

OLD CHINESE BASIC VOCABULARY: A HISTORICAL PERSPECTIVE

In this paper I would like to discuss a rather important methodological problem: does historical linguistics possess an objective procedure of evaluating proposed hypotheses concerning genetic relationship of languages?

The procedure that I propose below is the following:

a) to prove that two (or more) languages or linguistic families are related, we must know the set of regular phonetic correspondences connecting those languages. Otherwise any discussion is futile (all proposed equations may be due to chance). This is the standard demand of comparative linguistics;

b) the languages (or linguistic families) compared should share a significant part of basic vocabulary, and the items compared should match each other according to the set of correspondences demonstrated during the step a). This is also a common demand, but it is usually much less clear than the first one. What is basic vocabulary? What part of it is significant? I dare to propose here a test that appears (at least in my experience) to work in all cases of established genetic relationship.

As a rather quick way to test the results of comparison we may take the list of 35 most stable meanings proposed by S. Y. Yakhontov. They include the following (in English alphabet order): 'blood, bone, die, dog, ear, egg, eye, fire, fish, full, give, hand, horn, I, know, louse, moon, name, new, nose, one, salt, stone, sun, tail, this, thou, tongue, tooth, two, water, what, who, wind, year'. Actually, the stability of some items in Yakhontov's list raises doubts (this concerns, e. g., the items 'one' and 'this'). We could

easily choose some other list, but this one has an advantage of being already tested on a great many linguistic families of the world. The compared items should completely match semantically (i. e., correlations like 'fire' : 'hot' or 'water' : 'flow' are not taken into account - in order to exclude discussion of the semantic plausibility of comparisons).

I maintain that in all known cases of established genetic relationship this test yields the following results:

a) closely related languages (like Slavic or Germanic) have about 30 or more related items within the 35-wordlist;

b) more distantly related languages (on the level of Indo-European) have more than 15 related items within the 35-wordlist. To establish the precise nature of relationship (in order to distinguish, e. g., the Balto-Slavic level from the Indo-European level) we have to resort to other, more precise, statistical methods;

c) if the compared languages have from 5 to 15 related items within the 35-wordlist, it means that we can suppose a still more distant relationship between them. The precise nature of relationship is difficult to establish (it may be very archaic, like Nostratic, or somewhat more close, like Uralic or Altaic; other statistical methods should be used to obtain more precise results in cases like that);

d) if the languages compared have less than 5 common items in the 35-wordlist, it means either that they are not related at all (and the existing common items must be explained by pure chance or by borrowing), or that the common words may be in fact the 'Proto-World' heritage - if one believes in monogenesis. We will not discuss the latter hypothesis here: obviously, if one proposes a theory of genetically relating two languages, this implies that they are more closely related to each other than to all other languages of mankind.

Recently L. Sagart (CA) has presented evidence relating Chinese to Austronesian languages. This is, in fact, an attempt to revive an old theory of Conrady, put forward in 1916. To all ac-

quainted with the field, posing the problem in this way looks rather strange. Chinese is now generally accepted as a member of the Sino-Tibetan family, and comparing Chinese to Proto-Austronesian looks more or less like comparing English (as a member of Indo-European) to Proto-Uralic. In fact, in an earlier paper L. Sagart claimed that Old Chinese was closer related to Austronesian than to Tibeto-Burman [Sagart 1990: 29]. In this paper he is more cautious, saying that 'whether or not the same kind of evidence can be gathered from a comparison of RAN [*Reconstructed Austronesian* - S. S.] and the Tibeto-Burman languages, which are believed by many to be genetically related to Chinese, is not known' - thus allowing for the theoretical possibility of Tibeto-Burman being united with Old Chinese and Austronesian as a third related branch.

On the other hand, in 1982 I proposed a hypothesis uniting Sino-Tibetan (including, of course, Old Chinese), North Caucasian and Yenisseian languages into a single 'Sino-Caucasian' macrofamily (see [PYR]). The theory is also not exactly new (interestingly enough, it can also be traced back to 1916, see [Donner 1916]). However, I think this was the first time when phonetic correspondences were established between the languages in question. Later S. L. Nikolayev added Na-Dene languages to Sino-Caucasian (thus following the old ideas of E. Sapir), see [Nikolayev 1991], which is why the family is often called 'Dene-Caucasian'. Some other languages were also proposed to be included (Sumerian, Basque, Burushaski); since all of them are linguistic isolates without a possibility of intermediate reconstructions, I prefer not to resort to their evidence.

We have, therefore, two alternative theories, both corroborated by a set of phonetic correspondences, the first condition of genetic relationship thus being met in both cases. Let us now turn to Old Chinese basic vocabulary.

Below I will examine in detail every item of the 35-wordlist for Old Chinese with a discussion of proposed etymologies. I use my own reconstruction of Old Chinese (proposed in [Старостин 1989]). It is based mainly on the previous recon-

structions of S. Y. Yakhontov and E. G. Pulleyblank ([ЯХОНТОВ 1959, 1960a, 1960b, 1977a, 1977b], [Pulleyblank 1962]), but has the following distinctive features:

a) in the traditional 3d division (等) I reconstruct short vowels for OC, while long vowels are reconstructed in the 1st and 2d divisions;

b) besides plain *l-, a set of lateral initials is reconstructed (*ʎ, *ʎh, *ʎ̥, *ʎ̥h), based on a subdivision of the lateral *xiesheng* series into 2 types ([a]: MC d, th, ɖ, ɗ, ʃ, j - for OC *ʎ- and the like in different contexts; [b]: MC d, th, ʃ, ʒ, s, z, j - for OC *l-, *sl- in different contexts);

c) besides plain hissing consonants, a set of palatalized (or hushing) affricates is reconstructed (*č, *čh, *č̥, *č̥h), based on a subdivision of the hissing *xiesheng* series into 2 types ([a]: MC c, ch, ʒ, s, ʧ, s, ʃ - for OC hissing consonants; [b]: MC c, ch, ʒ, ʧ, ʧ̥, ʒ̥ - for OC hushing consonants);

d) I adopt the reconstruction (proposed by J. Norman, see [Norman 1974]) of OC voiced aspirated stops and aspirated resonants, based on the reflexes in the Min dialects (but I do not think that there is enough evidence for reconstructing another series of the so called "weakened" consonants);

e) a new final *-r is introduced (having merged with *-n in MC), on the evidence of OC rhymes and *xiesheng* series;

f) my reconstruction of main vowels in OC is almost identical to the one proposed by W. Baxter (see [Baxter 1992]), although we obtained this result independently.

The Sino-Tibetan protoforms are given in the shape reconstructed by me jointly with I. Peiros; they differ somewhat from the Tibeto-Burman reconstruction presented by P. Benedict in [STC], but the differences are not very essential for the purpose of the present paper, so I will not dwell on them in detail.

The North Caucasian protoforms are cited according to the [NCED]; Yenissonian protoforms are given according to the system proposed in [PYR]; Austronesian forms are cited according to [VL], with the revised versions according to [PANDYMC].

For correspondences between Sino-Tibetan, Yenissonian

and North Caucasian see [PYR], [HGC] and [Старостин 1989] (as well as the other paper of mine presented for this conference).

1. **blood** - 血 MC *xwiet*, OC *swhit*. The word goes back to **swhik*, with a suffixed -k (cf. other cases like 日 **nit* < **ni-k*, see below). This is the basic ST root for 'blood': ST **s-ʔw̃j*, cf., besides Chinese, Burm. *swijh*, Lush. *thi*, Kach. *sai*³¹ etc. Without the prefix **s-* (common for several ST body part names) the root is reflected, e. g., in Tsangla *yi*, Miri *iyi*, Vayu *vi*, Kiranti **hì* (possibly also Tib. *yi* 'spirit'). See IST 52, 135, 442; STC No 222 (**s-hỹwəy*).

The ST root can be compared with PNC (PEC) **hwěʔnV* 'blood', reflected in Av. *han* 'meat', And. *hin*, Akhv. *hini* 'blood', Gunz. *hāj*, Gin. *ijo*, Lak. *ul*, Darg. *hi*, Ag. *iʔ* id. etc. (see NCED 496-497). See HGC 18.

The basic Austronesian root for 'blood' is (VL) **[dɔ]ayah*, (PANDYMC) **[dD]áR_{1a}Q_{2ə}*, without known ST parallels.

