

TYPOLOGY OF NUMERALS: MINORITY LANGUAGES OF CHINA

Francisco A. Marcos-Marín
Catedrático de Lingüística General
Universidad Autónoma de Madrid (ESPAÑA)
Ap. 46348 E-28080 Madrid SPAIN

The Department of Linguistics of the Universidad Autónoma de Madrid (UAM) is conducting a research on the Typology of Numerals. This research embraces 500 languages, pertaining to the possible maximum number of families. A part of this research is included in the framework of the Institute of Studies and Exchange with the Far East, a nonprofit organization connected to several departmental activities, including the regular exchange of scholars and students between the UAM and the University of Foreign Studies at Beijing (P.R. China).

The research on numerals includes the preparation of a typological template, in which several features are considered, such as regularity in numeral formation and classification of peculiarities according to units, tens and hundreds, presence or absence of the name of the zero (0) inside the formulation of large names, chain structure (fusion, coordination or juxtaposition) and several other parameters. The proposed pattern also offers a basic typological characterization of the language according to the typical features of Subject Verb Object order and the position of the Adjective versus the Substantive. Some diachronical peculiarities regarding the evolution of addition and subtraction have as well been considered.

In this paper the author will expose some results achieved in the domain of languages of China not belonging to the Chinese group of the Sinitic languages. In order to give a broad vision of the scope of the task, three items will be presented: the typological scheme of the languages already studied (i), the presence of an Iranian language with a vigesimal system in Chinese territory (ii), and some observations concerning a characteristic of (at least) Altaic languages: the Oberstufenzählung (iii).

Typological templates

Abbreviations:

- * indicates an anomalous circumstance, f.i. in multiplication, column of tens, it indicates that they are formed irregularly (*vid. sal-ikur*).
- A** The higher number precedes.
- B** The lower number precedes.

BA Base of the numeration system. Remains of a former system are indicated by means of a bracketed number in the row above.

AS Adjective Substantive.

C Centena, hundred.

D Decena, decade, ten.

F Fusion.

FA linguistic Family, and branch:**IE** Indo-european, **IEIr** Indo-european, Iranian, f.i.

FEs Special Pattern. It indicates which numbers follow the pattern recorded in the line immediately below.

FRe Regular Pattern. It indicates which numbers follow the pattern recorded in the line immediately below.

G General.

Nu Number.

P Partial.

SA Substantive Adjective.

SiAd Adjectival Phrase.

Su Suffix.

Subst Subtraction

U Unit.

V Vinculation.

Y Juxtaposition.

Language Families

IEIr Indo-European, Iranian

STS1 Sino-Tibetan, Sinitic

ATu Altaic, Turkic

AMo Altaic, Mongolian

Amt Altaic, Manchu-tungus

Ktt Kam-Tai, Tai

Ktk Kam-Tai, Kam-Sui

Name	Fa	Ba	Subst.	Addition			Multiplication			Complementary Information											
				G	P	FEs	FRe	Nu	Su	Nu	Su	Nu	Su	MCDU	20	x10	Type	SiAd			
Salikur	IEIr	10	---	- - -		0 0 A	---		*	B00	OBO	---	MCDU	+	+	SOV	AS				
Wux	IEIr	20	---	11-19		21<				OBO	---	OBO	---	MCDU	+	-	SOV	AS			
(class)																					
Chinese	STS1	10	---	- - -		0 A A	---	OBO	---	OBO	---	+	MCDU	-	-	SVO	AS				
(modern)																					
Chinese	STS1	10	---	- - -		A A A	---	OBO	---	OBO	---	+	MCDU	-	-	SVO	AS				
Uygur	Atu	10	---	O0B	0 0 0	0 A 0	---		B00	OBO	---	-	MCDU	+	50	SOV	AS				
Kazakh	Atu	10	---	- - -	(50?)	≤99			0 0 0	0 A 0	---		B00	OBO		-	MCDU	+	50	SOV	AS
Salar	Atu	10	---	- - -	(50)				0 0 0	0 A 0	---		B00	OBO		-	MCDU	+	50	SOV	AS
Tatar	Atu	10	---	- - -	(50)				0 0 0	0 A 0	---		B00	OBO		-	MCDU	+	50	SOV	AS
Uzbek	Atu	10	---	- - -	(50)				0 0 0	0 A 0	---		B00	OBO		-	MCDU	+	50	SOV	AS
Kirghiz	Atu	10	---	- - -	(50)				0 0 0	0 A 0	---		B00	OBO		-	MCDU	+	50	SOV	AS
W. Yugur	Atu	10	---	B00	0 0 0	0 A 0	---		B00		OBO			MCDU	+	20	SOV	AS			
Mongol	Amo	10	---	- - -	(20)	<30			>30				B00	OBO		-	MCDU	-	90	SOV	AS
E. Yugur	Amo	10	---	- - -	(Amo)				0 A 0	0 A 0	---		B00	OBO		-	MCDU	-	90	SOV	AS
Baonang	Amo	10	---	- - -	(Amo)				0 0 0	A 0 0	---		B00	OBO		-	MCDU	-	90	SOV	AS
Dagur	Amo	10	---	- - -	(Amo)				0 0 0	0 A 0	---		B00	OBO		-	MCDU	-	90	SOV	AS
(Tu3)																					
Mongoru	Amo	10	---	- - -	(Amo)				0 0 0	0 A 0	---		B00	OBO		-	MCDU	-	90	SOV	AS
Dong-xiang	Amo	10	---	- - -														SOV	AS		
Oroqen	Amt	10	---	- - -		0 0 0		0 A 0	---		B00	OBO			MCDU	+	90	SOV	AS		
																			130		

