Word formation in the cardinal number systems across languages
(Mathematical problems in numbers)

Hoang Thi Chau*

College of Social Sciences and Humanities, VNU
336 Nguyen Trai, Thanh Xuan, Hanoi, Vietnam

Abstract. The cardinal number system is addressed in this paper as a small yet typical sign system in its larger sign super-system – language. The bifaceted nature of signs is clearly demonstrated: most numbers have double meanings, one is structural while the other is lexical. For instance, in Vietnamese, the number "forty four" lexically denotes "the next number after 43 in natural number chains", and structurally means $4 \times 10 + 4$. Meanwhile, in French, number "80" is *quatre-vingt* (4×20) with the lexical meaning of "eighty" while it structurally means "4×20"; and number "70" *soixant-dix* has the structure "$60 + 10$".

Deep under the addition and multiplication problems mentioned above are mathematical thoughts and languages of these nations: the Vietnamese is based on the decimal numeration while the French use the degisemal numeration.

The data used in this analysis are restricted within languages of ethnic minorities in Vietnam and two major language families in Southeast Asia, namely Austroasiatic and Austronesian, or to be more precise, Austro-Tai, which are closely related to the Vietnamese language. In order to clearly identify features of word formation in the numbers of isolating, analytic languages in Vietnam and Southeast Asia, comparative and contrastive analyses have been made against the number systems of Indo-European languages representative of the synthetic typology and familiar to us, including French, English and German, before the following conclusions are made:

1. Numbers are the most basic words among the basic, i.e. the most ancient. The number system of each nation has been developed through a long process of thousands of years. The 10 basic words have transformed into 100 others by employing one or more of these: morphological change, affixation, word combination, and word order change. The number systems are highly systematic with fixed structure. However, partial or total borrowings of an entire system are still possible.

2. This is an initial structural investigation of the linguistic form of the numeration systems with a view to identifying the steps in the formation of language and thought in a narrow sense.

3. The investigation of the number systems in different European and Asian regions, across various language families, helps reveal their universalities as well as typicalities in their language formulation and development process.

The cardinal number system is addressed in this paper as a small yet typical sign system in its larger sign super-system - language. The bifaceted nature of signs is clearly demonstrated:

_____

Tel.: 84-4-8531855

most numbers have double meanings, one is structural[1] while the other is lexical. For instance:

- In Vietnamese, the number "forty four" lexically denotes "the next number after 43 in natural number chains", and structurally means 4x10+4 (with the multiplication "x" and the plus "+" signs left implicit). Actually the multiplication "x" has been represented in "mươi" (with a level tone), which differs from "mười" (ten, with a falling tone): "mười bốn" (14) = 10+4 vs. "bốn mươi" (40) = 4x10.

- In French, number "80" is *quatre-vingt* (4x20) with the lexical meaning of "eighty" while it structurally means "4x20"; and number "70" *soixant-dix* has the structure "60+10".

- In Pazeh, an indigenous language in Taiwan, number "7" is formed by combining the two words *xaseb* – "5" and *dusa* – "2" together as *xasebidusa* (5+2). Their ancestors left Taiwan around 6,000 years ago for the Pacific islands which have become present nations of the Philippines, Indonesia, Malaysia, Brunei, New Zealand, etc. whose languages are commonly referred to as Malayo-Polynesian and, on a larger scale, Austronesian. In these languages, *xasebidusa* has been shortened as *pitu. [1, p.415-422]

Deep under the addition and multiplication problems mentioned above are mathematical thoughts and languages of these nations. The Vietnamese mathematical thought is expressed in "mươi", "mười", "một chục" (one ten) on the decimal numeration basis while the French use the degisemal numeration (80=4x20) and yet the Taiwanese count their fingers[2].

For these reasons, the paper is entitled "Word formation in the cardinal number systems" or "Mathematical problems in numbers".

The systemic feature and national characteristics are also clearly demonstrated, even when the number system is borrowed from another language. The number systems of Thai languages such as Thai in Thailand, Laotian, Thai in Yunnan, China, Thai in Northwestern Vietnam, Tay-Nung in Northeastern Vietnam, etc., all were borrowed from ancient Sino, though having their own peculiarities (which will be analyzed hereafter). The listeners, therefore, can recognize the numbers of Thailand's people as distinctive from those of the Nung people in Phan Sinh.