2. **bone** - 骨 MC *kot*, OC **kūt*. The root goes back to ST **kūt* 'hand bone', cf. Lush. *kut* 'hand', Vayu *got*, Kiranti **gù[t]* etc. (see IST 140, 144, 161).

Cf. further PNC (PEC) **qHwəntV* 'elbow; knee' > Tsez. *q̄lontu* 'knee', Lezg. *q̄lunt* 'elbow', Lak. *aInt* 'span' etc. (see NCED 925-926); PY **g[i]d* 'elbow, joint' > Ket. *ul-git* (pl. *ul-gerəŋ*), Yug. *ul-git* 'elbow' (*ul* 'arm'), Ass. *kenar-xat-ken* 'elbow' (*kenar* 'arm, hand'), *pul-gat-ken* 'knee' (*pul* 'leg, foot'). See PYR 208, HGC 25.

The basic ST root for 'bone' was apparently lost in OC. Cf. ST **ro(s)* > Tib. *rus*, Burm. *rəuh*, Lush. *ruʔ*, Kach. *n-rut*⁵⁵, Kiranti **rùt* (?-s), **sə-rut* 'bone', see IST 39, 411; STC No 6 **rus*. This can be compared with PEC **friçV* 'joint; cartilage' > Akhv. *raçi*, Tind. *raçal* 'joint', Ag. *jiçul* 'joint; vertebra' etc., see NCED 529-530, although there are phonetic problems (in ST we would rather expect a form like **riat* or **ret*).

Another, less widespread, ST root for 'bone' is **rāk / *rāŋ* (with a frequent alternation of **-k/*-ŋ*), reflected in OC 髀 **k-rāk*, MC *kāik* 'animals' bone', Kach. *n-raŋ* 'bone', Garo *g-reŋ* etc. (see IST 133). This root is a probable match for PNC **fireχwě* 'bone',

reflected in Ing. *teχk*, Av. *raχ:á*, Lak. *t:ark*, Tab. *jirk:i* 'bone' etc. (see NCED 528), PY **?ul-a3* 'rib' (Ket. *u'let*, pl. *u'lerəŋ*, Kott. *ulai*; **-a3* is a frequent suffix in body parts). Phonetic correspondences here are quite regular (NC **-χ-* = ST **-k*, PY **-l-*). See HGC 18.

The basic Austronesian word for 'bone' is (Dempwolff) **[t]ularŋ*, (Blust) **(tT)ularŋ*. I must admit that it has a certain similarity to ST forms in **raŋ* - but if **-ŋ* here is original, the reason for its development into OC **-k* is not clear. P. Benedict (AT 238) compares the PAN root with PT **?duuk* 'bone', which is rather dubious.

L. Sagart (CA 51) compares OC 骨 **kūt* with PAN **kukut* 'joint'. This comparison seems semantically and phonetically reasonable, and is at the very least a plausible alternative to our comparison with PEC **qHwəntV*, PY **g[i]d-*.

3. **die** - 死 MC *sij*, OC *sij?*. The root is a direct descendant of ST **sij* 'to die', cf. Tib. *sí*, Burm. *sij*, Lush. *thi*, Kach. *si*³³, Kiranti **si* etc. See IST 52, 124, 436; STC No 232 (**siy*).

The ST root can be probably connected with PNC **s̄HwV* 'to breathe; to get tired, die': cf. Kar. *s:uh-an-* 'to get tired', Lak. *s:iħ* 'breath', Abkh. *-p-ś(ə)-* 'to die', Ad. *pśə-* 'to get tired', etc., see NCED 961.

The basic Austronesian root for 'die' is (VL) **mataj* / **pataj*. P. Benedict (AT 269) compares this root with Proto-Thai **taay*, which seems quite plausible.

4. **dog** - 犬 MC *khwien*, OC **kh^win*. The root contains an archaic ST derivative suffix **-n* (either collective, cf. OC 民 **min* 'people' < ST **mi* 'person', or nominalizing, cf. Tib. *gśin* 'a dead man' < *śi* 'to die', *ńen* 'a near one, relative' < *ńe* 'near' etc.) and goes back to the basic ST root for dog, **qh^wij*: cf. Tib. *khji*, Burm. *khwijh*, Lush. *ui*, Kach. *gui*³¹ (cf. also *čă-khjon*³³ 'wild dog, wolf', possibly reflecting the same suffixation as OC), etc. (see IST 42, 408, 428, STC No 159 **kwiy*).

The ST root has a precise parallel in PNC **χHwěje* 'dog': cf. Batsb. *pħu*, Av. *hoj*, And. *χ^woj*, Tsez. *βI^waj*, Darg. *χIa*, Tab. *χuj*, Ad.

ha etc. (see NCED 1073-1074). See HGC 18.

The resemblance of the OC form to PIE **k̑wōn-* 'dog' is thus either purely fortuitous, or else reflects an extremely archaic relationship (if the Sino-Caucasian form **χHwěje* itself goes back to earlier **χHwějne*, it could be compared with Nostratic **küjnä* 'wolf, dog').

The basic word for 'dog' in Austronesian is (VL) **'at'u'*, (PANDYMC) **Wasu?*. In AT 272 the root is compared with PT **hma* - from a supposed proto-root **[wa]k̑əwma* - which is of course pure fantasy.

5. **ear** - 耳 MC *ńí*, OC *nhə?* (for initial **nh-* cf. Min forms: Xiamen *hi*⁶, Chaozhou *hĩ*⁴, Fuzhou *ŋe*⁶, Jianou *neiŋ*⁶, *neiŋ*⁸). The root directly reflects ST **nǎ* 'ear' > Tib. *rna*, Burm. *nah* id., Kach. *na*³³ 'ear, to hear' etc. (see IST 38, 23, 430, STC No 453 **g-na ~ r-na*).

Further cf. PNC (PEC) **ŷwǎnŷV* 'ear' > Av. *ŷin*, Khvarsh. *āhĩ*, Lezg. *jab*, Ud. *im-uχ* etc. (see NCED 239-240). The compound **ŷwǎnŷV-diķV*, witnessed in most Andian languages (cf. And. *hanṭika*, Kar. *hanṭika* etc.), links the NC and ST forms with PY **ʔən-gde* > **ʔəgde* id., see HGC 18-19. As in most other cases, here in ST we see a loss of the first syllable with an initial laryngeal.

The Austronesian root for 'ear' is (VL) **taliŋa'*, (PANDYMC) **Calíŋa?*. It has neither ST nor PT parallels. (P. Benedict (AT 277) attempts to compare the root with scattered PT material, but without much success).

6. **egg** - 卵 MC *lwǎn*, OC **rhōn?* (for **rh-* cf. Min forms: Xiamen *nŋ*⁶, Chaozhou *nəŋ*⁴, Fuzhou *lauŋ*⁶, Jianou *səŋ*⁶); the word is attested only since Late Zhou, but seems to be the original word for 'egg' in Chinese. A MC reading *lwá* is also attested - which theoretically makes a reconstruction like OC **rhōr?* possible (although we do not know of any other cases of initial and terminal **r* combined within one syllable).

The word has no apparent ST etymology; NC and Yenis-eian connections are also unknown. On the other hand, the basic Austronesian word for 'egg' is reconstructed by Dempwolff as

*[t]əluɣ, by Dyen - as *CelúR. This, as well as the exceptional phonology of the OC form, makes a loan from Austronesian rather plausible (see AT 86).

Tibeto-Burman languages have two main roots for 'egg': *t[u]j, reflected in Lush. *tui*, Kach. *dí*³¹, Kiranti *tì 'egg' etc., see IST 48, 442; STC 45, 135, 408; and *ɣo(w) (~ *ɣu), reflected in Tib. *s-go-ŋa*, Burm. *u?* 'egg'. Neither of them is reflected in Chinese; neither has any reliable external parallels.

7. **eye** - 目 MC *mük*, OC **mhuk* (Min dialects have only "literary" readings reflecting *m-, but cf. Meixian *muk*⁷, with the tone pointing to an "aspirated" *mh-). The word goes directly back to the basic ST root for "eye", ST **mjV̄k* (with a not quite clear vowel), cf. Tib. *mig*, Burm. *mjak*, Lush. *mit* (< **mik*), Kach. *mji*³¹, Kiranti **mik* 'eye' etc. (see IST 39, 124, STC No 402 **myak*).

