TYPOLOGY OF NUMERAL SYSTEMS

Bernard Comrie
University of California Santa Barbara
<comrie@linguistics.ucsb.edu>

Version of 2021 May 26

Where no source is given, I have used my own knowledge of the language in question, supplemented where necessary by checking (e.g. spelling) in a standard dictionary.

A. Restricted systems, with little or no internal structure

(1) No numerals

[*Pirahã*; Everett 2005: 623-627]

(2)

(2)	1	(ŋa)wumbawa
	2	ŋabaranwa
	3	ŋabaḷawa

[*Mangarrayi*; Merlan 1982: 92]

(3)	1	guman	
	2	jambul	
	3	dagul	
	4	yunggan.gunyjii *or* mugungabi	
	5	mala	'face of hand'

[*Yidiny*; Dixon 1991: 224]

Note: The forms for 4 and above may not be fully conventionalized.

(4)	1	püg (ma)	
	2	xep xep	
	3	ebapüg	2 + 1
	4	ebadipdip	2 + 1 + 1
	5	püg pögbi	'one hand'

[*Mundurukú*; Pica et al. 2004; see also section B]

(5)	1	towenyxa	
	2	asako	
	3	osorwawo	
	4	towtɨnke rye	'its brother twice over'
	5	kamorɨ ɨrakay (o) me	'half of our hands'

10 kamothɨrɨ tkatxehkaxe ro ‘our hands completely’

[*Hixkaryána*; Derbyshire 1979: 154-155]

Note: The forms for 4 and above are “non-basic” and “are also used without precision”.

(6) **subitize**
Subitizing is the rapid, accurate, and confident judgment of numbers
performed for small numbers of items
with a cut-off point around 4

B. Simple systems with addition only

(7)

1	paŋ	
2	mös	
3	mös paŋ	2 + 1
4	mös mös	2 + 2

[*Haruai*; Comrie: Own fieldwork]

(8)

1	añɨ (nöbö)
2	möhöp
3	möhau nögaŋ
4	möhau möhau

[*Kobon*; Davies 1981: 208]

(9)

1	bits	
2	iruʔ	
3	iruʔ da bits	2 + 1
4	iruʔ da iruʔ	2 + 2
5	iruʔ da iruʔ da bits	2 + 2 + 1
6	iruʔ da iruʔ da iruʔ	2 + 2 + 2
etc.		
10	dzi bangi marafain da dzi bangi marafain sib ‘my hand half and my hand half completed’	

[*Adzera*; Lean & Owens 2018: 299; partially checked against Smith (1988: 66), Holzknecht (1989: 127–128).]

C. More complex systems using multiplication and addition applied to a base

(10) General pattern:
For base b: (n x b) + m
where m < b
and usually n < b, or n ≤ b

Decimal (base 10)

(11)

wǔ-shí	sì
five-ten	four

54 [50 + 4]

[*Mandarin Chinese*]

Vigesimal (base 20)

(12)

kəlgən-qlekken	məngətken	ŋireq	parol
fifteen-twenty	ten	two	left

312 [(15 x 20) + (10 + 2)]

[*Chukchi*; Skorik 1961: 390]

Base 60

(13)

èna	ma	gàati	dàimita	mutò
one	and	ten	and	sixty

71 [60 + (10 + 1)]

(14)

muto	wìi
sixty	four

240 [4 x 60]

[*Ekari*; Drabbe 1952: 30]

Base 32

(15)

ɨ̀fɔ	wǎdhɨ̀
four	thirtytwo

128 [4 x 32]

[*Ngiti*; Kutsch Lojenga 1994: 355-358]

Base 12

(16)

ba-kuru	ba-ba	ná	CL-ā	CL-bā
PL-twelve	PL-two	plus	CL-this	CL-two

26 [(2 x 12) + 2]

[*Berom*; Bouquiaux 1970: 259]

Note: The last two elements -*ā* and -*bā* require a class prefix corresponding to the class of the head noun.

Base 8

(17) karnuʔ tenhiuɲ rnuʔ
three eight three
27 [(3 x 8) + 3]

[*Northern Pame*; Avelino 2006: 45]

Base 6

(18)

1	näbi	
2	eda/yda	
3	etha/ytho	
4	asar	
5	tabuthui	
6	nibo	
36	fta	6^2
216	taruba	6^3
1296	damno	6^4
7776	wärämäkä	6^5
46656	wi	6^6
72	eda fta	2×6^2
73	eda fta a näbi	$(2 \times 6^2) + 1$
50	näbi fta a eda nibo a eda	$(1 \times 6^2) + (2 \times 6) + 2$

[*Komnzo*; Döhling 218: 93-94; Yam-counting video: https://vimeo.com/54887315]

New Guinea Highland body-part counting systems ("bases" 18–74)

(19)

wañɨg nöbö	little finger	1		23	24		46
igwo	ring finger	2		22	25		45
igwo aŋ nöbö	middle finger	3		21	26		44
igwo mɨlö	index finger	4		20	27		43
mamɨd	thumb	5		19	28		42
kagoł	wrist	6		18	29		41
mudun	forearm	7		17	30		40
raleb	inside of elbow	8		16	31		39
ajɨp	biceps	9		15	32		38
siduŋ	shoulder	10		14	33		37
agɨp	collarbone	11		13	34		36
mögan	hole above breastbone		12			35	

[*Kobon*; Davies 1981: 206-208]

Notes:
a) The second half of the body on each pass across the body is distinguished by *böŋ* 'one side of'.
b) The second pass across the body is distinguished by *ñin juöl adog da* 'pulling out the hand, give back!', subsequent passes by modifying *ñin juöl* with a numeral.
c) The system is "symmetrical", i.e. the body parts used on one side of the body (in Kobon, for 1–11) are used in reverse order for the other side of the body (Kobon 13–23).

