books'), but more indiscriminate use of the genitive in numeral expressions is also found.

2.3.1. Genitive Influence

In Supyire (a Gur language, spoken in Mali and Burkina Faso; data from Carlson 1994: 202, 207) the numeral construction is basically a variety of the genitive, using the 'possessed tone' (which, amongst other allotones, changes (M) M → (L) H) on the possessor in a semantically possessive construction, and on the numeral in a numeral construction. The basic tones of these modifying items used here are *mEge* for 'the name' and *kE* for 'ten'). (Quantifiers show the same morphosyntax as numerals).

(25) Possessive Modification within an NP
kàn-he mɛ̀-gé
village-DEF name-DEF:POSSESSED.TONE
'the village's name'

(26) Numeral verb within an NP
cyèe kɛ́
women ten:POSSESSED.TONE
'ten women'

West Makian, a West Papuan language in the North Halmahera group from Eastern Indonesia (Voorhoeve 1982: 25, 36) Has the

order Noun-Numeral, or alternatively Numeral POSS Noun-Numeral. A simple possessive construction is shown in (27), employing the possessive marker *de*, which shows agreement for the features of the possessed item. In (28) the numeral follows the noun it modifies, and so does not require the use of a genitive construction, though the numeral does still necessarily agree with the noun in terms of class, here lower animate. In (29) the noun *oma* is doubly modified by the numeral *meminyé*, once prenominally in a genitive construction identical to that that seen in (27) and once postnominally, just as the numeral in (28) appears. From the translations provided in Voorhoeve (1982) it is hard to judge the relative definiteness or specificity of the constructions shown here in (28) and (29), but it appears that (28), with the single postnominal numeral, shows a greater degree of definiteness than does the doubly-marked noun in (29). If this assessment is accurate, it would constitute a counter-example to Greenberg's proposed universal 44, as detailed in section 1. Further examples of counter-examples will be presented in section 3; section 3.3, describing Fore, is particularly comparable to the West Makian data presented here.

(27) namu de esi
chicken 3SG.INAN.POSS egg
'chicken eggs'

(28) mo oma meminyé
3SG.AN.POSS child LOWER.AN.one
'her one child'

(29) situ terahir meminyé
until at.last LOWER.AN.one
do oma meminyé
3SG.INAN.POSS child LOWER.AN.one
'Until at last one child…'

Despite the evidence of numerals appearing modificationally in genitive constructions, to argue for a special link between the genitive construction and numerals is doomed to being inconclusive. Genitive constructions find their way into almost all other construction types about, so it is not surprising to detect their influence here as well. The conclusion that we can draw from these data is that the mapping of numerals as a class into a map of semantic prototypes shows a high degree of semantic plasticity. This is evidenced by the fact that numerals can appear with genitive constructions in some languages, a morphosyntactic strategy that is prototypically associated with nominals, the prototype of which is an object, while in other languages numerals appear modificationally with the same morphosyntactic treatment as verbs, the prototype of which is an action. If numerals can be satisfactorily mapped into an idealised 'semantic space', then they are clearly not linked, cross-linguistically, to any particular region in this map, but are subject to influence from a variety of different constructions.

In the following section I shall examine some cases in which numerals appear as ends of a continuum from nominal to non-nominal modification.

2.3.2. Quantifier Influence

We can also ask about the influence of numeral morphosyntax on that of other lexical classes. Not surprisingly, quantifiers show the greatest tendency towards 'contamination' from numerals. The obvious semantic correlations between quantifiers and numerals do not need to be belaboured here, but shall be exemplified with a pair of case studies.

Sahu is another North Halmaheran language, related to West Makian, which shows an interesting spread of both classifiers and genitive marking into the modificational numeral construction (data from Visser & Voorhoeve 1987: 37, 52, 44). (30) shows a classic

use of the genitive construction to mark possession, while (31) shows that numerals modify nouns by agreeing in the feature of noun class. A quantifier shows both possessive marking with *ma* as well as prefixal class agreement, as seen in (32).