In PST the meaning 'eye' is probably secondary < 'eye-sight' < 'eye witness' (cf. a probable archaism in Tib. (*d*)*mjug* 'to show'), cf. PNC (PEC) **wimqV* 'eye witness' > Chech. *baq* 'true', Ing. *boqo* 'truth', Av. (**muq* >) *nuŋ* 'eye witness', Darg. *biq-ri* id. etc. (see NCED 1050).

The basic Austronesian word for 'eye' is (VL) **mata'*, (PANDYMC) **mata*, having nothing in common with the ST root. P. Benedict (AT 283) compares the root with Thai **ta* ~ **thra* < **p(h)ra*, which seems quite possible.

8. **fire** - 火 MC *xwâ*, OC **s-m[ē]j?*. This is certainly the reflex of the basic ST word for 'fire', ST **mēj* > Tib. *me*, Burm. *mih*, Lush. *mei*, Kach. *mji*³¹ (in compounds), Kiranti **mì* 'fire' etc. (see STC 38, 124, 139; STC No 290 **mey*). The word, however, has no secure external parallels.

The basic Austronesian word for 'fire' is (VL) **'apuj*, (PANDYMC) **x₁apúye*. Its Para-Thai counterpart is **prway*, see AT 290.

9. **fish** - 魚 MC *ŋö*, OC **ŋha* (for **ŋh*- cf. Xiamen *hi*², Chaozhou *hi*²). This is the basic PST 'fish' **ŋ(j)ǎ*, cf. Tib. *ña*,

Burm. *ŋah*, Lush. *ŋha*, Kach. *ŋa⁵⁵*, Kiranti **ŋə* etc. (see IST 36, 123, 407, 429; STC No 189 **ŋya*).

Further connections of the word are somewhat problematic. It certainly has no parallels in Austronesian languages (where the basic root for 'fish' is (VL) **'ikan*, (PANDYMC) **ʔiSkan*).

One of the few (since the mountain dwellers are not fishermen nowadays) common PNC (PEC) words for 'fish' is **χwanhV*, reflected in And. *χ^wami*, Tind. *χ^wā*, Lak. *hawa*, Lezg. *ved* etc. (see NCED 1078). We may suspect that PST **ŋ(j)ǎ* (or **ŋh(j)ǎ*, cf. the aspiration in Kuki-Chin) is actually a contraction < **χVŋ(j)a*, corresponding to the NC form. Since, however, we do not know of other cases of such a contraction (in a word with initial post-velar and medial nasal; regularly a ST form like **qh^wǎ* would be expected), this solution is only tentative.

10. **full** - 盈 MC *jeŋ*, OC **leŋ*. With loss of initial *P- the root goes back to the standard ST word for 'full', ST **Plǎiŋ* (cf. Burm. *praŋʔ*, Kach. *phriŋ⁵⁵*, Kiranti **bhliŋ* etc., see IST 410, STC No 142 **bliŋ* ~ **pliŋ*).

The root has no apparent Caucasian or Yenisseian parallels. Proto-Sino-Tibetan possessed a suffix **-ŋ*, appearing very frequently in adjectival roots (cf. cases like Burm. *wa* 'yellow', *wanŋh* 'bright yellow' = OC 黃 **gh^wāŋ* 'yellow' etc.). If final **-ŋ* in ST **Plǎiŋ* is a historical suffix, it is very tempting to compare ST **Plǎ(i)-* 'full' with PIE **pelə-* (**pelH-*) id. This may be a very archaic root - common in Nostratic and Sino-Caucasian, but within the latter family preserved only by Sino-Tibetan. Another case like this see below ("name").

The basic Austronesian word for 'full' is (VL) **pənuh*, (PANDYMC) **pənúq*. It is actually distantly similar to IE **pelə-*, ST **Plǎi-* - but only if we admit a development **-l- > -n-* in Austronesian. Since we do not know the system of correspondences between Proto-Austronesian and Nostratic (or Sino-Tibetan), the question remains open.

11. **give** - OC has two common words for 'give':

a) 予 MC *jö*, OC **laʔ*. This root also means 'to be with, together' (written as 與) and corresponds to PST **lã* 'to give; to take' (cf. Lush. *la* 'to take', Kach. *la*⁵⁵ id., Newari *lã* etc., see IST 162, 434).

This word is an exact match for PEC **-iLV* 'to give' (Chech. *al-*, Av. *ʕ:e-*, Lak. *-ulu-*, Tsakh. *h-ile-* etc., see NCED 640-641). OC, therefore, seems to have preserved the archaic meaning (changed to 'be given' > 'take' in other ST languages). See HGC 20.

b) 𠬞 MC *pji*, OC **pij-s*. This is a reflex of the standard ST root meaning 'to give', **piaj* - cf. Tib. *sbji-n*, Burm. *pjih*, Lush. *pe*, Kiranti **bi* 'to give' etc., see IST 49, STC No 427.

No Caucasian or Yenisseian parallels for the root have been found; however, the Proto-Austronesian for 'to give' is (VL) **bajaj*, (PANDYMC) **bəRay*, and it is possible in this case to think of an early loan from Austronesian (if we can assume a development **-R- > -0-* in Sino-Tibetan), or even of common heritage. Note that the PAN word has a Thai parallel **hai* 'to give' (see AT 300).

12. **hand** - 手 MC *śów*, OC *ʕhuʔ* (for aspirated **ʕh-* cf. Min forms: Xiamen, Chaozhou, Fuzhou *chiu*³).

The etymology of this word is difficult to establish. Cf. perhaps Tib. *śog* 'wing, wing-feather' - although there are some problems (one would rather expect *lć-* or *lǰ-*, not plain *ś-* in Tib.).

For PST one may reconstruct two roots for "arm, hand / wing": **ʕuʔ* (OC 手, Tib. *śog*) and **lǎk* (OC 翼 **lak*, MC *jik* 'wing', Tib. *lag* 'hand, arm', Burm. *lak* etc., see IST 138, 409, 435, STC No 86 **g-lak*).

In PNC we also have two rather similar roots that could possibly influence each other: **HtuLV̄* 'sleeve; wing' (Lezg. *luw* 'wing', Darg. *dulga* 'sleeve' etc., see NCED 589) and **HluʕĚ* / **ʕulHĚ* 'arm' (Lezg. *ʕul* 'shoulder-blade', Av. *ruʕ:* 'arm', Darg. *duluy* 'elbow', Kab. *bLa* 'arm' etc., see NCED 588).

Phonetically the correspondence is precise: PNC **HtuLV̄* = PST **lǎk*; PNC **ʕulHĚ* = PST **ʕuʔ*. If these equations are correct, we have to admit that Chinese has here preserved the original

semantic distinction, while Tibeto-Burman languages have "reverted" 'wing' and 'arm, hand'. This "reversal", however, is only superficial: the latter root is basically only preserved in Chinese and was lost in most other Sino-Tibetan languages, whereafter the root **lāk* 'wing' changed its meaning to 'arm, hand'.

[The situation with 'hands' and 'arms' is in general extremely complicated in Sino-Caucasian languages (as, in fact, in most linguistic families). The basic Sino-Caucasian root for 'hand' seems to have been **q̄wVlʔV*, reflected in PNC **q̄w[ā]lʔV̄* 'hand; arm' (Ub. *qa-* 'hand'. Av. *q̄:wal* 'arm', Gin. *q̄ilu* 'shoulder' etc., see NCED 933-934), PY **xire* (cf. Ket. *īl*, Ar. *kar-*). In HGC 19 I compared this root with ST **Khwār* 'fist, handful' (OC 拳 **ghwran* 'fist', Tib. *skjor*, *khjor* 'hollow of hand'). At present I think that this comparison should be rejected (because of the irregular correspondence NC ㄨ : ST -r), and ST **Khwār* should be rather compared with the standard PEC word for 'hand': PEC **kwīlʔi* (Batsb. *ko*, Av. *k^wer*, Gunz. *koro*, Lak. *k^wa*, Arch. *kul* etc., see NCED 706-707). In Eastern Caucasian we must obviously suppose a semantic shift (**q̄w[ā]lʔV̄* 'hand' > 'arm'; **kwīlʔi* 'fist' (or 'handful') > 'hand').]