Apart from the number systems in all languages, there is a supra-national, supra-lingual system of digits, usually referred to as the Arabic digits, which serves as a major, advantageous source of reference and contrast for our research.

The data used in this analysis are restricted within languages of ethnic minorities in Vietnam and two major language families in Southeast Asia, namely Austroasiatic and Austronesian, or to be more precise, Austro-Tai, which are closely related to the Vietnamese language. In order to clearly identify features of word formation in the numbers of isolating, analytic languages in Vietnam and Southeast Asia, comparative and contrastive analyses have been made against the number systems of Indo-European languages representative of the synthetic typology and familiar to us, including French, English and German, before conclusions are made. Finally, through the comparison of the number systems across modern languages over a large space, the paper

---

[1] Russian linguists refer to this as "internal form of linguistic units".

[2] In Taiwanese indigenous languages, initially there were only 5 basic numbers from 1 to 5. The numbers from 6 to 9 are formed by placing the first five numbers next to each other, showing the addition (6=5+1, etc.), or substraction

(9=10-1, only with number 9), multiplication (6=3x2; 8=4x2, only with numbers 6 and 8).

seeks to enable the readers to visualize their evolution from about 10,000 years ago to 1,000 years ago, and their relations of ancient origin.

1. Austro - Tai language family

1.1. Thái Branch

As has been mentioned, all languages of the Thai branch did not preserve their primitive number systems; instead, they borrowed the number system of ancient Sino over 1,000 years ago, prior to the migration of the Thai people in their southward expansion to as far as Malay peninsula and westward to India so that they could occupy a vast territory over the present-day Southeast Asia.

At first glance, the number systems of the Thai, Tay-Nung in Vietnam and in Thailand seem similar, but a closer look reveals several phonetic, lexical and word combination differences across these three systems.

1.1.1. Phonetic differences

- Due to phonetic asystemic transformation:

sl/s (Tay-Nung / Thai, Thailand: *slong, lam, sli / song, sam, si* - "2", "3". "4")

nh/j (Thai / Thailand: *nhi / ji* - "2")

h/r (Thai / Thailand: *họi / rói* - "100")

p/ph (Thai / Thailand: *pắn / phan* - "1000")

- Due to word combination:

Thai: *sip-êt > met* ("11")

    song-síp > sảo ("20")

1.1.2. Lexical differences

Use of different synonyms

	Thai	Tay - Nung	Thailand
"100"	họi nùng	nâng pac	nựng rói
"20"	sảo	nhi slip	ji sip
"22"	sảo song	nhi slong	ji sip song

e.g. *song* and *nhi* (both meaning "2"); *nựng êt* - "1".

1.1.3. Differences in word combination

Due to Vietnamese influence on Thai and Tay-Nung, there are simplified combinations, e.g.

	Viet	Thai	Tay-Nung	Thailand
"25"	hăm lăm	sảo hà	nhi hà	ji sip ha
"140"	trăm tư		pac sli	nựng rói sị sip

1.2. Kadai branch

In both Vietnam and China, the languages are called Kadai as their speakers resided in an interwoven manner among other ethnic groups like Thai, Tay, Nung, Hmong, Yao, which increasingly narrowed down their areas of use and even posed them the threat of extinction. Normally, the old Laha people speak the Laha language, while younger generations have very poor proficiency in their mother tongue which is spoken only at home. For social communication, Thai and Vietnamese languages are used. In the population records of Lao Cai, Yen Bai provinces, several villages of the Lachi people are listed, but when we arrived there, we found no one speaking the Lachi language. Instead, the locals use Tay or Nung and the popular (i.e. Vietnamese) language in their day to day communication. However, wherever the mother tongue is preserved, its number system remains, including interesting cases of language mix like the number system of Laha.

	Laha	Pupéo	Lachi
"1"	căm	cja	caŋ
"2"	sa	sê	fu
"3"	Tʌ/aw	tăw	te
"4"	pa	pê	pu
"5"	hɛ ma	ma	m
"6"	Hɔk drăm	(ma)ɲăm	na
"7"	cet to	(ma)tu	te
"8"	pet mahu	(ma)zi	ba
"9"	daw sawa	(ma)sja	lju
"10"	Pat	pʌt	pe

The table above shows that numbers from "1" to "4" are monosyllabic Laha words, and numbers from "5" to "10" are disyllabic, with the former syllable being Thai word (actually ancient Sino) and the latter its synonym in Laha. This type of parallel or semantico-repetitive compounds (with two synonymous words of different origins) is very commonly encountered in Southeast Asian languages, like *chó má (dog dog* – Viet and Thai), *tre pheo (hamboo bamboo* – Viet and Muong).