(30) o nunu'u ma ngowa
ART.N dog POSS child
'a dog's pups'

(31) wala rimói
house CLF.HOUSE.one
'one house'

(32) nanau'u ma nga-repe
men 3SG.POSS CLF.HUM-many
'many men'

Table 8. Sharing Morphology in Sahu

	Possession	Quantifier	Numeral
GENITIVE	→		
CLASSIFIER			←

In Tukang Besi, an Austronesian of Southeast Sulawesi in Indonesia (Donohue 1999a, and also Donohue 1999b, Baker 2003 for data refuting the morphosemantic correlates of lexical types and discourse function proposed in Croft 1991), we find a set of numerals which must undergo partial reduplication to appear in any discourse function, whether it be reference, modification, *or* predication. If a numeral appears in its unreduplicated form, it behaves as a quantifier, floating in the clause and not within the NP that it modifies. Examples (33) and (34) show that heading an NP is only an option for a reduplicated numeral, (35) and (36) demonstrate the ungrammaticality of an unreduplicated numeral appearing

modificationally in an NP, and (38) and (39) show that for a numeral to function as a predicate requires that is show reduplication.

Reduplicated Numeral Functioning as Head of an NP

(33) No-rato=mo na to-tolu atu.
3R-arrive=PF NOM RED-three that
'Those three arrived.'

(34) * no-rato=mo na tolu atu.
3R-arrive=PF NOM three that

Reduplicated Numeral Functioning Modificationally

(35) No-moturu=mo na mia to-tolu iso.
3R-sleep=PF NOM child RED-three yon
'Those three children are (finally) asleep.'

(36) * no-moturu=mo na mia tolu iso.
3R-sleep=PF NOM child three yon

It is grammatical for a non-reduplicated numeral to appear modificationally in the NP (or right-adjoined to it), but only if it is affixed to a classifier.

(37) No-moturu=mo na mia tolu-mia iso.
3R-sleep=PF NOM child three-CLF yon
'Those three children are (finally) asleep.'

Reduplicated Numeral Functioning Verbally

(38) To-to-tolu='e!
1PL.R-RED-three=3P
'Let's do a third one!'
(literally, "Let's make them three!")

(39) * to-tolu=’e!
1PL.R-three=3P

In (40) and (41) we see a reversal of this near-obligatory derivation by reduplication, in that a reduplicated numeral cannot appear noncontiguously separated from its NP while a plain numeral, or numeral with classifier, can.

Non-reduplicated Numeral Floated from its Referent
(40) * totolu(-mia) no-moturu=mo (na ana).
three-CLF 3R-sleep=PF NOM child
‘The three (children) are asleep now.’

(41) Saba’ane/tolu(-mia) no-moturu=mo (na ana).
all /three-CLF 3R-sleep=PF NOM child
‘All of the(m)/The three (children) are asleep now.’

These results are significant; if quantifiers can be shown to share both genitive and ‘numeral’ types of constructions, we can posit their being ‘intermediate’ between objects and numerals. Of course, the problems raised earlier remain: there does not seem to be a consistent ‘semantic position’ occupied by numerals cross-linguistically.

Note that while the numeral appears in a ‘non-basic’ form when it modifies a noun in an NP, it does not appear with the same subordinating morphology that is expected of verbs in this position; the lack of any morphological marking of subordination is in fact more closely reminiscent of the behaviour of adjectives in this function.

The following sentences illustrate the fact that a verb must appear in a relative clause, marked with either the infix *-um-* or the prefix *i-*/*di-*/*ni-* (depending on the role of the head of the relative clause in the relative clause, see Donohue (1999: chapter 15). This is

shown in (42) and (43) with the verb *rato* 'arrive' (see also (33) earlier for a predicative use of this verb, in which *-um-* is not required). By contrast, an adjective such as *kalu* 'be tired' may modify a noun directly without any subordinating morphology. It is possible for *-um-* to appear with a modificational adjective, but in this case the adjective acquires a superlative reading, as seen in (45).

Verb Functioning Modificationally

(42) No-moturu=mo na ana [RC r<um>ato] iso.
3R-sleep=PF NOM child arrive< SI> yon
'Those children who arrived are (finally) asleep.'

(43) * no-moturu=mo na ana rato iso.
3R-sleep=PF NOM child arrive yon

Adjective Functioning Modificationally

(44) No-moturu=mo na ana kalu iso.
3R-sleep=PF NOM child tired yon
'Those tired children are (finally) asleep.'