Arithmetic properties (without any claims to psychological reality):

a) On the first pass across the body, any given number term *n* also denotes 24 – *n*, and similarly, mutatis mutandis, for subsequent passes across the body.
b) The system has a base 23.

(20)

agɨlɨŋ	little finger	1			19		
agɨlɨŋ rolyöbö	ring finger	2			20		
wölö mɨl	middle finger	3			21		
köñö nɨgɨb	index finger	4			22		
mömɨd	thumb	5			23		
urap cɨgɨb	wrist	6		18	24		36
mɨj	forearm	7		17	25		35
amɨñaxɨb	inside of elbow	8		16	26		34
mac	biceps	9		15	27		33
möib	shoulder	10		14	28		32
katlöi	collarbone	11		13	29		31
mɨgan	hole above breastbone		12			30	

[*Haruai*; Comrie: Own fieldwork]

[Some numerals, like the corresponding body parts, have alternative forms; *urap cɨgɨb* and *mac* are properly bracelets worn at the given body part.]

Notes:
a) The second half of the body on each pass across the body is distinguished by *adökwebö* 'of that side'.
b) The second pass across the body is distinguished by *höuöilɨb* 'turning' or *höbkalɨb* 'returning'.
c) The end of the first pass across the body (i.e. 23) can be identified by saying *padö kwo paŋyöbö dua* 'housepost there one went', subsequent passes by replacing *paŋyöbö* 'one' with a higher numeral and using the plural verb form *dumä*.
d) The system is "asymmetrical", i.e. the body parts used on one side of the body (in Haruai, for 1–11) are not all used in reverse order for the other side of the body (cf. in particular the fingers).

(20')

agɨlɨŋ	little finger	1		19			37
agɨlɨŋ rolyöbö	ring finger	2		20			38
wölö mɨl	middle finger	3		21			39
köñö nɨgɨb	index finger	4		22			40
mömɨd	thumb	5		23			41
urap cɨgɨb	wrist	6		18	24		36
mɨj	forearm	7		17	25		35
amɨñaxɨb	inside of elbow	8		16	26		34
mac	biceps	9		15	27		33
möib	shoulder	10		14	28		32
katlöi	collarbone	11		13	29		31
mɨgan	hole above breastbone		12			30	

Arithmetic properties:
a) There is no consistent relation between the numerical values of the same body part on different sides of the body. For 6–18, the two add up to 24; for 1–5 and 19–23, the total varies from 20 to 28.
b) The base shifts from 23 on the first pass across the body to 18 on subsequent passes, i.e. passes end at 23, 41, 59, etc.

See also (on Oksapmin): Saxe, Geoffrey B., with Indigo Esmonde, *Cultural development of mathematical ideas: Papua New Guinea studies*, New York, 2012: Cambridge University Press.

(21) Hypothesis: Arithmetic bases of numeral systems have either a somatic or a commercial (transactional) origin; lower bases are typically somatic, higher bases commercial, but New Guinea Highland body-part counting systems have relatively high somatic-origin bases.

10	fingers
20	fingers and toes; each finger twice (two phalanges/knuckles)
8	spaces between fingers (attested for some California languages)
12	phalanges or knuckles of fingers (excluding thumbs)

For higher bases with a commercial origin, cf. English *score* ‘20’, which in some varieties has made it into the numeral system.

(22) If the lowest multiplicative base is higher than about 12, and the system is not an extended body-part counting system, then the higher numbers below the lowest multiple base make use of an additive base, e.g. Georgian (lowest multiplicative base: 20); some languages use this even for a smaller lowest multiplicative base, e.g. Nauete (lowest multiplicative base: 10); see also Chukchi example (12) above.

(23)

1	ert-i	-i nominative case (NOM)
2	or-i	

5	xut-i	
6	ekvs-i	
9	cxra	NOM drops after a
10	at-i	
11	t-ert-met'-i	Ø-one-teen-NOM
15	t-xut-met'-i	
16	t-ekvs-met'-i	
19	[t-]cxra-met'-i	(t + c > c)
20	oc-i	
21	oc-da-ert-i	twenty-and-one-NOM
30	oc-da-at-i	twenty-and-ten-NOM
31	oc-da-t-ert-met'-i	twenty-and-Ø-three-teen-NOM
40	or-m-oc-i	two-Ø-twenty-NOM

[*Georgian*; Hewitt 1995: 51-52]

(24)

2	kai-rua	CL-two
3	kai-tolu	CL-three
5	kai-lima	CL-five
7	kai-lima resi kai-rua	CL-five and CL-two
8	kai-lima resi kai-tolu	CL-five and CL-three
10	weli-see	ten-one
12	weli-see resi kai-rua	ten-one and CL-two
17	weli-see resi kai-lima resi kai-rua	ten-one and CL-five and CL-two
20	weli-kai-rua	ten-CL-two

[*Nauete*; https://mpi-lingweb.shh.mpg.de/numeral/Naute.htm, with data provided by Geoffrey Hull in 1996; see also Schapper & Hammarström 2013: 429, and https://abvd.shh.mpg.de/austronesian/language.php?id=1365, the latter with data provided by Alexandre Veloso in 2016. There are differences of detail among the sources, not affecting the main point.]

Note: Nauete numerals 2-5 require a numeral classifier prefix, with *kai-* being the default used in counting.