Despite the number system being mixed, the separation of Sino and Thai elements reveals that Laha numbers are closer to the etymological numbers of Proto Malayo-Polynesian (PMP); Pupeo number system comes next, and last Lachi. More surprisingly, the phonetic correspondence in Laha numbers with PMP[3] is even clearer compared to Austronesian languages in Vietnam (e.g. Cham, Ede, Giarai, etc.)

It is probably not a mere coincidence when *mahu* – number 8 in Laha, *(mə)sja* – number 9 in Pupeo are very similar to *mahal* - number 8 and *(me)siya* - number 9 in Papora language in Taiwan. [1, p.414].

It should be added that the word "hand" in all Tai-Kadai languages is a homophone of number 5 in these languages and originated from proto Malayo-Polynesian

*lima > ma, mù, m

In most Taiwanese languages at present, *lima* - number 5 has replaced an ancient word *RaCep – number 5 which used to form the numbers from 6 to 9: 5+1; 5+2; 5+3; 5+4 [1, p.414-423] and later they were shortened as PMP numbers: *enem - "6", *pitu - "7", *walu - "8", and *siwa - "9".

Similarities in the number systems of Kadai languages in Vietnam and Malayo-Polynesian languages support Austronesian researchers' hypothesis on the genetic relationship between these two language groups and Taiwanese indigenous languages, and the migration path of the Tai-Kadai ancestors might have crossed Taiwan before their arrival in the mainland of present-day Southeastern China and Northeastern Vietnam about 6,000 years ago [1, 438-439]. This is also the date of the two archeological sites of Da But and Quynh Van in Thanh Hoa and Nghe An provinces respectively. Was this possibly the landing site of ancient Tai-Kadai people? They had lived on hunting and gathering for millennia of years there, leaving numerous dumps of shells as *shell hills and mountains* at present [2, p.8-9].

1.3. Malayo-Polynesian branch in Vietnam

Ancestors of the Cham people left Taiwan almost at the same time with the Tai-Kadai, yet on another direction to the Oceanic continent and stayed there for thousands of years, developing from hunting and gathering tribes to agricultural communities before making their way to Vietnam.

Sa Huynh Culture was discovered on the Central Coastal plains to Southeastern Vietnam, which dated back to around 500B.C.E and was identified as belonging to highly-developed agricultural inhabitants. What were buried together with them in their jar tombs include iron and copper arms and gold jewelries made with highly sophisticated techniques. "Those were the tribes which formed the ancient kingdom of Champa at the beginning of our Common Era" [3-5].

Hereafter is the comparison of the number systems in Malayo-Polynesian languages in Vietnam.

---

[3] Proto Malayo - Polynesian: primitive Malayo-Polynesian language

	PMP[4]	East Chăm[5]	West Chăm[6]	Chru	Roglai	Êdê	Giarai
"1"	*isa	tha	sa	sa	sa	sa	sa, ha
"2"	*dusa	twa	doa	dua	dua	dua	dua
"3"	*telu	klow	klau	klou	tlou	tlâo	klâo
"4"	*sepat	pa'	pak	pà	pàq	pă	pă
"5"	*lima	limɯ	lamɯ	loma	lumã	ĕma	(rɔ)ma
"6"	*enem	năm	năm	nam	năm	năm	năm
"7"	*pitu	taçuh	tajuh	tojuh	tijuh	kjuh	(to)juh
"8"	*walu	talipăn	tapăn	tolpan	lapat	sapăn	čopăn
"9"	*siwa	thalipăn	samlăn	solpan	salapat	duapăn	dua (rɔ)păn
"10"	*puluq	pluh	ha pluh	spluh	sapluh	pluh	pluh

The above table shows that, in comparison with the etymological form of Malayo-Polynesian languages (PMP), the number systems in Cham, Ede, Giarai, Chru, Roglai, etc., in Vietnam do have similar phonetic forms from 1 to 6 and 10 with a clear inclination toward monosyllabic forms. Number 7 - *tojuh* is identical among all the languages under comparison (with certain variants of the vowels i/u/o/a in their proto-syllables) whose origin is yet unknown. Numbers 8 and 9 are formed by combining numbers 1 and 2, and this combination can only be explained by such addition, substraction and multiplication as has been seen in Taiwanese languages. It might also have been the substraction: 8=10-2; 9=10-1.