Proto-Austronesian has two competing words for 'hand': (Dempwolff) **lima'*, (Dyen) **limáʔ*, and **tanjan*; the former probably corresponds to PT **mya*, see AT 309. Neither of them has any proposed cognates in ST or Chinese.

13. **horn** - 角 MC *kauk*, OC **krōk*. The word has a suffixed **-k* (cf. **swhī-k*, **ni-k* and many other similar cases) and goes back to PST **Kruā* 'horn' (sometimes with secondary loss of **K-*) > Burm. *khɹəw*, Tib. *rwa* 'horn' etc., see STC No 37 **kruw*. This is probably the original root for 'horn': another widespread root is **q^wrāŋ* (Kach. *n-run*³³, Kiranti **grāŋ* etc., see IST 406, 433, STC No 85 **rwan*), but it must have originally denoted 'horn as a vessel, drinking horn', cf. OC 觥 **k^wrāŋ* id.

PST **Kruā* corresponds to PNC **q^wīrhV* 'horn' (cf. Lak. *qi*, Ub. *qa* 'horn', Av. dial. *h^war* 'cock's comb' etc., see NCED 903), PY **qɔʔ* 'horn' (Ket. *qɔʔ*, Kott. *hau*). See HGC 20 (the PNC reconstruc-

tion has been somewhat modified since then, but the comparison is still valid).

There are two roots for 'horn' in Austronesian: **ta(n)duk* and (Dempwolff) **t'uru'*, (Lee) **Suru*. The former is compared by P. Benedict (AT 317) with PT **hmook* 'hump (of cattle)' or **no* 'horn'. Neither has any parallels in ST or Chinese.

14. I - OC has two roots for the 1st p. pronoun:

a) 我 MC *ŋā*, OC **ŋhāj?* (**ŋh-* is indicated by Min forms: Chaozhou *ua*³, Jianou *ŋuoi*⁸). This is the most common ST 1st p. pronoun (reconstructed as **ŋā* or **ŋāj*, the two forms probably reflecting original paradigmatic distinctions): cf. Tib. *ŋa*, Burm. *ŋa*, Kach. *ŋai*, Kiranti **o-ŋ*, **ŋa* etc.; see IST 36, 62, 123, 429; STC No 285 **ŋay*, No 406 **ŋa*.

The root has the closest parallel in Yenisseian, where **-ŋ* is the regular ending of the 1st p. singular object in verbs; in initial position **ŋ-*, as well as other nasal resonants, was prohibited in Proto-Yenisseian and changed to **m-* > **b-* (serving as a prefixed object marker or possessive affix of the 1st p.). Traces of this root in Caucasian can be perhaps observed in the Lak-Dargwa area (Lak. *na*, Darg. *nu* 'I' < PEC **nī*, see NCED 855).

The Yenisseian languages have preserved the distinction between suppletive forms **?a3V* 'I' (nominative stem: PY **?a3* > Ket. *āt*, Kott. *ai*, Pump. *ad*): **ŋV-* 'me, mine' (oblique stem: PY **b-/*-ŋ* > Ket. *b-*, Kott. *b-/ŋ*) (just like they have preserved the similar opposition **?aw* : **k-* in the 2nd person). North Caucasian languages have generalized the former form (as PNC **zō* 'I': Chech. *so*, Av. *du-n*, Lezg. *zu-n*, Abkh. *sa-* etc., see NCED 1084), while the Sino-Tibetan languages generalized the latter.

b) 予 MC *jō*, OC **la* 'I, we'. The semantic difference between 我 and 予 is still unclear. Both pronouns occur in the oldest texts; although S. Y. Yakhontov considers the former inclusive, and the latter exclusive, it can hardly be proved by actual examples, and a distinction like this is rather strange in a language that does not normally distinguish number.

The root has no apparent parallels in other Sino-Tibetan

languages. It can, however, be a trace of the former number distinction within the system of personal pronouns. If the original meaning of the root is 'we', it can be compared with PNC (PEC) **Lā* 'we (1st p. pl. inclusive), cf. Akhv. *iλ:i* 'we' (incl.), Tsez. *eli* 'we', Rut. *jä* 'we' (incl.) etc. (see NCED 786).

The Proto-Austronesian 1st p. pron. is (Dempwolff) **'aku'*, (Dyen) **ʔakú*, very probably corresponding to PT **kaw* 'I', see AT 203. There exists, indeed, a local Tibeto-Burman root **k(h)a* (Tib. *kho-bo* 'I, me', Lush. *ka* 'me, my', Dhimal *ka* - see IST 133), that can be compared to the Austronesian form. Its Sino-Tibetan antiquity, however, is very dubious. Moreover, **k* in Austronesian is present in all personal pronouns (see below), and is therefore most probably not a personal, but a deictic pronominal morpheme. Austronesian plural forms are also different from Sino-Tibetan: 'we' (incl.) is (Dempwolff) **kita'*, (Dyen) **kítà?*; 'we' (excl.) is (Dempwolff) **kami'*, (Dyen) **kamí?*.

15. **know** - 知 MC *ʈe*, OC **tre*.

This root goes back to ST **ria* 'to know, understand' (with a dental prefix: **T-ria* = Tib. *dra* 'experienced, learned'): Burm. *k-rah* 'to hear, get to know'; Lush. *hria* 'to know, hear, feel'; Kach. *rai* 'to be able, can'. Further etymology of the ST root is not clear.

The basic ST root for 'know' is **siə* (having changed the meaning to 'think' in OC: cf. OC 思 **sə*): Tib. *śe-s*, Burm. *si?* 'to know', Lush. *thei* 'can, be able', Kiranti **se-n* 'to learn, to see' etc. (see IST 52, STC No 182 **syey*). It corresponds (with a regular loss of the first weak syllable) to PNC **-āmçE* 'to know' (cf. Batsb. *-abç-*, And. *çin-*, Rut. *-açə-* 'to know', Darg. *umç-* 'to search', Ub. *ça-* 'to know' etc.; see NCED 262), PY **ʔVt-* (Ket. *ʔit-* etc.) id. See HGC 20.

Proto-Austronesian has two basic roots for 'know': (VL) **tahu'*, (PANDYMC) **taquh*, and (VL) **lala'*, (PANDYMC) **(ki)lala?*. Neither of them has parallels in ST and Old Chinese.

16. **louse** - 蝨 MC *ʃit*, OC **srit*.

The word goes back to the basic ST word for 'louse' - ST **śrik* (Tib. *śig*, Lush. *hrik*, Kach. *ci?*⁵⁵ etc., see IST 436, STC No 439

*ś-rik). ST *śrik itself is probably a suffixed form (on the suffix *-k see above), derived from *śar 'louse' (reflected in Burm. *sanh*, Kham *śâr*, Rgyarung *sar* and Kiranti *sèr).

ST *śar regularly corresponds to PNC (PEC) *šārī 'worm', reflected in Chech. *šera* 'earwig', Lak. *š:ira* 'ascarid', Darg. *širi*, Lezg. *šar* 'worm' etc. It is interesting to note that some EC languages also reflect a form *šār(V)-kV (cf. God. *š:irka*, Dag. *šulerk* 'helminth').

The basic Austronesian root for 'louse' is (Dempwolff) *kutu', (Dyen) *kúCuʔe, having no Chinese or ST parallels. The root is compared by P. Benedict (AT 333) with PT *thraw 'louse' < *[q](a)tru.

17. **moon** - 月 MC *ɲwət*, OC **ɲot*.

The root is absolutely isolated among ST languages. However, it seems to have quite reliable external parallels. Since ST *-t is a regular reflex of dental affricates, OC **ɲot* can be compared with the basic word for 'moon' in PNC (**wəmçö*, cf. Chech. *butt*, Av. *moç*., Tsez. *buci*, Darg. *baz*, Abkh. *á-mza* etc., see NCED 1044-1045) and PY (**?Vsuj*, cf. Ar. *ešuj*, *išuj*, Kott. *šui*, Pump. *tuj*), with a probable reconstruction **ɲəwçö* or **wəɲçö* for SC.