(25) Distribution of different bases across a sample of languages of the world

[See: Comrie (2013)]

D. Idiosyncrasies relating to bases

Portmanteau forms

(26) sorok
forty

40 [expected 4 x 10]

[*Russian*]

(27) eleven
11 [expected 10 + 1]

[*English*]

(28)

(28)	25	se-lae	‘one thread (of Chinese coins)’
	45	se-timan	‘one opium packet (costing 45 Chinese coins)’
	50	se-ket	‘one tie (i.e. two threads of 25 Chinese coins)’
	75	telung benang	‘three threads (of Chinese coins)’
	200	s-atak	‘one bundle of 200 Chinese coins’
	400	s-aman	‘one gold (coin worth 400 Chinese coins)’
	900	sanga	[etymology unclear]

[*Balinese*; Eiseman 1990: 162-168]

(29) Compare less spectacular irregularities

fif-teen (*five-teen)
five-ten
15

twenty; twelve

[*English*]

(30) In Hindi, arguably all the numerals 1–100 are irregular

	0	1	2	3	4	5	6	7	8	9
-		ek	do	tīn	cār	pāṁc	chah	sāt	āṭh	nau
10	das	gyārah	bārah	terah	caudah	pandrah	solah	satrah	aṭhārah	unnīs
20	bīs	ikkīs	bāīs	teīs	caubīs	paccīs	chabbīs	sattāīs	aṭṭāīs	untīs
30	tīs	ikattīs	battīs	taiṁtīs	cauṁtīs	paiṁtīs	chattīs	saiṁtīs	aṛtīs	untālīs
40	cālīs	iktālīs	bayālīs	taiṁtālīs	cavālīs	paiṁtālīs	chiyālīs	saiṁtālīs	aṛtālīs	uncās
50	pacās	ikyāvan	bāvan	tirpan	cauvan	pacpan	chappan	sattāvan	aṭṭhāvan	unsaṭh
60	sāṭh	iksaṭh	bāsaṭh	tirsaṭh	cauṁsaṭh	paiṁsaṭh	chiyāsaṭh	sarsaṭh	aṛsaṭh	unhattar
70	sattar	ikhattar	bahattar	tihattar	cauhattar	pachattar	chihattar	sathattar	aṭhhattar	unyāsī
80	assī	ikyāsī	bayāsī	tirāsī	caurāsī	pacāsī	chiyāsī	sattāsi	aṭṭhāsī	navāsī
90	nave	ikyānve	bānve	tirānve	caurānve	pacānve	chiyānve	sattānve	aṭṭhānve	ninyānve

[*Hindi*; McGregor 1972: 61-62; note that some numerals have minor variants]

Isolated “bases”

(31) quatre-vingt-douze

four-twenty-twelve
92 [(4 x 20) + 12]

[*French*]

(32) deu-naw
two-nine
18 [2 x 9]

[*Welsh*; King 1993: 113]

Overrunning

(33) disąt-nocti
10-teen
20 [10 + 10]

Note: 21 is janü disąt-nocti, i.e. [1 + 20], not [11-teen]

[*Polabian*; Polański & Sehnert 1972: 52, 73]

(34) soixante-dix
sixty-ten
70 [60 + 10]

(35) soixante-douze
sixty-twelve
72 [60 + 12]

(36) soixante-dix-sept
sixty-ten-seven
77 [60 + (10 + 7), rather than [(60 + 10) + 7]

[*French*]

E. Exponentiation and other higher bases

(37)	10^1	10^2	10^3	10^6
	ten	hundred	thousand	million

[*English*]

Absence of exponentiation

(38) qliq-qlikkin
twenty-twenty

400 (20 x 20) – highest numeral in traditional system

[*Chukchi*; Skorik 1961: 388, 391]

Effectively monomorphemic series of bases

(39)	10	dáśa
	100	śatá
	1,000	sahásra
	10,000	ayúta
	100,000	lakṣá
	1,000,000	prayúta
	10^7	kóṭi
	10^8	arbudá
	10^9	mahārbuda
	10^{10}	kharvá
	10^{11}	nikharva

[*Sanskrit*; Whitney 1889: 177-178, who notes further "The series of decimal numbers may be carried still further; but there are great differences among the different authorities with regard to their names; and there is more or less of discordance even from *ayúta* on."]

Note: Indian English currently uses: lakh 10^5; crore 10^7; in India, 133,435,360 would be written as 13,34,35,360 (13 crore, 34 lakh, 35 thousand, 360).

(40)	wàn	万		萬	10^4	
	yì	亿		億	10^8	
	zhào		兆		10^{12}	
	jīng		京		10^{16}	etc.

[*Mandarin Chinese*. This system is used in Chinese, Japanese, Korean; but see also (92)]

(41)	cem-pōhual-li	one-twenty-ABS	20
	cen-tzon-tli	one-four.hundred-ABS	400 (20^2)
	cen-xiquipil-li	one-eight.thousand-ABS	8000 (20^3)

[*Classical Nahuatl*; Andrews 1975: 397-398, 464, 482, 484]

Note: *pōhua-* is also a verb 'to count'; *tzon-* also means 'hair'; *xiquipil-* also means 'bag, sack'.

(Semi-)productive systems

(42)

		long scale	short scale
million	first	10^6	10^6
billion	second	10^{12}	10^9
trillion	third	10^{18}	10^{12}
quadrillion	fourth	10^{24}	10^{15}
[general pattern]	nth	10^{6n}	$10^{3(n+1)}$

[*English*]

See http://www.isthe.com/chongo/tech/math/number/howhigh.html for a proposal on how to count indefinitely high using (pseudo-)Latin prefixes.