In the compared languages, number 1 is *tha, sa, sɔ*, number 2 is *dua, da, ta*. Yet number 10 in the table shows up as *pluh*, which might have been a new form in substitution of an older one. Number 10 – *labatan* in Ketagalan in Northeastern Taiwan[7] is closer to the latter part of numbers 8 and 9 of the compared languages above *labatan > lapan, lipan, pan, păn*, etc.

It is possible to divide the compared languages into 3 groups:

1.3.1. East Cham, West Cham, Chru and Malay group with numbers 8 and 2 (*da, ta, to*) (numbers 9 and 1 *tha, sa, sɔ*). It can be assumed that these are results of substraction: 10-1 = 9; 10-2 = 8

	"8"	"9"
East Chăm	dalipăn	thalipăn
West Chăm	tapăn	samlăn
Chru	tolpăn	sopăn
Malay	lapan	somilan

1.3.2. Ede and Giarai Group uses these numbers in the reverse order, as number 1 in forming number 8, and number 2 in forming number 9. This could be temporarily explained that the Ede and Giarai people, when borrowing numbers 8 and 9 from the Cham, were confused, without any knowledge that numbers 2 and 1 were actually inside numbers 8 and 9.

	"8"	"9"
Êdê	sapăn	duapăn
Giarai	čopăn	dualăn, dua rɔpăn

1.3.3. Roglai language:

"8"	"9"	"10"	"11"
lapat	salapat	sa pluh	sa pluh sa
2x4	1+8	one ten	one ten one

We hypothesize that number 8 - *lapat* could have resulted from the combination of numbers 2 and 4 with a multiplication: *dua *sepat (2x4 = 8). Number 9 is 1+8: *salapat*.

[4] Proto Malayo-Polynesian
[5] Chăm in Ninh Thuận and Bình Thuận provinces
[6] Chăm in South Vietnam
[7] The habitat of Malayo-Polynesian ancestors prior to their scattering over the Pacific islands [1, p.437, Fig 2; 431 Table 4]

The four numbers above represent two different ways of thinking about the number system:

Number 8 - *lapat* (2x4), 9 - *salapat* (1+8): counting the fingers, a more ancient way.

10 – *sa pluh* (one ten), 11 – *sa pluh sa* (one ten one): based on the decimal system, which was newer, more popular, has been in existence until now and built up the number system to infinity. A closer investigation can show us the clear mismatch in the two combined parts in Roglai number system. The mismatch lies in number 9 - *salapat* and number 10 - *sa pluh*. *Salapat* - "9" implies addition (1+8). If such pattern had continued, *sa pluh* would have been 1+10 = 11. Yet, in reality, *sa pluh* - "one ten" lies in the upper part of the system, was constructed later and aligned with *dua pluh* - "two ten", *tlou pluh* – "three ten", etc.

Next, we can see *một chục* (one ten) used in place of 'mười' (ten) is a characteristic of the number systems in Mon-Khmer language which emphasizes the decimal system – a advance in numbering.

2. Austroasiatic Family and Mon-Khme Branch in Vietnam

The number systems of Mon-Khme languages in Vietnam were also built upon those two mathematical thoughts. Khme numbers typify finger-counting, while other use the decimal system. However, the number are quite identical across all these languages.

2.1. Numbers from 1 to 5

Numbers from "1" to "5" display high identicality across all the language groups, with the exception of Katu as in number 5 - *sŏng*.

| Branch \\ Numbers | Khmer | Bahnar South | Bahnar North | Katu | Viet-Muong | |
	Khmer	Chrau	Rơngao	Bru	Rục	Việt
"1"	muôi	muôi	mŏi	muôi	mộc	một
"2"	pi	var	bar	bar	hal	hai/ vài
"3"	bây	pe	pĭ	pái	pɔ	ba
"4"	buôn	puơn	pú	pŏn	pon	bốn
"5"	pram	prăm	podăm	sŏng	dam	năm

2.2. Numbers from 6 to 9

On the basis of the numbers from 6 to 9, it is possible to categorize the following:

- Khmer branch: combining numbers
- Bahnar North: had proto syllable
- Bahnar South and Viet-Muong: had initial consonant clusters, very similar to Vietnamese language.