Tibeto-Burman languages normally reflect another root, ST **slā* (lost in Chinese): cf. Tib. *zla-ba*, Burm. *laʔ*, Lush. *thla*, Kach. *šəta*³³, Kiranti **lá* etc., see IST 57, 124, 138, 435, STC No 144 **s-gla*). It also has a reliable NC parallel: PNC **śVI?V* (~-t-, -fi-) 'light (in particular, moonlight), ray' (cf. Chech. *sa*, Darg. *šala* 'light', Kab. -p-sə- 'to shine', Urart. *šēl-ardə* 'moon, moon deity' etc., see NCED 974).

Proto-Austronesian has **bulan* 'moon', with a quite plausible parallel in PT **?blian* id., see AT 423, and without ST cognates.

18. **name** - 名 MC *mjeŋ*, OC **mheŋ* (for *mh- cf. Shaowu *mian*⁷, see Norman 1974). This is a direct descendant of PST **māiŋ* 'name', cf. Tib. *miŋ*, *mjiŋ*, Burm. *maí*, Lush. *hmiŋ*, Kach. *mjiŋ*³³, Kiranti **mìŋ* / **nìŋ* etc. (see IST 38, 124, 134, 407, 430; STC No 83

**r-min*).

The word has no known parallels in Caucasian or Yeniseian languages; however, just as 盈 **leŋ* 'full' (ST **Pläiŋ*), it seems to have possible cognates in Nostratic (IE **enom-*, Uralic **nimi*, Altaic **nüma* etc.), and thus may be a remnant of a very archaic root.

Proto-Austronesian has (VL) **[ʼ]agʼan*, (PANDYMC) **ŋájaL*, without ST parallels. A possible Thai (Kam-Sui) parallel is **?daan* 'name', see AT 343].

19. **new** - 新 MC *sjin*, OC **sin*. The root directly corresponds to Burm. *sać* 'new' (PLB **sikH* < **siŋH*; cf. other cases like Burm. *sać* 'tree' < ST **sŋ* etc.), Tib. *g-sin* 'good, fine', Limbu *ku-sŋ* 'new' < PST **sŋ* (/ **sŋ*) 'new'.

The most widespread Tibeto-Burman root for 'new' is **chär* (Tib. *gsar-ma* 'new', Burm. *chanh* 'new (of moon)', Lush. *thar* 'new' etc., see IST 52, 443, STC 147). This root, however, also means 'fresh', and its Chinese reflex is OC 鮮 **shar* 'fresh' (MC *sjen*; for **sh-* cf. Min forms: Xiamen, Chaozhou *chĩ*). Chinese obviously preserves the more archaic situation, and we may safely reconstruct for Sino-Tibetan **sŋ* (/ **sŋ*) 'new' opposed to **chär* 'fresh'.

PST **sŋ* (/ **sŋ*) 'new' has a very good match in PNC **čän?V* 'new' (Chech. *čina*, Av. *çija-*, Tind. *çihu-*, Rut. *çin-* etc., see NCED 357-358). Also related is Yug. *tul-im* 'new' (< PY **tur-*; since, however, in PY both **-r-* and the combination **-nH-* yield *-r-*, the word may correspond either to PST **sŋ* (/ *-n*) or to PST **chär*). See PYR 216-217, HGC 21.

The basic Proto-Austronesian root for 'new' is (VL) **bayu*, (PANDYMC) **baqeR_{1U}*, without ST parallels. P. Benedict compares PT **hmai* 'new' < **q/m-báo*.

20. **nose** - 鼻 MC *bji*, OC **bhij-s* (for **bh-* cf. Xiamen, Jianou *phi*⁶, Chaozhou *phi*⁶, Fuzhou *phe*⁵).

The word so far has no reliable ST or any other etymology. The common Tibeto-Burman root for 'nose' is **s-na* (Tib. *sna*,

Burm. *hna*, Kiranti **nə* etc., see IST 38, 123, 134, 430, STC No 35 **s-na*), with no Chinese parallels and no known Caucasian or Yenisseian cognates.

The basic Proto-Austronesian word for 'nose' is (VL) **'ig'un*, (PANDYMC) **qijún*, having no similar words in ST.

21. **one** - — MC *?jit*, OC **?it*. This is one of several common Sino-Tibetan roots for 'one', reflected also in Burm. *ac* 'a unit, one', Kanauri *id* 'one' (see STC 94) < PST **?it*.

Since PST **-t* regularly corresponds to NC affricates, it seems quite justified to compare the root with PNC **cHǎ* (< **Hǎcǎ*) 'one' (cf. Chech. *ča?*, Av. *co*, Khvarsh. *has*, Lak. *ca*, Lezg. *sa*, Ub. *za* etc., see NCED 323-324) and PY **χu-sa* 'one' (Ket. *qūś*, Kott. *hūčā*, Ar. *qusej*, Pump. *xuta*); see HGC 21.

We should note, however, that in this case the Proto-Austronesian form is also quite close: (VL) **'ət'a'*, **'it'a'*, (PANDYMC) **?esá?*, **?isá?* 'one' (within Para-Thai similar forms can be found in Laqua *tiə*, Kelao *si*, *tsi* - see AT 211). This may be, therefore, a case of "Urverwandtschaft" on a very deep chronological level.

22. **salt** - OC has several words for 'salt', and it is rather hard to establish which is the basic one:

a) 鹵 MC *ló*, OC **rhā?* (for **rh-* cf. Min forms: Xiamen *lo⁶*, Fuzhou *lo⁶*). The word has no ST, NC or Yenisseian parallels, and L. Sagart (CA 22) may be right in comparing it with PAN (Dempwolff) **t'ila[h]*, (Dyen) **siraq* 'condiment, salt' (P. Benedict in AT 369 connects this root with PT **klia*).

b) 鹽 MC *jem*, OC **lam*. The closest parallel is Burm. *jamh* 'powder, salt-petre' allowing to reconstruct PST **jām*. A prefixed form is **r-jam*, reflected in Tib. *rgjam-chwa* 'crystal-like salt', Kiranti **rùm* 'salt', and, possibly, with further prefixation (**K-r-jam*), in OC 鹹 **grām* (MC *γām*; for unaspirated **g-* cf. Min forms: Xiamen *kiam²*, Fuzhou *ker²* etc.) 'salt, salty'. See IST 436, STC No 245 (**g-ryum*).

Further etymology of the root is not clear (P. Benedict in

AT 107 adduces some scattered AN forms: Old Javanese *garem*, Proto-Thai **xrom* 'bitter': a loan from Tibeto-Burman is most probable here).

c) 鹺 MC *zâ*, OC **₃(h)āj*. The root also has a quite reliable Sino-Tibetan etymology: cf. Tib. *chwa* 'salt', Burm. *čhah* id., Kanauri *cha* etc. < PST **Cāj* (/ **Cuāj*), see IST 124, 504, STC No 214 (**tsa*).

PST **C(u)āj* can be directly compared with the basic NC root for 'salt', **čwěnhV* (reflected in Av. *č:am*, Akhv. *č:ani*, Kar. *č:āji*, Gunz. *cō*, Lak. *č^wu* 'salt', Abkh. *a-čá* 'salty' etc., see NCED 71-372). For the development of **-nh-* > -0 (-j) see the other paper of mine presented for this conference. In PY cf. perhaps **sin-* in Kott. *śin-čēt* 'salt'.

The basic Proto-Austronesian root for 'salt' is (VL) **'at'in*, (PANDYMC) **'asiLə*, without obvious Chinese or ST parallels.

23. **stone** - 石 MC *žek*, OC **diak* (for unaspirated **d-* cf. Min forms: Xiamen *cio?*⁸, Chaozhou *cie?*⁸ etc.). If the final **-k* is an original suffix (see above), the root can be compared with Tib. *r-do* 'stone' < PST **d(i)ă* (~ **t-*). The root has no further etymology and is a Chinese-Tibetan isogloss.

The basic ST root for 'stone' is **λəŋ* / **λək*, reflected in Burm. *kjauk* (PLB **k-luŋH* / **k-lukH*), Lush. *luŋ*, Kach. *n-luŋ*³¹, Kiranti **lùŋ* 'stone' etc. (see IST 24, 69, 434, STC No 88 **r-luŋ*). The probable Chinese reflex of this root is 礪 MC *dāŋ*, OC **L(h)āŋ-s* 'veined stone'. PST **λəŋ* has a plausible NC parallel: PEC **λānχwV* 'cobble-stone', cf. Gin. *λiχ^win* 'cobble-stone', Akhv. *λaxa* (dial. *tanχa*) 'ruins' etc., see NCED 774. L. Sagart compares OC 礪 with Proto-Austronesian **balaŋ* 'striped', but the word means actually 'spotted' and should be probably kept apart.