(43) Contrast innovative / metric

kilo-	10^3
mega-	10^6
giga-	10^9
tera-	10^{12}
peta-	10^{15}
exa-	10^{18}
zetta-	10^{21}
yotta-	10^{24}

[*English / International*]

Sequence of bases that are not (all) powers of a single base

(44)

cxr-as	otxm-oc-da-cxra-met'i
nine-hundred	four-twenty-and-nine-teen

999 [(9 x 100) + (4 x 20) + (10 + 9)]
Bases: (10,) 20, 100

[*Georgian*; Hewitt 1995: 54]

(45)

kàmpwòò	ná	ŋ̀kwuu	sicyɛɛré	'ná	béé-tàànrè	ná	kɛ́	'ná	báár-ìcyɛ̀ɛ̀rè
fourhundred	and	eighty	four	and	twenty-three	and	ten	and	five-four

799 [i.e. 400 + (4 x 80) + (3 x 20) + {10 + (5 + 4)}]'
Bases: (5, 10,) 20, 80, 400

[*Supyire Senoufo*; Carlson 1994: 169]

(46) Note that the next higher base (nhb) is nearly always equal to or less than the square of the current base (cb):

nhb ≤ cb^2

But note the older Germanic “long hundred”:

(47)

10	tíu	
30	þrír tigir	
100	tíu tigir	
110	ellifu tigir	
120	hundrað	(“long hundred”)
240	tvau hundrað	
1200	þúsund	(“long thousand”)

[*Old Norse*; Gordon 1957: 292-293]

(48) A higher multiplicative base is nearly always a product of the lowest multiplicative base. An exception is provided by Coahuiltecan, with multiplicative bases 3 (for numerals 12-19) and 20; surprisingly, the lower base 3 is itself complex (2 + 1).

(49)

12	4 x 3
13	4 x 3 +1
14	4 x 3 + 2
15	5 x 3
18	6 x 3 (?)
20	20
30	20 + 10
40	2 x 20

(50)

puwāntz’an	axti-k-pil’	ko	pil’
four	two-and-one	and	one

13 [i.e. {4 x (2 + 1)} + 1]

[*Coahuiltecan*; Swanton 1940: 48]

Alternating bases

(51)

10	dḙ́sat	10
20	dwísti	2 x 10
30	trә́sti	3 x 10
40	dwákrat dwísti	2 x 20
50	patardú	5 x 10
60	trḭ́krat dwísti	3 x 20

[*Resia Slovenian*; Steenwijk 1992: 125]

F. Other arithmetic processes

Subtraction

(52) un-de-viginti
one-from-twenty
19 [20 – 1]

[*Latin*]

Division (actually: multiplication by fraction)

(53)

hanner	cant
half	hundred

50 [½ x 100]

[*Welsh*; King 1993: 113]

Subtraction and addition

(54)

éks	bə́nsaŋ	kiʔ
twenty	without	hundred

80 [100 – 20]

(55)

ɨ́nam	ə́kam	éks	bə́nsaŋ	kiʔ
two	left.over	twenty	without	hundred

82 [(100 – 20) + 2; NB: *not* 100 – (20 + 2)]

[*Ket*; Georg 2007: 179-181]

(56) Successive approximation, cf. time expressions in some languages

drie	(uur)
three	hour

03:00

(57)

half	drie
half	three

02:30

(58)

vijf	over	half	drie
five	after	half	three

02:35

[*Dutch*]

Overcounting

(59) halv-tred-sinds-tyve

half-third-times-twenty
50 [half of the third, times twenty]

[*Danish*; Allan et al. 1995: 127; now usually: halvtreds]

(60)

paüne	tini	šata
less.quarter	three	hundred

275 [less a quarter of the third hundred]

[*Oriya*; Karpuškin 1964: 38]

Pairing

(61)

1	seénu, wepul		
2	woói		
3	báhi		
4	naíki		
5	mámni		
6	búsani		
7	wóo-búsani	two-six (i.e. ‘second six’?)	
8	wóh-naíki	two-four (i.e. 2 x 4)	←
9	bátani		
10	wóh-mámni	two-five (i.e. 2 x 5)	←

[*Yaqui*; Dedrick & Casad 1999: 229]

(62)

1	hito	2	huta
3	mi	6	mu
4	yo	8	ya

[*Japanese*]

Non-arithmetic structures

(63)

10^8	arbudá-	
10^9	mahārbuda-	(*maha*- ‘big’)

[*Sanskrit*; Whitney 1889: 177]

(64)

10^3	mille (PL mila)	
10^6	milione	(*-one* AUGMENTATIVE)

[*Italian*]

G. Ordering of constituents

From larger to smaller

(65) sān-bǎi wǔ-shí sì
three-hundred five-ten four
354 [i.e. 300 + 50 + 4]

[*Mandarin Chinese*]

From smaller to larger

(66) efatra amby dima-mpolo sy telo-njato
four plus five-ten and three-hundred
354 [i.e. 4 + 50 + 300]

[*Malagasy (Standard)*; Rajaonarimanana 2001: 67]

From smaller to larger for smaller combinations, from larger to smaller for larger combinations

(67) drei-hundert-vier-und-fünf-zig
three-hundred-four-and-five-ten
354 [i.e. 300 + 4 + 50]

(68) zwei-hundert-sechs-und-fünf-zig-tausend-drei-hundert-vier-und-sieb-zig
two-hundred-six-and-five-ten-thousand-three-hundred-four-and-seven-ty
256 374 [i.e. (200 + 6 + 50) x 1000 + (300 + 4 + 70)]

[*German*]

(69) order smaller-larger (contrasting with higher larger-smaller)

through 12 (ðó-ðeka, cf. 13 ðeka-trís)	*Modern Greek*
through 15 (quin-ce, cf. 16 diec-i-séis)	*Spanish*
through 16 (se-dici, cf. 17 dici-as-sette)	*Italian*
through 19 (nine-teen)	*English*
through 99 (neun-und-neun-zig)	*German*

[For Modern Greek: Holton et al. 1997: 103-104]

From larger to smaller for smaller combinations, from smaller to larger for larger combinations

(70) limam-polo roe amby, amby telon-jato
five-ten two plus plus three-hundred
'352 [i.e. 50 + 2 + 300]'

[*Malagasy (Nosy Be)*; Dahl 1968: 14]

(71) Hypothesis: The order from larger to smaller is preferred because it gives earlier recognition of the approximate quantity involved, i.e. in 354 the 300 is more significant than the 4. Local inversion of lower positions (e.g. tens and units) is minimally disruptive. So we expect prevalence of the order from larger to smaller, with possible local inversion of the lower positions.