(45) No-moturu=mo na ana k<um>alu iso.
3R-sleep=PF NOM child tired< SI> yon
'Those most tired children are (finally) asleep.'

Compare the use of overt subordinating morphology with verbs, but not adjectives, when they modify a noun with the use of a numeral to modify a noun, seen earlier in (35)-(37). When the numeral appears with a classifier, no further morphology is required. When the numeral appears without a classifier it must be (partially) reduplicated, but does not have to (and, indeed, cannot) appear with the infix *-um-* that is found with verbs in this function.

2.4. Conclusions: the Semantic 'Type' of Attributive Numerals?

Numerals are sufficiently idiosyncratic as to defy any simple mapping of their semantic type.

The only defensible conclusion that we can draw is that those (less theoretical, more traditional) grammars that included a separate class of 'numerals' in their description of 'parts of speech' are quote accurate. Ascribing numerals to a particular semantic grouping is not accurate, either language-internally or cross-linguistically, as they often show different morphosyntactic behaviour in different discourse functions, and often show the morphosyntax of a combination of different construction types or the ability to appear in different construction types.

3. Post-nominal Numerals and Unique Reference

In the following sections I shall present data from a range of languages about the meaning inherent in different positions of the numeral with respect to their head noun in languages in which both pre-nominal and post-nominal numerals are found. The examples are by no means exhaustive, but are representative of the kind of variation found cross-linguistically. In these different languages we find ample documentation of the fact that the post-nominal position is used for more highly specific references. This data is obviously a problem for the predictions of Greenberg's Universal 44, as described in section 1.

3.1. Indonesian

Indonesian (a major Austronesian language of South-east Asia)

(Sneddon 1996) typically has post-nominal modification, with the exception of quantifiers and numerals, which precede the noun; quantifiers cannot follow a noun, but numerals can, in certain specific situations. Examples of noun phrases with different types of modification are shown in (46)-(49).

(46) kota besar itu
city big that
'that big city'

(47) orang yang sudah tiba
person REL already arrive
'the people who have already arrived'

(48) kota besar
city big
'big city'

(49) (pada) suatu jam
at an hour
'sometime'

Compare (50)-(52); in (50) we can see the use of *suatu* as a development of *satu* into (roughly) an indefinite article, which must be prenominal. (51) shows the use of a numeral pre-nominally, and may be more or less exact. Finally, in (52) we see that the post-nominal use of the numeral refers to an exact time: no longer one or some hours, but uniquely identifying the one particular hour.

(50) (pada) suatu jam
at an hour
'sometime'

(51) satu jam
one hour
'(about) one hour'

(52) jam satu
hour one
'one o'clock'

The same difference between prenominal and post-nominal modification can be seen in (53) and (54), which is not atypical in its use of pre- and post-nominal numerals to indicate general and specific cases. Note that a construction with a classifier as well as a numeral is also possible with a prenominal numeral, as seen in (55), but is not possible with a postnominal numeral, as shown in (56) (*se-* is the form taken by *satu* when cliticised to another root, numeral or classifier).

(53) satu potlot
one pencil
'one/a pencil'

(54) potlot satu (itu)
pencil one the/that
'the pencil that has been designated as "one"'

(55) se-batang potlot
one-CLF pencil
'one/a pencil'

(56) * potlot se-batang (itu)
pencil one-CLF the/that
'the pencil that has been designated as "one"'

Right-adjoined to the NP, *sebatang* is grammatical, in a position that a simple numeral cannot occur and with a set-member sense. Compare (56), in which *sebatang* precedes the NP-final demonstrative *itu*, with (58).

(57) * potlot itu satu terjatuh.
pencil the/that one fell
'The pencil that has been designated as "one" fell.'

(58) potlot itu se-batang terjatuh.
pencil the/that one-CLF fell
'One of the/those pencils fell.'

Extremely similar behaviour is found in English; compare the interpretation of (for instance) 'eleven buses', which refers to a quantity of buses (but not specifically any particular set of buses) and 'bus eleven', which uniquely identifies a bus, or at the most the set of buses that travel the number eleven route. In any case, it is certain that the use of the number post-nominally is associated with a more definite reading than the prenominal use.

