

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** Indoeuropean, **IEIr** Indoeuropean, 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 Indoeuropean, Iranian

STSi 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 FYV	P FYV	FES F Y V	FRe F Y V	U Nu FYV	Su Su FYV	D Nu FYV	C Nu FYV	Su Su FYV								
Salikur	IEIr	10	---	---	- - -	0 0 A	---	---	---	*	B00	0B0	---	---	MCDU	+	+	SOV	AS
Wux	IEIr	20	---	---	11-19 0 A 0	21< 0 0 A	---	---	---	0B0	---	0B0	---	---	MCDU	+	-	SOV	AS
(class) Chinese	STSi	10	---	---	- - -	0 A A	---	---	---	0B0	---	0B0	---	+	MCDU	-	-	SVO	AS
(modern) Chinese	STSi	10	---	---	- - -	A A A	---	---	---	0B0	---	0B0	---	+	MCDU	-	-	SVO	AS
Uygur	Atu	(50?) 10	---	---	S99 00B 0 0 0	0 A 0	---	---	---	B00	0B0	---	---	---	MCDU	+	50	SOV	AS
Kazakh	Atu	(50?) 10	---	---	0 0 0	0 A 0	---	---	---	B00	0B0	---	---	---	MCDU	+	50	SOV	AS
Salar	Atu	(50) 10	---	---	0 0 0	0 A 0	---	---	---	B00	0B0	---	---	---	MCDU	+	50	SOV	AS
Tatar	Atu	(50) 10	---	---	0 0 0	0 A 0	---	---	---	B00	0B0	---	---	---	MCDU	+	50	SOV	AS
Uzbek	Atu	(50) 10	---	---	0 0 0	0 A 0	---	---	---	B00	0B0	---	---	---	MCDU	+	50	SOV	AS
Kirghiz	Atu	(50) 10	---	---	0 0 0	0 A 0	---	---	---	B00	0B0	---	---	---	MCDU	+	50	SOV	AS
W. Yugur	Atu	(20) 10	---	---	<30 B00 0 0 0	>30 0 A 0	---	---	---	B00	---	0B0	---	---	MCDU	+	20	SOV	AS
Mongol	Amo	10	---	---	0 0 0	0 A 0	---	---	---	B00	0B0	---	---	---	MCDU	-	90	SOV	AS
E. Yugur	Amo	10	---	---	- - -	0 A 0	---	---	---	B00	0B0	---	---	---	MCDU	-	90	SOV	AS
Baonang	Amo	10	---	---	0 0 0	A 0 0	---	---	---	B00	0B0	---	---	---	MCDU	-	90	SOV	AS
Dagur	Amo	10	---	---	0 0 0	0 A 0	---	---	---	B00	0B0	---	---	---	MCDU	-	90	SOV	AS
(Tu3) Monguor	Amo	10	---	---	0 0 0	0 A 0	---	---	---	B00	0B0	---	---	---	MCDU	-	90	SOV	AS
Dong- xiang	Amo	10	---	---			---	---	---				---	-	1			SOV	AS
Oroqen	Amt	10	---	---	0 0 0	0 A 0	---	---	---	B00	0B0	---	---	---	30 MCDU	+	90	SOV	AS

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	FYV	FYV	FYV	FYV	0	MCDU	20	x10	Type	SiAd	
Evenki	Amt	10	---	---	0 0 0	0 A 0	---	---	---	B00	0B0	---	-	MCDU	30	+	90	SOV	AS
(Hezhen) Nanai	Amt	10	---	---	15 x 0 0	0 A 0	---	---	---	B00	0B0	---	-	MCDU	30	+	90	SOV	AS
(Manchu) Sibo	Amt	10	---	---	15 x 0 0	0 A 0	---	---	---	B00	0B0	---	-	MCDU	30	+	50	SOV	AS
Dai	Ktt	10	---	---	0 0 0	0 A 0	---	---	---	0B0	---	0B0	---	+	MCDU	+	---	SV0	SA
Zhuang	Ktt	10	---	---	0 0 0	0 A 0	---	---	---	0B0	---	0B0	---	+	MCDU	-	---	SV0	SA
Bouyei	Ktt	10	---	---	0 0 0	0 A 0	---	---	---	0B0	---	0B0	---	+	MCDU	-	---	SV0	SA
Máonàn	Ktk	10	---	---	0 0 0	0 A 0	---	---	---	0B0	---	0B0	---	+	MCDU	-	---	SV0	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 Indoeuropean 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 Tāshikūěrgāntājiké and neighbouring zones (Gǎo: 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 *wǎhǎn*, name of a valley in the North-East of Afghanistan (Gǎo: 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. /aǝ |
| 2. bu(i) | 7. mb |
| 3. tru(i) | 8. at |
| 4. tsmbur | 9. nau |
| 5. pandz | 10. ǝas |

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

- | | |
|----------------|-------------|
| 11. ǝas jiu | 16. ǝas /aǝ |
| 12. ǝas bui | 17. ǝas mb |
| 13. ǝas trui | 18. ǝas at |
| 14. ǝas tsmbur | 19. ǝas nau |
| 15. ǝas 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) ǝas (1 x 20 + 10) |
| 40. bu bist (2 x 20) |
| 50. bu bist e(t) ǝas (2 x 20 + 10) |
| 60. tru bist (3 x 20) |
| 70. tru bist e(t) ǝas (3 x 20 + 10) |
| 80. tsmbur bist (4 x 20) |
| 90. tsmbur bist e(t) ǝas (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) ǝas trui |
| (58) bu bist e(t) ǝas at |
| (95) tsmbur bist e(t) ǝas 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'.