H. Finiteness and gaps (Comrie 2020)

(72) "Every language has a numeral system of finite scope."
(Generalization 1, Greenberg 1978)

(73)

	Short scale	Long scale	
		without intermediaries	with intermediaries
10^6	million	million	million
10^9	billion		milliard
10^{12}	trillion	billion	billion
10^{15}	quadrillion		billiard
10^{18}	quintillion	trillion	trillion

(74)
a. a new single-word term is available each time one multiplies the previously highest single-word term by 1,000,000, up to *trillion*, i.e. there are no terms **quadrillion* and beyond;
b. there are no intermediaries, e.g. **milliard*, so 10^9 can only be expressed as *thousand million*;
c. numerals from *million* up can be combined to express multiplication, e.g. *million trillion* ($[10^6 \times 10^{18}] = 10^{24}$);
d. there is a strong preference to implement the "Packing Strategy" (Hurford 1975: 67), i.e., informally speaking, to avoid combining terms with lower numerical value when there is an available appropriate term with higher numerical value; thus, although *million million* is possible, *billion* is preferred;
e. it is not possible for a higher-value term to precede a lower-value term, e.g. for 10^{30} one can say *billion trillion* but not **trillion billion*.

(75)

a.	10^6	million
b.	10^9	thousand million
c.	10^{12}	billion
d.	10^{15}	thousand billion
e.	10^{18}	trillion
f.	10^{24}	million trillion
g.	10^{30}	billion trillion
h.	10^{36}	trillion trillion
i.	10^{54}	trillion trillion trillion
j.	8×10^{67}	eighty billion trillion trillion trillion

Traditional British long-scale system (my version)

(76) Every number *n* (0 < n < L) can be expressed as part of the numerical system in any language.

[Generalization 2, Greenberg 1978]

(77) decillion, vigintillion, centillion
What is the 23rd term in the series?
tresvigintillion [https://en.wikipedia.org/wiki/Names_of_large_numbers]

I. Ambiguity

Parsing ambiguities

(78) a million and a half (apples)
(i) 1½ million, i.e. 1,500,000
(ii) 1,000,000 ½

[*English*]

(79)

un	fil	a	thri	ugain	o	asynod
one	thousand	and	three	twenty	of	asses

(i) 1060 asses (possible interpretation)
(ii) 61,000 asses (Numbers 31.34 – intended interpretation)

(80)

deuddeg	a	thri	ugain	mil	o	eidionau
twelve	and	three	twenty	thousand	of	cattle

72,000 cattle (Numbers 31.33)

(81)

saith	mil	a	phedwar	ugain	mil
seven	thousand	and	four	twenty	thousand

87,000 (I Chronicles 7.5)

[*Welsh (Biblical)*; Hurford 1975: 192, 184, 176]

Abbreviation

(82)

(nɨ̀ŋ)	phan	sɔ̌ɔŋ
one	thousand	two

(i) 1002
(ii) 1200
(iii) *1020

(83)

(nɨ̀ŋ)	phan	sɔ̌ɔŋ	rɔ́ɔy
one	thousand	two	hundred

1200

(84)

(84)	(nɨ̀ŋ)	phan	kàp	sɔ̌ɔŋ
	one	thousand	with	two
	1002			

[*Thai*; Smyth 2002: 173]

(85)	(a)	sān-bǎi	liù-shí
		three-hundred	six-ten
	(b)	sān-bǎi	liù
		three-hundred	six
	360		

(86)	sān-bǎi	líng	liù
	three-hundred	zero	six
	306		

Note: Obligatory use of *líng*, creating a partial place-value system, avoids ambiguity.

(87)	sān-qiān	líng	liù
	three-thousand	zero	six
	3006		

(88)	sān-qiān	líng	liù-shí
	three-thousand	zero	six-ten
	3060		

[*Mandarin Chinese*]

Note: An older usage required *líng* to be repeated as many times as there are zeroes in the "Arabic" numeral representation, so (87) would be *sān-qiān líng líng liù* (Chao 1968: 575).

Diachronic merger

(89) thirty, thirteen

[*English*]

(90)	7	sāt
	100	sāt > janiɣ sāt (janiɣ 'big')

[*Northern Mansi*; Honti 1993: 125]

Semantic change

(91) billion

(i) older, “long scale”	10^{12}
(ii) newer, “short scale” = US English	10^{9}

(long scale: 10^{6n}; short scale: $10^{3(n+1)}$)

[*British English*]

(92)

	万/萬	亿/億	兆	京
	wàn	yì	zhào	jīng
(i)	10^{4}	10^{5}	10^{6}	10^{7}
(ii)	10^{4}	10^{8}	10^{12}	10^{16}
(iii)	10^{4}	10^{8}	10^{16}	10^{24}
(iv)	10^{4}	10^{8}	10^{16}	10^{32}

[*East Asian*]

[Row (ii) is the one in current use for Chinese, Japanese, and Korean, though with occasional variation. Rows (i), (iii), (iv) are presented by Martzloff (1997: 99), based on the Chinese mathematical classic *Shushu jiyi*.]