3.2. Basque

Basque, a language isolate from western Europe, is another language in which we find that numbers may either precede or follow the head. Saltarelli (1988) notes that in a non-partitive numeral construction, the NP referring to the class of items being enumerated is the head of the construction and optionally marked for number. The numeral appears as a prenominal complement.'

Examples of unproblematic prenominal numerals in Basque are given in the examples seen in (59)-(61):

(59) bi mutil(-ak)
two boy-PL
'two boys'

(60) hiru neska(-k)
three girl-PL
'three girls'

(61) bost liburu(-ak)
five book-PL
'five books'

Saltarelli continues (1988: 168-169):

> *Bat* 'one', and in Bizkaiera *bi* 'two', are always treated as determiners and follow the head noun.

So far this marks only a difference in the number system; but we might importantly note the translations given for the numerals modifying the noun *ora* 'hour' in following sentences (sentence (63) also shows regular prenominal modification of *dei* 'call' by *bost* 'five') (Saltarelli 1988: 195).

(62) Azken bi ora-etan hemen ego-n d-a.
last two hour-PL.LOC here be-PRF 3A-PRS
'S/he has been here within the last two hours.'

(63) Ordu-bete-an bost dei iza-n d-it-u-t.
hour-one-S.LOC five call be-PRF 3A-A.PL-AUX2-1SG.E
'I have had five calls in a single hour.'

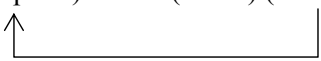
From (62)-(63) we can see that the use of the numeral 'one' post-nominally has a highly specific meaning, specifying not just one

hour, but a particular length of time used as a defining measure. Clearly the position of the postnominal numerals in the determiner position is more highly specific, and less approximate.

3.3. Numerals Allowed both Prenominally and Postnominally

Fore is a Trans New Guinea of the Goroka group, spoken in the eastern highlands of Papua New Guinea. The structure of the noun phrase is amenable to a templatic organization, shown in (64) (which excludes any NP-final case marking)

(64) NP:(Dem/Genitive)(Descriptive) HEAD(-POSS) (Numeral)



Examples illustrating these different prenominal orders, for all modifiers other than numerals, are given in (65)-(70) (all examples are drawn from Scott (1978), though some glosses have been inferred, rather than transferred—see Donohue & Donohue 1997). The morphology shall be discussed later.

(65) má: ntamá
má:´ -N na:máN
this -OBL/GEN house
'this house'

(66) na:mánka:né
na:máN -wá:´N -e
house -3SG.POSS -INDIC
'(it is) his house'

(67) naba:néné
na -pa:´ -né´N -e
1SG.INAL.POSS -father -1SG.POSS -INDIC
‘(it is) my father’

(68) kabá:re ntáninta:we
kabá:re´ -N naninta: -e
Kabare -OBL/GEN food -INDIC
‘Kabare’s food’

(69) aogi nama
aogi na:máN
good house
‘a good house’

(70) pai 'taeguntá: 'kiná
paiQ ta -egu´ -nt´´ -á:´N kináQ
long.ago 1PL.OBJ -hit -PERF -3PL.SUBJ.EMPH being
‘people who fought us long ago’

Scott (1978: 93) explicitly states that both Noun-Numeral and Numeral-Noun orders are possible (indeed, even Numeral-Noun-Numeral is found, with the (necessarily same) numeral occurring twice on both sides of the noun). Numerals are the only modifiers within the NP that may follow the noun. The post-nominal position is the marked position; Scott continues, noting that when the post-nominal position is used ‘a speaker gives it [the numeral-MD] more prominence than when it occurs preceding the head’. Examples of the appearance of a numeral on either or both sides of the nominal are given in (71)-(73), and some additional sentential examples (from Scott 1974: 94) in (74) and (75).

(71) tara yá
two tree
'two trees'

(72) ya: tára
tree two
'<u>two</u> trees'

(73) tara yá tára
two tree two
'<u>TWO</u> trees'

(74) Tara ya'ku-'tasa aeguyúwe.
two firewood-INSTR I:hit:him
'I hit him with two (pieces of) firewood.'