The basic PAN root for 'stone' is (VL) **batu'*, (PANDYMC) **batú?ə*; connected by P. Benedict, AT 398, with PT **pat* 'gem', but having no known ST parallels.

24. **sun** - 日 MC *ńit*, OC **nit*. The word contains a suffixed **-k* (OC **nit* < **ni-k*) and goes back to PST **nij* 'sun, day', cf. Tib. *ńi*,

Burm. *nij*, Lush. *ni*, Kach. *šə-ni*⁵⁵ etc. See IST 37, 429, STC No 81 (**niy*). The word has no known parallels outside Sino-Tibetan.

Proto-Austronesian has two words for 'sun': (VL) **a(n)daʷ*, **ha(ŋ)g'av*, (PANDYMC) **qajaw*, **qaLjaw*; (VL) **vayi'*, (PANDYMC) **waRih*. The former is connected by P. Benedict (AT 402) with PT **?daaw* 'star' or **[t]hraaw* 'open sky'; the latter - with PT **wan* 'day', see AT 266.

25. **tail** - 尾 MC *mwíj*, OC **məj?*. The word goes back directly to the basic PST word for 'tail', **mǎj*: cf. Burm. *mrih* (< **r-mi-h*), Lush. *mei*, Kach. *mai*³¹, Kiranti **mé* etc. (see IST 38, 408, 430; STC No 282 **r-may*).

ST **mǎj* regularly corresponds to the basic PNC (PEC) root for 'tail': PEC **mēBǎ* (cf. Batsb. *muв*, Av. *maв*, Tsez. *mahi*, Lak. *maв* etc., see NCED 801). PY has no initial **m-* and has changed it to **p-* in nominal roots (but to **w-* in auxiliary monosyllables); therefore, PY **puG-aʒ* 'tail' can also be safely compared (-*aʒ* is a frequent suffix in body part names). Cf. Ket. *hu:t* (pl. *hūrəŋ*), Yug. *fu:t* (pl. *fūdiŋ*), Kott. *fugai* / *phugai*, Ar. *phugaj* 'tail'. See HGC 22.

The Proto-Austronesian word for 'tail' is (VL) **'ikuy*, (PANDYMC) **waikuR₁*, having no ST or Chinese parallels.

26. **this**. OC had a whole set of demonstrative pronouns, whose meanings can be more or less precisely expressed as 'this'. They are:

a) 此 MC *chjé*, OC **chéj?* = Tib. *ce-s* 'so, thus' (PST **zhěj*). Further etymology of the root is unclear.

b) 斯 MC *sje*, OC **se*. The ST root is probably **šě* or **šǎ*, occurring with various suffixes: cf. Kach. *ši*³³ 'he, she, it', *ši-ŋ*³¹ 'thus, in this way', Burm. *sañ* 'this, that', *səw?* 'this', *su* 'he, she', Lush. *so* 'that', Tib. *so-s* 'the other'.

The root can be probably compared with PEC **žwV*, reflected in Tsez. *žo* 'this', Khin. *sā* 'that', Darg. *i-š* 'this' - although this root is not widespread and can be related to a nominal stem meaning 'thing' (cf. Av. *žo*, Lak. *za*), see NCED 1087-1088.

c) 是 MC *žé*, OC **de?* 'this is', 時 MC *ží*, OC **də* 'this' (cf.

also 之 MC *čí*, OC **tə* 'him, her, it', functioning as an object). The ST root is **tě* or **tǎ*, also occurring with different suffixes: cf. Tib. *do* 'this', *de* 'that', Burm. *thəw* 'this', Lush. *ti?* 'that', Kach. *dai*³³ 'this, that' etc. (see IST 133, STC No 21 **day*).

The root can be directly compared with PNC **dV-* 'that, this': Chech. *dɣā*, Av. *do-* 'that', Tsez. *je-da* 'this', Tab. *du-* 'that', Abkh. *-da* 'a deictic stem (that, there)' etc., see NCED 404-405. The Yenisseian parallel is PY **tu-* 'this' (Ket. *tu-*, Yug. *tu-* 'this', Ar. *ita-ŋ* 'they'). See HGC 22. We must note that while the ST and PY root basically denote 'this' (near deixis), PNC **dV-* was probably a far deixis pronoun ('that').

d) 伊 MC *?ji*, OC **?ij* < PST **?i*, cf. Burm. *?i*, Lush. *i* 'this'.

The pronoun **?i* is well known in NC languages: Chech. *i-*, Tsez. *je-*, Darg. *i-*, Lezg. *i*, Ub. *jə-* 'this' etc. (see NCED 214-215).

We should make a general note that most deictic morphemes are quite universal, and it is generally rather easy to find probable cognates. Thus, in Indo-European we may find **so-* 'that' (cf. PST **šě*, PNC **zəwV*), **to-* 'this, that' (cf. PST **tě* / **tǎ*, PNC **dV*, PY **tu-*), **e-/i-* 'this, a deictic stem' (cf. PST **?i*, PNC **?i*). Similar morphemes can be found in a great many linguistic families, so this material must be used with caution.

Similarly, in Proto-Austronesian we find the following two basic pronouns meaning 'this':

a) (VL) **'ini'*, (PANDYMC) **?ini?* (compared by P. Benedict in AT 408 with PT **ni/nay* 'this');

b) (VL) **'i[t]u'*, (PANDYMC) **?itu(h)* (compared in AT 406 with PT **(n)tu* 'they, that'). It is certainly possible to compare **?i-*, occurring in both of these pronouns, with PST **?i-* and **-tu(h)* - with PST **tě* / **tǎ*. This must be, indeed, a common heritage reflecting a very archaic relationship (see below). [L. Sagart (CA 45) cites PAN **idi* 'that, there' for OC 時 **də*, and PAN **ati* 'there (distant)' for OC 之 **tə*. What we said above, however, is true also for these pronominal stems.]

27. **thou** - 爾 MC *né*, OC **nhej?* (for **nh-* cf. the archaic form in Jianou, *ni*⁸). OC also has 汝 **nha?* (MC *ńö*) 'thou' and some

other forms in *nh- probably reflecting original paradigmatic distinctions.

The root is no doubt related to the basic ST 2d person pronoun **naŋ* (cf. also 戎 **nhuŋ* 'you, your' in OC). A direct descendant of the form with *-ŋ is perhaps OC 乃 **nhā?* < **nhōŋ?* 'your' (the final *-ōŋ? is not allowed in OC, while *xiesheng* points to *-ŋ in the series). Cf. Burm. *naŋ*, Lush. *naŋ*, Kach. *naŋ*³³ 'thou, you' etc. (see IST 38, 118, 410, 430; STC No 407 **naŋ*).

The PST 2d person pronoun **nǎ* / **nǎŋ* does not have reliable external parallels (unless, of course, we start comparing it with demonstrative pronouns - which, in my opinion, is absolutely inadmissible; therefore, L. Sagart's (CA 23) comparison of OC **nha?* with PAN **ina(q)* 'that, there' can not be taken seriously).

The archaic paradigm of the suppletive 2d person pronoun must have been **uV* (direct stem, nominative) : **K(w)V* (oblique stem), cf. PY **?aw* (direct stem) : **ku-* (possessive stem, marker of the 2d p. object). In PNC we have **uō* : **kwV* with a similar distribution. Some Tibeto-Burman languages, indeed, have preserved **K(w)V-* in the 2d person: cf. Tib. *khji-d*, *khjo-d* 'thou, you', Burm. *kwaj* 'you', *kha-ŋ* id., Gurung *kjàŋ* 'thou'. But the original **uV-* vanished completely, being replaced by **nǎ* of unclear origin.

Proto-Austronesian has (VL) **kav*, (Capell) **kaw* 'thou'; however similar, it certainly can not be compared with Tibeto-Burman (and North Caucasian / Yenisseeian) **K^wV-* - because, as we said above, initial **k-* here is a general marker of all personal pronouns. The second part, **-aw*, may be perhaps compared with Sino-Caucasian (not Sino-Tibetan!) **uV*, but the matter still needs further investigation.