Specialized use

(93) bak
usually 400
but 360 days (long calendar)

[*Mayan*]

[Tozzer 1921: 97, and more generally 97-103. The forms cited by Tozzer are more specifically Yucatec Maya; note that Tozzer uses the spelling *baq*. The Classical Mayan inscriptions and codices do not clearly include large numbers other than in calendrical accounts (Chrisomalis 2010: 292-294).]

(94) kilo-
1000
but: kilobyte ‘1024 bytes’

[*English/International*]

Body part systems

(95) siduŋ ‘shoulder’ = 10, 14, 33, 37, 56, 60

[*Kobon*; Davies 1981: 206-208]

J. Internal structure and psychological reality

(96) Issues at interface with psychomathematics (psychoarithmetic)

a) To what extent are problems caused by having a linguistic representation that does not correspond to the arithmetic notation, e.g., assuming a decimal notation in descending order
(i) if a formation is non-decimal, e.g. French quatre-vingt-onze, literally 'four twenties and eleven', for 91?
(ii) if the order is (partially) inverted, e.g. German sechsundfünfzig, literally 'six and fifty', for 56?

b) When people do arithmetic, do they operate with linguistic representations of numerals (in which case problems (a) might arise) or do they operate with abstract quantities or arithmetic notations (in which case they should not)?

c) But note that some arithmetic calculations might be easier in a non-decimal base, e.g. 68 + 7 might be easier in a vigesimal base, where one does not have to "carry over" a number; in decimal notation, 68 + 7 = 75; in vigesimal notation 38 + 7 = 3F (where F = 15)

(97) Writing 'six and fifty' for 56

Dutch	zesenvijftig
German	sechsundfünfzig

Dutch speakers typically write the 6 and then go back to "fill in" the 5.
German speakers typically write the 5 then the 6, and are explicitly instructed to do this at school.

(98) But note that in some instances the overall structure of the numeral system means that the formal structure of individual numerals can be overridden.

(99)

	Cardinal	Ordinal	
10	ti	tiende	
20	tyve	tyvende	
30	tred(i)ve	tred(i)vte	
40	fyrre	fyrretyvende	
50	halvtreds	halvtredsindstyvende	cf. tredje 3rd
60	tres	tresindstyvende	cf. tre 3
70	halvfjerds	halvfjerdsindstyvende	cf. fjerde 4th
80	firs	firsindstyvende	cf. fire 4
90	halvfems	halvfemsindstyvende	cf. femte 5th
100	(et) hundred(e)	hundrede	

[*Danish*; Allan et al. 1995: 122]

(100) There can be other discrepancies between formal structure (often historical/etymological) and current analysis for native speakers, cf.

(101)

	10	at-i	
	20	oc-i	
	100	as-i	
	1000	at-as-i	ten-hundred-NOM
	2000	or-i at-as-i (not *oc-as-i)	two-nom-ten-hundred-NOM
			(*twenty-hundred-NOM)

[*Georgian*; Hewitt 1995: 51-54]

(102)

10^6	éna ekato-míri-o	one hundred-ten.thousand-SG
10^7	ðéka ekato-míri-a	ten hundred-ten.thousand-PL
		(not *thousand-ten.thousand)

[*Modern Greek*; LKN 1998: 427]

(103) Possible competing pedagogical advantages

a) Having a numeral system that corresponds to the arithmetic notation facilitates learning arithmetic.

b) Having a numeral system that does not correspond to the arithmetic notation provides extensive practice in arithmetic and leads to a higher level of arithmetic skill.

References

This list includes references from the body of the handout, plus some other more general works.

Allan, Robin, Philip Holmes, & Tom Lundskær-Nielsen. 1995. *Danish: A comprehensive grammar*. London: Routledge.
Andrews, J. Richard. 1975. *Introduction to Classical Nahuatl*. Austin: University of Texas Press.
Avelino, Heriberto. 2006. The typology of Pamean number systems and the limits of Mesoamerica as a linguistic area. *Linguistic Typology* 10: 41-60.
Barriga Puente, Francisco. 1998. *Los sistemas de numeración indoamericanos: un enfoque areotipológico*. México: Universidad Nacional Autónoma de México.
Bouquiaux, Luc. 1970. L*a langue birom (Nigeria septentrional): Phonologie, morphologie, syntaxe*. Paris: "Les Belles Lettres".
Carlson, Robert. 1994. *A grammar of Supyire*. Berlin: Mouton de Gruyter.
Chan, Eugene (compiler). *Numeral systems of the world's languages*. https://mpi-lingweb.shh.mpg.de/numeral/.