¹ From 11 on, they are Chinese loan-words.

Name	Fa	Ba	Subst.	Addition		Multiplication						Complementary Information		
				G	P	FEs	FRe	U	D	C				
				FYV FYV	F Y V	F Y V	FYV FYV	Nu Su Nu Su Nu Su	Nu Su	MCDU	20 x10 Type SiAd			
Evenki	Amt	10	---	-	-	0 0 0	0 A 0	-	-	-	B00 OB0 -	-	MCDU +	90 SOV AS
(Hezhen)						15								30
Nanai	Amt	10	---	-	-	x 0 0	0 A 0	-	-	-	B00 OB0 -	-	MCDU +	90 SOV AS
(Manchu)						15								30
Sibo	Amt	10	---	-	-	x 0 0	0 A 0	-	-	-	B00 OB0 -	-	MCDU +	50 SOV AS
Dai	Ktt	10	---	-	-	0 0 0	0 A 0	-	-	-	OB0 --- OB0 ---	+ MCDU +	- ---	SVO SA
Zhuang	Ktt	10	---	-	-	0 0 0	0 A 0	-	-	-	OB0 --- OB0 ---	+ MCDU -	- ---	SVO SA
Bouyei	Ktt	10	---	-	-	0 0 0	0 A 0	-	-	-	OB0 --- OB0 ---	+ MCDU -	- ---	SVO SA
Mánorán	Ktk	10	---	-	-	0 0 0	0 A 0	-	-	-	OB0 --- OB0 ---	+ MCDU -	- ---	SVO SA

Vigesimal numeration in wux

An ample variety of languages is spoken in the autonomous region of Xinjiang, in China. Those languages belong, usually, to the Turkic family of the Altaic group. In the Western part of the region, at the border of the Sovietic Republic of Tadzhikistan, a Indo-european language is used. It belongs to the Pamir subgroup of the Iranian group, inside the Indo-iranian family. The wux language is spoken in the Autonomous Province of Tâshkùérgântâjiké and neighbouring zones (Gão: 1985, 1-2). This minority is called tajik by the Chinese, who relate it to the one that speaks this last language. The tajik language is spoken, out of China, in the Soviet Tadzhikistán and Uzbekistan, in Afghanistan and Pakistan, but it must not be confounded with the wux. Wux and salikur, another language of the Iranian family, south-eastern group, are languages different from the tajik language, which belongs to the South-Western branch of the Iranian family (Comrie: 1981, 161). Another name for the wux is wahan, (Russian vakh, not to be mistaken for a khanty (ostyak) dialect belonging to the Finno-Ugric branch of the Uralic group called in Russian by this name, Comrie: 1981, 106, sole Uralic language showing a clear ergative construction). The wux presented a regular vigesimal pattern, exposed after the variety called in Chinese wâhâñ, name of a valley in the North-East of Afghanistan (Gão: 1985, 108-109). This dialect serves as an exchange language in the border of China, Afghanistan, Pakistan and the USSR, at the South of the Pamir plateau. It is fragmented in subdialects (Central, Eastern, Western and Northern). The variant to be described is mainly spoken in the Dábùdá village, inside the P.R. of China.

The first decade is formed by:

² When multiplied by the unit, this follows the hundred or the thousand: thousand one = 1000.

³ See note 2.