It is a complete difference in Katu branch

| Branch \\ Numbers | Khmer | Bahnar North | Bahnar South | Viet-Muong | Katu | |
	Khmer	Rơngao	Chrau	Rục	Việt	Bru
"6"	prăm muôi	tơdrú	prau	phrau[1]	sáu	tapơâ
"7"	prăm pi	tơpâih	poh	paj[4]	bảy	tapul
"8"	prăm bây	tơham	pham	tham	tám	takual
"9"	prăm buôn	tơchĭn	sɯm	chin[3]	chín	Takêh
"10"	đóp	mŏi jât	mât	mươj[2]	mười	muôi chɯt

These number ranges are ordered from left to right, reflecting phonetical changes from ancient to modern time.

2.3. Number 10 - "one ten" in Món - Khmer languages

Number 10, with the exception of *đóp* in Khmer, and "mười" in Viet-Muong group, is the same "muôi chit" ("one ten") in all other languages with such phonetic variants as:

muôi chit / mười chưt / mồi chat...

mồi jêt / mồi jât / moi jơt / mát...

"One ten" is a critical number of the decimal system, marking a novel progress from finger counting, and forming the basis for continuation:

In Bru for instance:

"11": *muôi chưt la muôi (one ten and one)*

"12": *muôi chưt la bar (one ten and two)*

"20": *bar chưt (two ten)*

Also 'one ten' is used as a basis to count larger numbers like 'one hundred, one thousand', and 'one million'. In other languages, Indo-European ones, for example, just 'hundred, thousand, million', etc., are used without the preceding word 'one'.

In Chrau language (Bahnar South), "10" is *mát*, shortened from *muôi jât* – "one ten"; "12" is *mát var*; but "20" is *var jât*, "30" is *pe jât* [6, p.76-77]. It is possible to hypothesize that in Vietnamese in the past, "10" was *mười chât/jât*, which was repeated in various combinations – 11 (one ten and one), 12 (one ten and two), etc., and finally reduced to only two words – the initial and the final ones; "mười" stands at the front to bear the lexical meaning of the whole combination and takes on a new meaning 'one ten'. This is similar to the lexical semantic change of "đun" and "thổi" to "nấu" in modern Vietnamese:

"thổi lửa nấu cơm" > "thổi cơm";
"thổi" changed its meaning into "nấu"
blow fire boil rice > blow rice
"Cook rice."
"đun củi nấu nước" > "đun nước"; "đun"
changed its meaning into "nấu"
push wood boil water > push water
"boil water."

Today, electricity is used to cook rice and boil water, but the language remains "thổi cơm" and "đun nước".

Forming numbers with addition and multiplication is differentiated in Vietnamese by changing the tones and word orders:

mười hai	mười ba	mười bốn	hai mươi	ba mươi
ten two	ten three	ten four	two ten	three ten
10+2	10+3	10+4	2x10	3x10

or by phonetical change, reduction or blending:

"hai mươi" > "hăm"; "ba mươi" > "băm"

hăm mốt	băm hai	hăm lăm/nhăm	bốn tư
twenty one	thirty two	twenty five	forty four
20+1	30+2	20+5	40+4

and "một trăm tư" (one hundred four – 140) distinct from "một trăm linh bốn" (one hundred and four – 104).

These investigations reveal that word formation in the number systems in Vietnamese as well as other isolating languages does not only involve the combination of word and word order change, as we have so far believed, but also phonetical transformation similar to morphological change of Indo-European languages, and this phonetical change has been used in a very flexible way.

2.4. Khmer numbers

Khmer number system is a perfect model of the combination of finger counting and the decimal system, of basic Khmer words and loans from ancient Sino. It is interesting that

ancient Sino borrowings are used as foundation for upper orders of the decimal system, and function as the basic word in word formation. What were borrowed were the materials to build up Khmer number system, which differs from Thai languages that borrowed the whole ancient Sino number system.