Chao, Yuen Ren. 1968. *A grammar of spoken Chinese*. Berkeley: University of California Press.
Chrisomalis, Stephen. 2010. *Numerical notation: A comparative history*. Cambridge: Cambridge University Press.
Closs, Michael P. (ed.). 1986. *Native American mathematics*. Austin: University of Texas Press.
Comrie, Bernard. 2005. Endangered numeral systems. In Jan Wohlgemuth & Tyko Dirksmeyer (eds.): *Bedrohte Vielfalt: Aspekte des Sprach(en)tods* [*Endangered diversity: Aspects of language death*], 203–230. Berlin: Weißensee Verlag.
Comrie, Bernard. 2013. Numeral bases. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wals.info/chapter/131.
Comrie, Bernard. 2020. 'Revisiting Greenberg's "Generalizations about Numeral Systems" (1978)'. *Journal of Universal Language* 21.2: 43-84.
Dahl, Otto Chr. 1968. *Contes malgaches en dialecte Sakalava: Textes, traduction, grammaire et lexique*. Oslo: Universitetsforlaget.
Davies, John. 1981. *Kobon*. Amsterdam: North-Holland.
Dedrick, John M. & Eugene H. Casad. 1999. *Sonora Yaqui language structures*. Tucson: University of Arizona Press.
Dehaene, Stanislas. 2011. *The number sense: How the mind creates mathematics*. Rev. edn. Oxford: Oxford University Press.
Derbyshire, Desmond C. 1979. Hixkaryana. Amsterdam: North-Holland.
Dixon, R.M.W. (comp. & ed.) 1991. *Words of our country: Stories, place names and vocabulary in Yidiny, the Aboriginal language of the Cairns-Yarrabah region*. St Lucia: University of Queensland Press.
Döhling.Christian. 2018. *A grammar of Komnzo*. Berlin: Language Science Press
Donohue, Mark. 2008. Complexities with restricted numeral systems. *Linguistic Typology* 12: 423-429.
Drabbe, Peter. 1952. *Spraakkunst van het Ekagi, Wisselmeren, Nederlands Nieuw Guinea*. The Hague: Martinus Nijhoff.
Eiseman, Fred B. Jr. 1990. *Bali: Sekala and Niskala, Volume II: Essays on society, tradition, and craft*. Singapore: Periplus.
Everett, Caleb. 2017. *Numbers and the making of us: Counting and the course of human cultures*. Cambridge, MA: Harvard University Press.
Everett, Daniel L. 2005. Cultural constraints on grammar and cognition in Pirahã: another look at the design features of human language. *Current Anthropology* 46: 621-646.
Georg, Stefan. 2007. *A descriptive grammar of Ket, Part 1: Introduction, phonology and morphology*. Folkestone: Global Oriental.
Gordon, E.V. 1957. *An introduction to Old Norse*. 2nd ed. rev. by A.R. Taylor. Oxford: Clarendon Press.
Greenberg, Joseph H. 1978. Generalizations about numeral systems. In Joseph H. Greenberg, Charles A. Ferguson, & Edith A. Moravcsik (eds.), *Universals of human language, Volume 3: Word structure*, 249–295. Stanford: Stanford University Press. Reprinted (with corrections) in Keith Denning & Suzanne Kemmer (eds.), *On language: Selected writings of Joseph H. Greenberg*, 271–309. Stanford, 1990: Stanford University Press.

Hammarström, Harald. 2010. Rarities in numeral systems. In Michael Cysouw & Jan Wohlgemuth (eds.), *Rethinking universals: How rarities affect linguistic typology*, 11-59. Berlin: De Gruyter Mouton.
Hanke. Thomas. 2005. *Bildungsweisen von Numeralia: eine typologische Untersuchung*. Berlin: Weißensee Verlag.
Hewitt, B. G. 1995. *Georgian: A structural reference grammar*. Amsterdam: John Benjamins.
Holton, David, Peter Mackridge, & Irene Philippaki-Warburton. 1997. *Greek: A comprehensive grammar of the modern language*. London: Routledge.
Holzknecht, Susanne. 1989. *The Markham languages of Papua New Guinea* (Pacific Linguistics C-115). Canberra: Department of Linguistics, Research School of Pacific Studies, The Australian National University.
Honti, László. 1993. *Die Grundzahlwörter der uralischen Sprachen*. Budapest: Akadémiai Kiadó.
Hurford, James R. 1975. *The linguistic theory of numerals*. Cambridge: Cambridge University Press.
Hurford, James R. 1987. *Language and number: The emergence of a cognitive system*. Oxford: Blackwell.
Karpuškin, Boris Mixajlovič. 1964. *Jazyk orija*. Moscow: "Nauka".
Kutsch Lojenga, Constance. 1994. *Ngiti: A Central-Sudanic language of Zaire*. Cologne: Rüdiger Koppe.
Lean, Glendon. 1988. *Counting systems of Papua New Guinea*. 17 vols. Lae: Papua New Guinea University of Technology.
Lean, Glen & Kay Owens,. 2018. Appendix B: Details of counting systems discussed in chapter 8. In Kay Owens & Glen Lean, with Patricia Paraide & Charly Muke, *History of number: Evidence from Papua New Guinea and Oceania*. Cham (Switzerland): Springer
LKN. 1998. *Leksikó tēs koinḗs neoellēnikḗs*. Thessaloniki: Aristotéleio Panepistḗmio Thessaloníkēs, Institoúto tōn neoellēnikṓn spoudṓn.
Martzloff, Jean-Claude. 1997. *A history of Chinese mathematics*. Berlin: Springer. [Translated from French original, 1988].
McGregor, R.S. 1972. *Outline of Hindi grammar with exercises*. Oxford: Clarendon Press.
Meeuwissen, Marjolein. 2004. *Producing complex spoken numerals for time and space*. Nijmegen: Max Planck Institute for Psycholinguistics.
Merlan, Francesca. 1982. *Mangarayi*. Amsterdam: North-Holland.
Owens, Kay. 2001. The work of Glendon Lean on the counting systems of Papua New Guinea and Oceania. *Mathematics Education Research Journal* 13: 47-71.
Pica, Pierre, Cathy Lemer,Veeronique Izard & Stanislas Dehaene. 2004. Exact and approximate arithmetic in an Amazonian indigene group. *Science* 306: 499-503.
Polański, Kazimierz & James Allen Sehnert. 1967. *Polabian-English dictionary*. The Hague: Mouton.
Rajaonarimanana, Narivelo. 2001. *Grammaire moderne de la langue malgache*. Paris: Langues & Mondes – L'Asiathèque.
Saxe, Geoffrey B., with Indigo Esmonde. 2012. *Cultural development of mathematical ideas: Papua New Guinea studies*. New York: Cambridge University Press.
Schapper, Antoinette & Harald Hammarström. 2013. Innovative numerals in Malayo-Polynesian languages outside of Oceania. *Oceanic Linguistics* 52: 423-456.