- | | |
|-----------|---------|
| 1. ji(u) | 6. /aš |
| 2. bu(i) | 7. ɯb |
| 3. tru(i) | 8. at |
| 4. tsɯbɯr | 9. nau |
| 5. pandz | 10. ɔas |

The digits (11) to (20) are formed by juxtaposition of 'ten' and 'unit':

- | | |
|----------------|-------------|
| 11. ɔas jiu | 16. ɔas /aš |
| 12. ɔas bui | 17. ɔas ɯb |
| 13. ɔas trui | 18. ɔas at |
| 14. ɔas tsɯbɯr | 19. ɔas nau |
| 15. ɔas pandz | 20. bist |

From twenty on, decades are formed as multiples of twenty, plus ten, in due case:

- | |
|--|
| 20. ji bist (1 x 20) |
| 30. ji bist e(t) ɔas (1 x 20 + 10) |
| 40. bu bist (2 x 20) |
| 50. bu bist e(t) ɔas (2 x 20 + 10) |
| 60. tru bist (3 x 20) |
| 70. tru bist e(t) ɔas (3 x 20 + 10) |
| 80. tsɯbɯr bist (4 x 20) |
| 90. tsɯbɯr bist e(t) ɔas (4 x 20) |
| 100. pandz bist (5 x 20), and also sad |

The combination of the decades, from (20) on, and the units is achieved by vinculation, with the conjunction et:

- (21) bist et jiu
- (22) bist et bui
- (23) bist et trui
- (43) bu bist et trui

Note that, due to the nature of the vigesimal system, when the decade is not a pure multiple of twenty, the number is expressed by:

pure multiple + et + ten + unit
 (33), so, is "twenty and thirteen".

This observation is utmost needed for the correct typological interpretation, given the possibility of misinterpreting the result as that of a juxtaposed type; actually, it exists coordination, and accordingly to it the construction belongs to those called *by vinculation*.

- (33) ji bist e(t) ɔas trui
- (58) bu bist e(t) ɔas at
- (95) tsɯbɯr bist e(t) ɔas pandz

The wux language belongs to the many tongues possessing a scheme for the addition between (10) and (20) different from the one employed from (20) on.

Oberstufenzählung

We might translate this technical term as *overstepcounting* or, more simply, *overcounting*. It designates a type of construction not to be confounded with the subtractive one: one number (usually above ten) is named by the corresponding unit and the following decade. The Altaic languages offer very clear examples (Ehlers: 1983, PhTF, JbO):

bir yigirmi, in Old Turkish, 11, is, lit., 'one twenty', i.e. 'the first number belonging to the decade that ends in the twenty', as it were, instead of 'one ten' expected.
äki yigirmi, 12, 'two twenty'.

...
toquz yigirmi, 19, 'nine twenty'.

It was believed (Pritsak, *HbO*, 5,1, 1963, p. 40) that this pattern run up to 89, due to the expected ambiguity in the decade from ninety on (Clauson: 1959, 25-26). Clauson himself believes it to be an argument in favour of a system taking us to a period when the Turks did not count above fifty. Ehlers, however, has demonstrated that the system applied also to the tens between 91 and 99, account taken of the following:

92, that will become (overcounting) *ékjí: yü:z* 'two hundred', being then mistaken for 'two hundreds, 200', allows no such construction. It is replaced by the unit followed by the word *örki*, maybe related to the nominal base *ör*, 'height, high', or the verbal *ör-* 'heighten'. The resulting *ékki örki* is, therefore, free from ambiguity. This pattern is valid for 91-99, 191-199, etc.

The system is found in Old Turkish together with the 'regular' constructions on *toquz* (19, ten nine) (Annemarie von Gabain in *PhTF*, I, 32).

Kaare Thomsen (*PhTF*, 321) documents overstepcounting in the language of the *sary uygur* ("Yellow Uygur"), *pär yiþirma 11* (lit. 'one twenty'); *peþiþirma 15* (lit. 'five twenty'); *pär otus 21* (lit. 'one thirty'). The last numeral proves how old is *otus* for 'thirty', whose modern form is *uþon* (lit. 'three ten'). It seems that this language shows overcounting only up to 29.

We have documented overcounting in Western Yugur, as spoken in the Sunan Yugur Autonomous County in Gansu Province (Chen Zong Zhen et al.: 1985, 74-75):

11	<i>birejyirmi</i> (1 20)	16	<i>ahldejyirmi</i> (6 20)
12	<i>þigejyirmi</i> (2 20)	17	<i>jidejyirmi</i> (7 20)
13	<i>hygyjyirmi</i> (3 20)	18	<i>sagisejyirmi</i> (8 20)
14	<i>diorgiyirmi</i> (4 20)	19	<i>dohgiþejyirmi</i> (9 20)
15	<i>beþejyirmi</i> (5 20)	20	<i>jiyirmi</i> and sometimes <i>þigon</i>

For 21 - 29 both possibilities exist; the old form for 30, *ohdis*, is used:

21	<i>birohdís</i> (1 30)	26	<i>ahldohdís</i> (6 30)
22	<i>þigohdís</i> (2 30)	27	<i>jidohdís</i> (7 30)
23	<i>hudzohdís</i> (3 30)	28	<i>sagisohdís</i> (8 30)
24	<i>diorohdís</i> (4 30)	29	<i>dohgiþohdís</i> (9 30)
25	<i>beþeohdís</i> (5 30)	30	<i>hudzun</i> ~ <i>hudzon</i>

Sometimes the regularized forms (*þigon bir* 21 etc.) are used as well.