(75) Ya'ku tára-rasa aeguyúwe.
firewood two-INSTR I:hit:him
'I hit him with <u>two</u> (pieces of) firewood.'

This example is crucial, since it establishes the prominence and 'approximateness' associated with a numeral in pre- and post-nominal position, when the one language allows for both positions. Given a language with a choice of positions, and the explicit statement by Scott that the appearance of a numeral in the different positions is associated with different pragmatic contexts, we can test Greenberg's hypothesis in a control and empirical manner. In fact, when we try this, we find the facts do not support the proposed universal, but rather the opposite view.

4. The Grammaticalisation of Numerals to Articles: Asymmetries

Numerous adjectives display an indefinite article that is related in form to the numeral 'one', or uses of the numeral 'one' in an indefinite article-like manner. Given this common pattern of grammaticalisation, we might expect the order of the noun and an indefinite article that has its origins in a numeral to reflect the order that most clearly expresses indefiniteness; from Greenberg's Universal 44, we would expect a preponderance of Noun-Article orders.

In fact this is the opposite of the attested case; there are numerous examples of languages with an indefinite article preceding the noun, with that article derived from a numeral. In contrast, there are NO attested examples of a post-nominal indefinite article that is derived from a numeral. An example of the former is the case in Dutch, in which the indefinite article is clearly derived from the numeral *een* 'one' (the fact that English shows a similar grammaticalisation reflects both the close genetic links between English and Dutch, and a universal trend regarding the grammaticalisation of 'one' as an indefinite article). Compare the numerals in (76)-(77) with the indefinite article in (78). Note that, while the numeral 'one' and the indefinite article are orthographically identical, and are indeed pronounced identically in slow, careful speech, in most cases *een* the article is pronounced with a short schwa and an assimilating nasal coda (or even just a nasalised vowel), while the numeral *een* has a long front vowel and a fixedly alveolar nasal coda.

(76) *een* *boek* [eːn bʊk]
one book
'one book'

(77) twee boek-en
two book-PL
'two books'

(78) een boek [əm bʊk]~[ə̃ (m)bʊk]
INDEF.SG book
'a book'

If we were to assume that a post-nominal numeral contained a more indefinite reading, as suggested by Greenberg, then we would expect there to be examples of languages with post-nominal numerals grammaticalising into indefinite articles, yet this is in fact not the case. The absence of this type of language points to a more indefinite reading being found with pre-nominal numerals.

5. Conclusions

From the data seen here we can say that Greenberg's Universal 44 is clearly not an absolute universal, and is at best a statistical tendency. It can be treated as a statement on distribution, but has no validity as an absolute universal constraining the pragmatic interpretation of numerals in different positions; the Fore data alone are enough to ensure that we cannot interpret the universal as an absolute condition, but the other languages cited give strong support for the idea that there is a strong tendency for postnominal numerals to be interpreted in highly specific, highly definite ways. This is not a wholly negative criticism (see Dryer 1997 for reasons why

statistical universals are no less significant, and indeed more predictive, than absolute universals), but does force a reformulation of our view of this 'universal' on numeral systems. I offer a partial solution to why numerals show such different behaviour to that predicted by Greenberg: they do not occupy a unified semantic position, being propositions that show some behavioural properties associated with objects, some associated with properties, and some that are typical of actions.

Clearly a more detailed 'semantic map' of modifier types needs to be available before we can discuss the position of numerals in a typology of universal constraints on modification. We have seen many works in the last few decades discussing semantic maps, and semantic continua within verbs (Blume 1998; Testelec 1998; Tsunoda 1981, 1985, 1999; Malchukov 2005), there have been very few attempts to categorise and map different modifier types within the NP. I hope that this offering goes some way towards illustrating some of the possible methodologies, and pitfalls, that are to be found in an investigation into the typology of modification.

References

Baker, M. 2003. *Lexical Categories: Verbs, Nouns, and Adjectives*. Cambridge: Cambridge University Press.

Bauer, W., W. Parker, & T. Evans. 1993. *Maori*. London: Routledge.

Blume, K. 1998. A Contrastive Analysis of Interaction Verbs with Dative Complements. *Linguistics* 36, 253-280.