Skorik, Petr Ja. 1991. *Grammatika čukotskogo jazyka, Čast´ pervaja: Fonetika i morfologija imennyx častej reči*. Moscow-Leningrad: Izd-vo Akademii nauk SSSR.
Smith, Geoffrey P. 1988. Morobe counting systems. In *Papers in Papua New Guinea linguistics No. 26* (Pacific Linguistics A-76). Canberra: Department of Linguistics, Research School of Pacific Studies, The Australian National University.
Smyth, David. 2002. *Thai: An essential grammar*. London: Routledge.
Steenwijk, Han. 1992. *The Slovene dialect of Resia: San Giorgio*. Amsterdam: Rodopi.
Swanton, John R. 1940. *Linguistic materials from the tribes of southern Texas and northeastern Mexico*. Washington DC: United States Government Printing Office.
Whitney, William Dwight. 1889. *Sanskrit grammar, including both the classical language and the older dialects of Veda and Brāhmaṇa*. 2nd ed. Leipzig: Breitkopf & Härtel.
Tozzer, Alfred M. 1921. *A Maya grammar with bibliography and appraisement of the works noted*. Cambridge MA: Peabody Museum.
Yasugi, Yoshiho. 1995. *Native Middle American languages: An areal-typological perspective*. Osaka: National Museum of Ethnology.

Index of Languages

The classification follows Glottolog 4.3 <http://glottolog.org>. The three-letter abbreviations are ISO 639-3 codes; parentheses indicate a partial match.

Adzera	ADZ	Oceanic, Austronesian	Morobe Pr., Papua New Guinea
Balinese	BAN	Malayo-Sumbawan, Austronesian	Bali, Indonesia
Berom	BOM	Benue-Congo, Atlantic-Congo	Plateau State, Nigeria
Chukchi	CKT	Chukotko-Kamchatkan	Chukotka, Russia
Classical Nahuatl	NCI	Uto-Aztecan	Mexico [extinct]
Coahuilteco	XCW	isolate	Texas, USA; Mexico [extinct]
Danish	DAN	Germanic, Indo-European	Denmark
Dutch	NLD	Germanic, Indo-European	Netherlands; etc.
Ekari	EKG	Paniai Lakes, Nuclear Trans New Guinea	Papua, Indonesia
English	ENG	Germanic, Indo-European	England; USA; etc.
French	FRA	Italic, Indo-European	France; etc.
Georgian	KAT	Kartvelian	Rep. of Georgia
German	DEU	Germanic, Indo-European	Germany; etc.
Haruai	TMD	Piawi	Madang Pr., Papua New Guinea
Hindi	HIN	Indo-Aryan, Indo-European	North-Central India
Hixkaryána	HIX	Cariban	Amazonas, Brazil
Italian	ITA	Italic, Indo-European	Italy

Japanese	JPN	Japonic	Japan
Ket	KET	Yeniseian	W. Siberia, Russia
Kobon	KPW	Madang, Nuclear Trans New Guinea	Madang Pr., Papua New Guinea
Komnzo	(TCI)	Morehead-Maro, Yam	Western Pr., Papua New Guinea
Latin	LAT	Italic, Indo-European	Rome [extinct]
Malagasy, Nosy Be (variety of Sakalava Malagasy	(SKG)	Basap-Greater Barito, Austronesian	NW Madagascar
Malagasy, Standard (= Plateau Malagasy)	PLT	Basap-Greater Barito, Austronesian	Madagascar
Mandarin Chinese	CMN	Sinitic, Sino-Tibetan	China
Mangarrayi	MPC	Mangarrayi-Maran	Northern Terr., Australia
Mayan		[language family]	Mesoamerica
Modern Greek	ELL	Greek, Indo-European	Greece, Cyprus
Mundurukú	MYU	Tupian	Amazonas/Mato Grosso/Pará, Brazil
Nauete	NXA	Central Malayo-Polynesian, Austronesian	East Timor
Ngiti	NIY	Central Sudanic	Orientale Pr., DR Congo
Northern Mansi	(MNS)	Mansic, Uralic	W. Siberia, Russia
Northern Pame	PMQ	Pamean, Otomanguean	San Luis Potosí, Mexico
Oksapmin	OPM	Asmat-Awyu-Ok, Nuclear Trans New Guinea	Sandaun Pr., Papua New Guinea
Old Norse	NON	Germanic, Indo-European	Scandinavia; Iceland [extinct]
Oriya	ORI	Indo-Aryan, Indo-European	Orissa, India
Pirahã	MYP	isolate	Amazonas, Brazil
Polabian	POX	Slavic, Indo-European	Germany [extinct]
Resia Slovenian	(SLV)	Slavic, Indo-European	Udine Pr., Italy
Russian	RUS	Slavic, Indo-European	Russia
Sanskrit	SAN	Indo-Aryan, Indo-European	India [extinct]
Spanish	SPA	Italic, Indo-European	Spain; Latin America
Supyire Senoufo	SPP	Senufo, Atlantic-Congo	Mali
Thai	THA	Tai-Kadai	Thailand
Welsh	CYM	Celtic, Indo-European	Wales, United Kingdom
Yaqui	YAQ	Uto-Aztecan	Mexico
Yidiny	YII	Pama-Nyungan	Queensland, Australia
Yucatec Maya	YUA	Mayan	Yucatán, Mexico