28. **tongue** - 舌 MC *žet*, OC **lat*. The closest parallel is the Kach. form *šij-let*³¹ 'tongue', pointing to PST *(*s*)-*lāt* / *(*s*)-*lēt*. The Chinese and Kachin form can hardly be taken away from the rest of Sino-Tibetan forms, pointing to **s-lǎj* / **m-lǎj* : cf. Tib. *lce*, Burm. *hlja*, Lush. *lei* (PKC **m-lei*), Dimasa *salai*, Rawang *phəle*, Newari

me ~ *mye* etc., see IST 40, 121, 124, 136, 166, 434; STC No 231 (**m-lay* ~ **s-lay*).

Since *s- is a frequent prefix in names of body parts, we must suppose that **mlāj* is the original form (later replaced by **slāj* in some dialects). This form can be quite reliably compared with the basic PNC word for 'tongue', PNC **měļčĩ* (cf. Chech. *mott*, Av. *maç*; Tsez. *mec*, Tab. *melz*, Ub. *bža* etc., see NCED 802-803).

Some questions may be raised by the correspondence *-č- : *-j (normally *-t would be expected, and, indeed, the Chinese and Kachin form possibly reflect the more regular variant **mlāt* ~ **slāt*). This was probably conditioned by the position of *-č- within the cluster *-lč-, and in addition there exists another extremely similar case: PNC **měļžV* 'place, ground' (Chech. *mott*, Tsez. *moči*, Arch. *mač* etc.) corresponding to PST **m-laj* 'earth, ground' (Burm. *mrij*, Lush. *lei*, Nung *māli* etc., see IST 24, 184, STC No 152 **mliy*).

L. Sagart (CA 36) compares OC **lat* - without taking into account the Tibeto-Burman parallels - with PAN **dilat* 'to lick'. It is not clear whether this root is related to PAN (VL) **dilah*, (PANDYMC) **dilaq* 'tongue' (the comparison of this root with Ong-Be **lek* 'to lick' is not quite persuasive); in any case, initial **ml-* in PST makes the Austronesian parallel rather improbable.

29. **tooth** - 齒 MC *čí*, OC **thə?*. So far no satisfactory etymology was proposed for this root. *Xiesheng* (phonetic 止 **tə?*) seems to point to OC **thə?*; Min dialects, however, have a quite unexpected velar initial (Xiamen, Chaozhou, Fuzhou *khi³*). Even if we assume a non-standard cluster like **khl-* here, the word still stays an etymological mystery.

The basic ST root for tooth is, no doubt, **s-H^wa* (Tib. *so*, Burm. *swah*, Lush. *ha*, Kach. *wa³³* etc., see IST 53, 124, 135, 407, 430; STC No 437 **s-wa*). It has, however, no Chinese reflex and no reliable external parallels (cf. perhaps Proto-Lezghian **sāχ^w* 'molar tooth' > Ag. *seχ^w*, Kryz. *saχ*; the PNC antiquity of the root is somewhat dubious.)

The basic Proto-Austronesian word for 'tooth' is (VL)

*[ʼ]ipən, (PANDYMC) *ɲipən, very probably corresponding to PT *van ~ *ven id., see AT 412, but having no cognates in ST or Chinese.

30. **two** - 二 MC *ni*, OC **nij*-s. The word is a very clear descendant of PST **K-nij*(s), cf. Tib. *gnis*, Burm. *hnać*, Lush. *hniʔ*, Kach. *ni*³³, Dimasa *gi-ni*, Kiranti **ni(k)* 'two' etc. (see IST 37, 135, 411, 429; STC No 4 **g-nis*).

A direct parallel in NC is **nāwši* 'two; a two-year-old animal'. The root is preserved as a numeral (with metathesis) in Nakh **šin*- (e. g. Chech. *šiʔ*, *šina*-) and Hurr. *šin*- 'two'. Other languages have only preserved the derivate 'two-year-old animal'. Cf. Chech. *šina-ra* '2 y.-old heifer' = Lak. *nuwš:a* '2 y.-old ram' = Rut. *nüšej* id. = Ad. *nəšə* 'a ram slaughtered in honor of the guest' (*'a grown-up ram, a 2 y.-old ram'). See NCED 845-846.

The PY form for 'two' is **xina* (Ket. *in*, Yug. *in*, Kott. *ina*, Ar. *kina*, Pump. *hine-an*). This is probably an exact match for PST **K-nij*(s), although the loss of *-s is unclear (we would rather expect **xinas*).

The comparison of PST **K-nij*(s), PY **xina* and PNC **nāwši* suggests that the first two forms reflect some archaic compound. Indeed, another widespread NC root for 'two' is **qHwā* which can correspond to PY **xi*- and PST **K*-. It is interesting that for numbers from one to four most Sino-Caucasian languages actually reflect two sets of numerals; their discussion, however, goes beyond the scope of this paper. Suffice it to say that numerals reveal quite obvious parallelism in these languages.

The standard PAN root for 'two' is (VL) **duva*', (PANDYMC) **DewS3aʔ*; a very similar form is Proto-Li **draw*, see AT 211. The root may be very archaic (cf., e.g., the similar-looking PIE **dwō(u)* 'two'), but has nothing in common with the PST and OC numeral 'two'.

31. **water** - 水 MC *świ*. This word is usually reconstructed with a lateral initial (cf., e. g., in DEZC 570: **hljuajʔ*; in our reconstruction it would be **slujʔ* or **λujʔ*). However, there is no *xi*-

esheng evidence corroborating this reconstruction, and Min dialects reveal here an affricate reflex: Xiamen, Chaozhou, Fuzhou *cui*³, pointing unambiguously to an unattested Early MC form **ćwi* < OC **tuj*?. MC in this case obviously reflects a late dialectal development (*t-) > *ć- > ś- (likewise in some other cases like 書 MC *śö* 'letter, book' < **ta*, cf. Xiamen *cu*¹, Chaozhou *ci*¹, Fuzhou *cü*¹; 叔 MC *śük* 'father's younger brother' < **tik*^w, cf. Xiamen *cik*⁷, Chaozhou *cek*⁷ etc. - all containing dental stops in *xiesheng* series).

OC **tuj*? 'water' is a quite regular reflex of one of the widespread ST words for 'water', ST **tŭj*, cf. Lush. *tui* 'water', Kach. *madi*³³ 'fluid, liquid', Bodo *doi*, Kanauri *ti* 'water' etc. (see IST 48, 442; STC No 55, pp. 45, 134).

A very probable parallel is PNC (PEC) *-*V̄t̄wV* 'to pour, soak' (cf. Batsb. -*opt̄*- 'to soak', Av. *te-*, Tsez. *eṭ^w-*, Lak. -*uṭi-* 'to pour' etc., see NCED 1034-1035).

In PAN two basic words for 'water' are reconstructed: (VL) **[d̄d̄]anum*, (PANDYMC) **DaNúme* (with a quite interesting parallel in PT **na(a)m* 'water', see AT 420); (VL) **vajəy*, (PANDYMC) **wahiR* (with a not quite secure parallel in Proto-Li **ya* 'river', see AT 420). Neither of them reveals any resemblance to the Chinese word. L. Sagart (CA 44) compares it with PAN (VL) **aluy*, (PANDYMC) **aluR* 'waters, waterflow': since, however, OC certainly had a dental, not a lateral, initial here, this comparison should be rejected.

32. **what** - 何 MC *γâ*, OC **ghāj* (there is also a number of other interrogative forms in **gh-*, reflecting probably original paradigmatic distinctions; for aspirated **gh-* cf. Min reflexes of MC 胡 *γo*, OC **ghā* 'how, what': Xiamen *ɣ²*, Chaozhou *ou*²).

The ST protoform is **qhā-*, with various suffixes: cf. Tib. *ga-ru* 'whether', *ga-ŋ* 'who, which', *ga-na* 'where', Lush. *ē-ŋ* 'what', Kiranti **he* 'what'.

A probable external parallel is PNC **xV* 'who, what' (Av. *λ:i-* 'who (obl. stem)', Inkh. *λu* 'who, what', Tab. *fu* 'what', Ad. *xa-t* 'who' etc., see NCED 1062-1063).

The Proto-Austronesian form is (VL) **'apa'*, (PANDYMC)

*ʔapaʔ, having of course nothing in common with the above root.