These data, together with those offered by Pritsak (1955), prove that overstepcounting is a characteristic of the Altaic numerals and help us understand the seemingly complicated evolution of the naming of numbers in those languages.

Overcounting is not to be misunderstood as subtraction. The last construction is also documented in some dialects of (neo)uygur, for some forms between 20 and 100. O. Pritsak (*PhTF*, 547) presents it as a continuation of Chagatai Turkish, in Aksu, *ikki kám otuz*, 28 (lit. 'two for thirty'), *ikki kám qırıq*, 38 (lit. 'two for forty'), variants in other dialects for 39 are *bir qäm qırq*, *bi kem q'x* (lit. 'one for fourty').

Given the huge number of languages studied, it is very convenient to present the partial results of the research to groups of scholars more deeply concentrated in other factors of the languages studied, in order to verify, not the raw data, even if they are not so easily available, but the

connections established between the system of numerals and the general structure of language.

REFERENCES

Note: This research would not have been accomplished, had I not had access to the Libraries of Stanford University, California, University of California, Berkeley, and of the universities of Heidelberg and Kiel, in Germany. I thank the financial assistance provided by the Comité Conjunto Hispano-Norteamericano and the EUROTRA programme of the E.C. (Stanford, 1987), the Comunidad Autónoma de Madrid (California, Berkeley, summer 1989) and the Alexander von Humboldt Stiftung (Heidelberg, between 1987 and 1991, Kiel, 1991).

Abbreviations

AGI	Archivio Glottologico Italiano.	PICCS	Proceedings of the International Conference on Semitic Studies held in Jerusalem 19-23 July 1965.
CRAI	Comptes Rendus de l'Académie des Inscriptions et Belles-Lettres, Paris.	RA	Revue d'Assyriologie et d'Archéologie Orientale, Paris.
HbO	Handbook der Orientalistik, Leiden/ Colonia.	SbÖAW	Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften zu Wien. Phil.-Hist. Klasse.
IJAL	International Journal of American Linguistics.	U-AJb	Uralo-Altaische Jahrbücher.
JAOS	Journal of the American Oriental Society, New Haven, Conn.	WPLU	Working Papers in Language Universals, Stanford.
JJS	The Journal of Jewish Studies, Londres.	WZKM	Wiener Zeitschrift für die Kunde des Morgenlandes.
JNES	Journal of Near Eastern Studies, Chicago.	ZDMG	Zeitschrift der Deutschen Morgenländischen Gesellschaft.
JRAS	Journal of the Royal Asiatic Society.	ZrPh	Zeitschrift für romanische Philologie.
JSS	Journal of Semitic Studies, Manchester.		
MUSJ	Mélanges de la Faculté Orientale de l'Université Saint-Joseph, Beirut.		
PhTF	Philologiae Turcicae Fundamenta. Cf. Bazin et al., Deny et al.		

TITLES

1. Bazin, Louis; Alessio Bombaci; Jean Deny; Tayyib Gökbilgin; Fahir Iz; Helmuth Scheel 1964 *Philologiae Turcicae Fundamenta* tomum secundum curavit et auxiit, indices adiecit Pertev Naili Boratav, Wiesbaden [Aquis Mattiacis]: Franz Steiner. Cf. Deny et al.
2. Chén Zōng Zhèn; Léi Xuǎn Chūn; Biān Zhù 1985 *Xībù yùgù yǔ jiānzhi*, Beijing: mínzú chǔbǎnshé chǔbǎn (dist. Xīn Huá).
3. Clauson, sir Gerard 1959 "The Turkish Numerals," *JRAS*, 19-31.
4. Clauson, sir Gerard 1972 *An Etymological Dictionary of Pre-Thirteenth-Century Turkish*, Oxford: Clarendon Press. Indexes (2 vols.) by A. Róna-Tas, Szeged: Studia Uralo-Altaica, 14 & 15, 1981.
5. Comrie, Bernard 1981 *The languages of the Soviet Union*, Cambridge: University Press.