The use of alien loan words to make the 'ten' like *sam sǎp* – "30", *se sǎp* – "40" would have resulted in their loss of word-forming meaning, leaving only the lexical one, i.e. total elimination of multiplication (3x10), which leaves only the result (30). The entire system merely uses simple addition, as simple as word combination. For instance, *sam sǎp muôi* – "31" (three ten - one). It suffices for the user just to add the 'ten' digit to the unit digits in Khmer language to count till 100[8].

dóp	"10"	muôi	"1"	*môphey muôi* = "21"
môphey	"20"	pi	"2"	
sam sǎp	"30"	bây	"3"	*sam sǎp buôn* = "34"
se sǎp	"40"	buôn	"4"	
ha sǎp	"50"	prǎm	"5"	*se sǎp prǎm muôi* = "46"
hôt sǎp	"60"	prǎm muôi	"6" (5+1)	
chét sǎp	"70"	prǎm pi	"7" (5+2)	*ha sǎp prǎm bây* = "58"
pét sǎp	"80"	prǎm bây	"8" (5+3)	*hôt sǎp prǎm buôn* = "69"
kâu sǎp	"90"	prǎm buôn	"9" (5+4)	
muôi roi	"100"			

Another comparison of the numbers from 1 to 100 in Indo-European languages can provide us with an impressive and comprehensive view of the formation and content of the language substance.

3. Number systems in Indo-European languages

3.1. The following table can provide a useful comparison of the numbers less than 10, over 10

and tens in French, English, German an Vietnamese:

French	3.trois	4.quatre	5.cinq
	13.treize	14.quatorze	15.quinze
	30.trente	40.quarante	50.cinquante
English	3.three	4. four	5.five
	13.thirteen	14.fourteen	15.fifteen
	30.thirty	40.forty	50.fifty
German	3. drei	4.vier	5.fünf
	13.dreizehn	14.vierzehn	15.fünfzehn
	30. dreizig	40.vierzig	50.fünfzig
Vietnamese	3.ba	4.bốn	5.năm
	13.mười ba	14.mười bốn	15.mười lăm
	30.ba mươi	40.bốn mươi	50.năm mươi

3.2. Through the presented examples, it i. possible to conclude as follows

In French and all the three Indo-Europea languages under comparison, the numbers les: than 10 are common roots, while the number: above 10 and the 'tens' are derivatives in boi ways: morphological change and suffixation.

e.g.   trois > treize

trente

English and German mostly use word combination, with slight phonetical change like "mười" and "mươi" in Vietnamese. In addition Vietnamese changes the word orders, following the orders of the number, i.e. the unit stand: after the ten.

| "13" | mười ba | thirteen | dreizehn |
| "30" | Ba mươi | thirty | dreizig |

3.3. Apart from the afore-mentionec similarities, the compared number systems shou the following differences:

3.3.1. In German, the order of the odc numbers above 10 (from 21 to 99) is reversed the unit is place before the ten

"21" - *ein und zwanzig (one and twenty)*

"99" - *neun und neunzig (nine and ninety)*

3.3.2 In French, the numbers less than 70 are formed on the basis of the decimal numeration, while those above 70 use the vigesimal numeration:

"60" - *soixante (sixty)*

"70" - *soixante-dix (sixty-ten)*

"71" - *soixante-onze (sixty-eleven)*

"80" - *quatre-vingt (four-twenty)*[9]

"90" - *quatre-vingt-dix (four-twenty-ten)*

"99" - *quatre-vingt-dix-neuf (four-twenty-en-nine)*

Such uniqueness and originality of the German and French number systems were originated from the ancient language of Celtic class, particularly Gauls[10], the language of the indigenous inhabitants who used to dwell upon the territory of present-day France and Germany thousands of years B.C.E. until the Roman Empire expanded to all over Europe.

3.4. The Gaul number system is wonderful evidence which enables us to visualise a pattern for building up the numbers from 1 to 100 using the degisemal numeration, which is carried out order by order like this:

Order 1: Use 10 basic numbers, just like other Indo-European languages:

1. un   2. dau   3. tri   4. pedwar 5. pump
6. chwech 7. saith 8. wyth 9. naw   10. deg

Order 2: Combine numbers to make 11 to 15 in a consistent manner throughout the system: the basic number 10 - *deg* always assumes the final position.