Börjars, K. & M. Donohue. 2000. Much Ado about Nothing: Features and Zeroes in Germanic Noun Phrases. *Studia Linguistica* 54, 309-353.

Burrow, T. & S. Bhattacharya. 1970. *The Pengo Language*. Oxford: Clarendon Press.

Carlson, J. 1994. *A Grammar of Supyire*. (Mouton Grammar Library 15). Berlin: Mouton de Gruyter.

Conrad, J. & K. Wogiga. 1991. *An Outline of Bukiyip Grammar*. Canberra: Pacific Linguistics C-113.

Croft, W. 1991. *Syntactic Categories and Grammatical Relations: The Cognitive Organization of Information*. Chicago, IL: The University of Chicago Press.

Donohue, M. 1999a. *A Grammar of Tukang Besi*. (Mouton Grammar Library no. 20). Berlin: Mouton de Gruyter.

_____. 1999b. Syntactic Categories in Tukang Besi. *Revue Québécoise de Linguistique* 27, 71-90.

Donohue, C. & M. Donohue. 1997. Fore Case Marking. *Language and Linguistics in Melanesia* 28, 69-98.

Dryer, M. 1988. Object-verb Order and Adjective-noun Order: Dispelling a Myth. *Lingua* 74, 185-217.

_____. 1991. SVO Languages and the OV: VO Typology. *Language* 27, 443-482.

_____. 1997. Why Statistical Universals are Better than Absolute Universals. *Chicago Linguistic Society* 33, 123-145.

Fortune, F. 1942 (1977). *Arapesh*. Publications of the American Ethnological Society Volume XIX. Reprinted New York: AMS.

Greenberg, H. 1978. Generalizations about Numeral Systems. In J. Greenberg & E. Moravcsik (eds.), *Universals of Human Language 3: Word Structure* 249-295. Stanford, CA: Stanford University Press.

Jones, B. 2002. A Grammar of Wangkajunga, a Language of the Great Sandy Desert of North Western Australia. Ph.D. Dissertataion, Sydney: University of Sydney.

Kimball, D. 1991. *Koasati Gramma*r. Omaha, NE: University of Nebraska Press.

Malchukov, A. 2005. Case Pattern Splits, Verb Types, and Construction Competition. In M. Amberber & H. de Hoop (eds.), *Competition and Variation in Natural Languages: The Case for Case* 73-117. London: Elsevier.

McGregor, E. & A. McGregor. 1982. *Olo Language Materials*. Canberra: Pacific Linguistics D-42.

Mosel, U. & E. Hovdhaugen. 1992. *Samoan Reference Gramma*r. Oslo: The Institute for Comparative Research in Human Culture,

Saltarelli, M. 1988. Basque. *Croom Helm Descriptive Grammars*. London: Croom Helm.

Scott, G. 1978. *The Fore Language of Papua New Guinea*. Canberra: Pacific Linguistics B-47.

Sneddon, N. 1996. *Indonesian: A Comprehensive Grammar*. London: Routledge.

Testelec, G. 1998. On Two Parameters of Transitivity. In L. Kulikov & H. Vater (eds.), *Typology of Verbal Categorie*s. 29-45, Tübingen: Niemeyer

Tsunoda, T. 1981. Split Case-marking Patterns in Verb-types and Tense/Aspect/Mood. *Linguistics* 19, 5. 389-438.

_____. 1985. Remarks on Transitivity. *Journal of Linguistics* 21, 2. 385-396.

_____. 1999. Transitivity. In K. Brown & J. Miller (eds.), *Concise Encyclopedia of Grammatical Categories* 383-391. Amsterdam: Elsevier.

Visser, E. & C. Voorhoeve. 1987. *Sahu-Indonesian-English Dictionary and Sahu Grammar Sketch* (Verhandelingen van het Koninklijk Instituut voor Taal-, Land- en Volkenkunde deel 126). Dordrecht: Foris.

Voorhoeve, L. 1982. The West Makian Language, North Moluccas, Indonesia: A Fieldwork Report. In C. Voorhoeve (ed.), *The Makian Languages and their Neighbours* 1-74. Canberra: Pacific Linguistics D-46.