6. Deny, Jean; Kaare Grønbech; Helmuth Scheel; Zeki Velid 1959 *Philologiae Turcicae Fundamenta* tomus primus, 1 mapa, Wiesbaden [Aquis Mattiacis]: Franz Steiner. Cf. Bazin et al.
7. Ehlers, Gerhard 1983 "Notabilia zur alttürkischen Oberstufenzählung," *U-AJb*, n.f. 3, 81-87.
8. Fritz, Johann Friedrich 1748 *Orientalisch- und occidentalischer Sprachmeister, welcher nicht allein hundert Alphabeta nebst ihrer Aussprache, so bey denen meisten europäisch-asiatisch-africanisch- und americanischen Völkern und Nationen gebräuchlich sind, auch einigen tabulis polyglottis verschiedener Sprachen und Zahlen vor Augen leget, sondern auch das Gebet des Herrn, in 200 Sprachen und Mund-Arten mit dererselben Characteren ... mittheilet ... Aus glaubwürdigen auctoribus zusammengetragen*, Leipzig: C.F. Gessner.
9. Gāo, Ěrqiāng 1985 *Tājiké yǔ jiānzhi*, Beijing: mínzú chūbānhé chūbān (dist. Xin Huá).
10. Greenberg, Joseph H. 1978 "Generalizations about Numeral Systems," en *Universals of Human Language*, ed. J.H. Greenberg, Stanford: Stanford University Press, 3, 249-295.
11. Hamdamov, P. 1983 *Hozirgi Uzbek tilida numerativlar*. Masül muharrir M. A. Asqarova. Toshkent: Uzbekiston SSR "Fan" Nashrieti.
12. Hamp, Eric P. 1970 "On the Altaic Numerals," en *Studies in General and Oriental Linguistics Presented to Shirō Hattori on the Occasion of His Sixtieth Birthday*, ed. Roman Jakobson y Kawamoto Shigeo, Tokio: TEC, 188-197.
13. Jacob, J. 1965 "Notes on the Numerals and Numeral Coefficients in Old, Middle and Modern Khmer," *Lingua*, 15, 143-162.
14. Jakobson, Roman, Gerta Hüttle-Worth & John Fred Beebe 1957 *Paleosiberian peoples and languages. A bibliographical guide*, New Haven, res, de Karl H. Menges en *U-AJb*, 32, 1960, 133-135.
15. Marcos Marín, Francisco 1989 "Los numerales: Contraste y Tipos," *L.E.A.*, XI, 13-53.
16. Marcos Marín, Francisco 1989b "Cuando los numerales no representan número," *Lexis*, XIII, 161-222.
17. Marcos Marín, Francisco 1990 "Notas al azar: para simular lo par y lo impar," en David Hook y Barry Taylor (eds.) *Cultures in Contact in Medieval Spain: Historical and Literary Essays Presented to L.P. Harvey*, Londres: King's College London Medieval Studies, 151-156.
18. Marcos Marín, Francisco 1990b "Lexicología en la formación del sistema de los nombres del número," *Hispanica Posnaniensis*, I, 105-131.
19. Marcos Marín, Francisco 1990c "Remarks on the Syntax of Numerals," *Proceedings of the Fourteenth International Congress of Linguists. Berlin/GDR, August 10- August 15, 1987*, Berlin: Akademie-Vg., I, 792-801.
20. Menges, Karl H. 1968 "Die Tungusischen Sprachen," *HbO*, 5, 3, 21-256.
21. Pott, August Friedrich 1847 *Die quinäre und vigesimalen Zählmethode bei Völkern aller Welttheile*, Wiesbaden: Dr. Martin Sändig.
22. Pott, August Friedrich 1868 *Die Sprachverschiedenheit in Europa. An den Zahlwörtern nachgewiesen, sowie die quinäre und vigesimale Zählmethode*, Halle: Verlag der

- Buchhandlung des Waisenhauses. (Rep. Amsterdam: Rodopi, 1971).
- 23. Pritsak, Omelian 1955 "Die Oberstufenzählung im Tungusischen und Jakutischen," *ZDMG*, 105, 184- 191.
 - 24. Ramsey, S. Robert 1987 *The languages of China*, Princeton, N.J.: Princeton University Press.
 - 25. Róna-Tas, A. cf. Clauson, sir Gerard.
 - 26. Skalička, Vl. 1979 *Typologische Studien*, Braunschweig: Vieweg,
 - 27. Spuler, Bertold et al. (ab 1952) *Handbook der Orientalistik*, Leiden: Brill.

Note: Further references may be found in Marcos Marín: 1989b.