"11" - *undeg (one ten)*

"12" - *daudeg (two ten)...*

"15" - *pumtheg (five ten) (pd > th)*

Order 3: Form numbers 16 to 19 on the basis of 15

"16" - *un ar bymtheg (one and fifteen, 1+15) (p > b)*

"19" - *pedwar ar bymtheg (four and fifteen, 4+15)*

Order 4: Start to build up numbers using the degisemal numeration, taking 20 as the most basic number for the whole system: "20" - *ugain* (changed into *vingt* in French). From now on, *ugain* always takes the final position, except for number 50. The next basic numbers include:

"40" - *deugain (two – two ten, 2x20)*

"60" - *trigain (three – two ten, 3x20)*

"80" - *pedwar ugain (four – two ten, 4x20)*

Between these landmark numbers are combinations with 1 to 19.

e.g. "21" - *un ar hugain (one and two ten, 1+20)*

"22" - *dau ar hugain (two and two ten, 2+20)*

This inverse order was imported into German to form numbers from 21 to 99.

e.g. "21" - *ein und zwanzig (one and twenty, 1+20)*

"99" - *neun und neunzig (nine and ninety, 9+90)*

Numbers 30, 70 and 90 are not marked basic numbers like those in the decimal numeration; rather, they are merely ordinary numbers, e.g. in Gauls.

"30" - *deg ar hugain (ten and two ten, 10+20)*

"70" - *deg a trigain (ten and three – two ten, 10+3x20)*

"90" - *deg a pedwar ugain (ten and four – two ten, 10+4x20)*

in French:

"70" - *soixante-dix (sixty-ten, 60+10)*

"71" - *soixante-onze (sixty-eleven, 60+11)*

---

[9] In English, for a time, *score* was also used as number 20, and number 80 was also formed by combining 'four scores', very much like French.

[10] The language of the Gauls is still used as a living language (local tongue) in Wales, Southwest of the UK. See reference 28.

"90" - quatre-vingt-dix (four-twenty-ten, 4x20+10)

"91" - quatre-vingt-onze (four-twenty-eleven, 4x20+11)

It is clear that French uses the degisemal numeration of the Gauls to build up the latter part of their number system (from 70 to 99) (see 3.2.2).

From number 100 – *cant*, a common number among Indo-European languages, the Gauls people created number 50 – *hanner cant*, i.e. ½ of 100. This is probably a new number 50 in substitution for an older one formed by using the degisemal numeration.

3.5. The universality of the number system structure

3.5.1. The structure of the Gauls numeration system enables us to visualize an impressive five-storey ancient castle, with the foundation and upper floors, which differs from modern squarish high-rise blocks, i.e. the decimal system which is becoming increasingly popular and allows one to count until infinity thanks to its simplicity and convenience in massive assemblies. The structure of the degisemal numeration just examined in Gauls is self-evident with regards to its age. It starts from counting parts of the human body: the two hands, and then the two feet. Such counting may possibly have existed long among the primitive tribes living on hunting and gathering. When humans knew how to domesticate animals and did agricultural farming, the number systems were topped up with higher floors with various ways of thinking on the basis of degisemal or binary numeration in different localities, as expressed in their languages.

3.5.2. The number systems in Austroasiatic and Austronesian (Austro-Tai) families also display similar structure in the foundation – finger counts. In Austroasiatic languages, only

Khmer retains this finger-count manner number 6 - *prăm muôi* (5+1); "7" - *prăm ¡* (5+2); "8" - *prăm hây* (5+3); "9" - *prăm buc* (5+4) as a kind of museum. In other Moi Khmer languages, they are replaced wii numbers "six, seven, eight" and "nine", formin 10 basic numbers which serve as the foundatic for the decimal numeration, and a special fori is given to number 10: *muôi chit* (one ten).

3.5.3. Austro-Tai languages build up the numbers in two ways:

- Languages of the Thai branch borrowe the numbers from ancient Sino which used th decimal numeration.

- Kadai and Malayo-Polynesian language possess a primitive number system: finger coun (6 = 5+1,..., 9 = 5+4), which is still retained ii several indigenous languages in Taiwan. Th contracted results of this addition problem (5+1 5+2, 5+3, 5+4) are the four numbers "6" *enem; "7" - *pitu; "8" - *walu; "9" - *siwa ii PMP, which are being used in the remaininį Taiwanese languages and present-day Kada and Malayo-Polynesian languages. Then thι system continues to build upon the decima system. Kadai and Malayo-Polynesiar languages in Vietnam also preserve such tracε: in its own way.

In sum, the differences among Austroasiatic and Austronesian languages in the bases of thε numbers lie in that Austronesian use thε contracted forms of old numbers whilε Austroasiatic use new replacements.

4. Conclusion

1. Numbers are the most basic words among the basic, i.e. the most ancient. The numbeɪ system of each nation has been developec through a long process of thousands of years. The 10 basic words have transformed into 100 others in the following manners:

- Morphological change
- Affixation
- Word combination
- Word order change

The number systems are highly systematic with fixed structure. However, partial or total borrowings of an entire system are still possible.

2. This is an initial structural investigation of the linguistic form of the numeration systems with a view to identifying the steps in the formation of language and thought in a narrow sense.

3. The investigation of the number systems in different European and Asian regions, across various language families, helps reveal their universalities as well as typicalities in their language formulation and development process.

References

[1] Reinecke, A., Nguyễn Chiều, Lâm Thị Mỹ Dung, "New Findings of Sa Huynh Culture", Linden Soft, 2002.

[2] Thomas, D., Thô Sàng Lục, "Chrau Vocabulary", SIL 1966.

[3] Cohen. P., "Jeh Vocabulary", SIL, 1970.

[4] Heine - Geldern, R., "Primitive Geographical Territory and the Earliest Migrations of the Austronesians", Anthropers. vol. XXVII (1932), 543-619.

[5] National Center for Social Sciences and Humanities, History Institute, "Vietnam - Historic Events (from ancient time to 1858)", Education Publisher, 2001.

[6] Smith, K.D., "A phonological reconstruction of Proto - North - Bahnaric. Language Data", Asian Pacific Series No2, SIL, 1972.

[7] Awoi-hathe, Aviong..., Roglai Glossary, Collection of Vietnamese Ethnic Minority Languages, vol. 3, part 1, p. 99.

[8] Benedict, Paul K, "Austro - Thai: Language and Culture", New Heaven: HRAF.

[9] Benedict, Paul K. "Thai - Kadai and Indonesian: a new alignment in Southeastern Asia", American Anthopologist, No 44,2, 1942.

[10] Bùi Khánh Thế (ed), "Việt – Chăm Dictionary", "Chăm - Việt Dictionary", Social Sciences Publisher, Hanoi, 1996.

[11] Burusphat S., "Elementary Thai for foreigners", Mahidol University, 1991.

[12] Đoàn Văn Phúc, "Phonetics of Êđê", Doctoral Dissertation, Hanoi, 1993.

[13] Edmondson, Jerold A., Devid B. Solnit (eds), "Comparative Kadai: linguistics studies beyond Tai", Dallas: SIL/UTA Series in Linguistics.

[14] Fuller, E., "Chru Vocabulary", SIL. 1997.

[15] Gregerson, Kenneth, Marilyn, "Rengao Vocabulary", SIL, 1977.

[16] Hoàng Lương, "A glimpse at the Kadai ethnic communities of Vietnam: an anthropological study", Discussions in Kadai and SE Asian Linguistics. Coordinator: Jerold A. Edmondson IV, 10/1994.

[17] Hoàng Tuệ, (ed), "Learning Pakôh – Taôih languages" textbook, People's Committee of Bình Trị Thiên Province, 1986.

[18] Hoàng Văn Ma, Tạ Văn Thông, "The Language of Bru - Vân Kiều", Social Sciences Publisher, Hanoi, 1998.

[19] Hoàng Văn Ma, Vũ Bá Hùng, "Pupeo Language", Social Sciences Publisher, Hanoi,1992.

[20] Kosaka R., "A Descriptive Study of the Lachi Language - syntactic description historical reconstruction and genetic relation", Doctoral Dissertation, University of Foreign Languages, Tokyo, 2000.

[21] Lâm Đồng Province, "Việt – Koho Dictionary", 1983.

[22] Mussay, G., "Dictionnaire Căm - Vietnamien - Francais", Phan Rang, 1971.

[23] Nguyễn Phú Phong, Trần Trí Dõi, Ferlus M., "Lexique Vietnamien Ruc-Francais", Paris 1988.

[24] Remah, Del, "Việt – Giarai Dictionary", Social Sciences Publisher, Hanoi, 1977.

[25] Sagart L. "The Higher Phylogeny of Austronesian and the Position of Tai - Kadai", Oceanic Linguistics, Vol. 43, No.2.

[26] Smith, K.D., "Sedang Vocabulary, Sedang - Viet - English", SIL, Saigon 1967.

[27] Vy Thị Bé, Saul, J.E., Wilson, N.F., "Nung Phan Stihng English Dictionary", SIL, 1982.

[28] Walter, H., L'aventure des langues en Occident, Robert Laffont, S.A., Paris, 1994, pp. 90-93.

[29] Watson, Richard, Saundra, "Pacoh - Viet - English Glossary", SIL. 1979.