33. **who** - 誰 MC *ʒwi*, OC **duj* (for unaspirated *d- cf. Min forms: Xiamen *cui*², Fuzhou *sui*²). The root **du-* also occurs with several suffixes, reflecting original paradigmatic forms (cf. also a suffixless - but containing a *-r-infix - OC 疇 **dru* 'who').

The PST form is **tū* 'who, which' (also having some paradigmatic modifications, notably **tū-j*), cf. Tib. *du* 'how many', Lush. *tu* 'who; whose'; Kach. *gəde*³¹ 'how many', Burm. *a-ti* 'which'.

PST **tū* can be compared with PNC **ʔādV* 'interrogative pronoun', with a usual dropping of the initial syllable with laryngeal (cf. Cham. *ed* 'what', Tsez. *didi-ju* 'which', Kab. *da-r* 'which', *da-na* 'where' etc. - see NCED 244-245). In PNC, however, this is not one of the basic interrogative pronominal roots: it is rather a specialization of **ʔādV* 'thing' (cf. Kar. *hede-la* 'thing', Bezht. *hada* 'instrument', Hurr. *edə* 'thing' etc., see *ibid.*). The same semantic development occurred probably in some ST dialects, where the root **tū* has replaced the original **su* 'who' (Tib. *su*, Burm. *a-su*, Kiranti **su* etc., see IST 135). The latter corresponds to PNC **saj* 'what', PY **ʔas-* / **sV-* id. (see HGC 23).

It is interesting to note the semantic and phonetic proximity of Proto-Austronesian (VL) **[tʰ]a[j]i[lʰ]*, (PANDYMC) **[cs]a[qh0]i* 'who'; Old Chinese, however, has not preserved any traces of this archaic pronominal root.

34. **wind** - 風 MC *püŋ*, OC **prəm*. The word rhymes in *-əm in OC, but **pəm* can not be reconstructed because of distribution restrictions; there is also some *xiasheng* evidence in favour of *-r- here.

The root is difficult to etymologize. Cf. perhaps Kach. *n-puŋ*³³ 'wind, air', Midzhu *m-boŋ* id. etc. (see IST 49, 180, 427). The words are quite similar to the MC form, but do not contain any -r-, and a development *-m > -ŋ is not typical for the languages involved. The comparison thus is rather hypothetical. In any case, OC **prəm* has replaced the basic ST root for 'wind', **lij* (reflected

in Burm. *lij*, Lush. *thli* 'wind', Kach. *buŋ-li⁵⁵* 'breeze' etc., see IST 195, STC No 454 **g-liy*, and corresponding to PNC **λwāl?V* 'wind', PY **?ul-* 'whirlwind', see HGC 30).

The root for 'wind' in Proto-Austronesian is (VL) **'aŋin*, (PANDYMC) **h₂aŋin*, without any ST or OC parallels.

35. **year** - 年 MC *nien*, OC **nhin* (for **nh-* cf. archaic Min forms: Shaowu *nin⁷*, Chaozhou *hi²*). Also means 'harvest' in OC, but this meaning is secondary: cf. the basic ST root for 'year', **snin* (Burm. *hna^c* < **hnik* - for this development see above, Kach. *šəniŋ³³*, etc. - see IST 38, 410, STC No 368 **s-nin*).

The root combines the meanings 'year' and 'old, ancient' (cf. Tib. *riniŋ* - although *r-* instead of *s-* is not quite clear here). We may think, therefore, that **-ŋ* is actually an archaic adjectival suffix here (see above), and the original root was **sīn* 'year' (**sīn* + **-ŋ* > **snin*, with a usual in these cases shift of the final resonant).

The root has a very probable Sino-Caucasian etymology: cf. PNC **šwänš* 'year' (Av. *son*, Lak. *šin*, Rut. *sän*, Ub. *š^wa* etc., see NCED 975-976); PY **sin-* 'old, ancient' (Ket. *šin*, Yug. *sin*), **s[i]-Ga* 'year' (with a temporal suffix **-Ga*; cf. Ket. *šī*, Yug. *sī*, Kott. *šēga*, Ar. *šej*, Pump. *ciku*). See PYR 220-221, HGC 35.

The standard Proto-Austronesian word is (VL) **tahun*, (PANDYMC) **taqwén*, without established ST or OC parallels.

On the basis of the evidence and discussion presented above following conclusions can be made:

A. The closest relationship can be evidently established between Old Chinese and Tibeto-Burman, thus proving the commonly accepted Sino-Tibetan theory. Old Chinese has at least 24 words within the 35-wordlist exactly matching the Tibeto-Burman parallels:

- 1) blood **s-whī-t* : **s-?wīj*
- 2) die **sij* : **šij*
- 3) dog **kh^wi-n* : **qh^wij*
- 4) ear **nhə?* : **nǎ*

- 5) eye **mhūk* : **mjV̥k*
- 6) fire **s-m[ē]j?* : **mēj*
- 7) fish **ηha* : **η(j)ǎ*
- 8) full **leη* : **Plǎiη*
- 9) give **pij-s* : **piaj*
- 10) horn **krō-k* : **Kruā*
- 11) I **ηhāj?* : **ηā(j)*
- 12) louse **srit* : **śrik*
- 13) name **mheη* : **mǎiη*
- 14) new **sin* : **sīη (/ -η)*
- 15) one **ʔit* : **ʔit*
- 16) salt **ʒ(h)āj* : **C(w)āj*
- 17) sun **ni-t* : **nīj*
- 18) tail **mǎj?* : **mǎj*
- 19) this **də, de?* : **tǎ*
- 20) thou **nha?* : **na-*
- 21) tongue **lat* : **mlāj / mlāt*
- 22) two **nij-s* : **K-ni(s)*
- 23) water **tuj?* : **tūj*
- 24) year **nhīn* : **snīη*

We can perhaps also add interrogative pronouns (25: what **ghā-* : **qhā-*, 26: who **du-* : **tū-*), but their precise meaning (which is 'who' and which is 'what') in ST is not quite clear.

A number of other words of the Chinese list has also reliable parallels in Tibeto-Burman, although the semantic match is not exact (**kūt* 'bone' : **kūt* 'hand'; **la?* 'give' : **lǎ* 'take'; **ʃu?* 'hand' : Tib. *śog* 'wing'; **tre* 'know' : **T-ria* 'hear, know, feel'; **lam* 'salt' : **jam* 'powder, salt-petre'). In two cases ('stone' and 'wind') the semantic and phonetic matches are plausible, but the words seem to be not basic and widespread in Tibeto-Burman.

We must note, however, that the exact position of Chinese within Sino-Tibetan is not clear from this comparison. It may be a separate branch, equally distant from all other subgroups of Sino-Tibetan (in that case a division into 'Sinitic' and 'Tibeto-Burman' would be justified), or it may be just one of several

Sino-Tibetan branches. We believe that the second is the case, but the discussion of this point would take us beyond the scope of the present paper.

B. The next in proximity to Old Chinese (and Sino-Tibetan in general) are North Caucasian and Yenisseian languages. Old Chinese has at least 13 precise phonetic and semantic matches with PNC in the 35-word-list:

- 1) blood *s-whī-t (< *s-ʔ^wīj) : *hwěʔnV
- 2) dog *kh^wī-n (< *qh^wīj) : *χhwějV
- 3) ear *nhəʔ (< *nǎ) : *ʕwǎnʕV
- 4) give *laʔ (< *lǎ) : *-iLV
- 5) horn *krō-k (< *Kruā) : *qwīrhV
- 6) moon *ηot : *wǎmçö
- 7) new *sin (< *sīn /-η) : *çǎnʔV
- 8) one *ʔit (< *ʔit) : *Hǎcǎ (/ *cHǎ)
- 9) salt *ʒ(h)āj (< *C(w)āj) : *çwěnhV
- 10) tail *mǎjʔ (< *mǎj) : *mēɸV
- 11) tongue *lat (< *mlăt) : *mēlçǐ
- 12) two *nij-s (< *K-ni(s)) : *nǎwǰi
- 13) year *nhīn (< *snī-η) : *śwǎnǐ

We can perhaps add 14 (fish *ηha : PNC *χwanhV), although there are some phonetic doubts (see above), and the interrogative pronoun (what *ghā- : PNC *χV), although the precise meaning of the root in PNC is not quite clear. Several other basic OC words have correlates in NC, although there are semantic differences, or else the compared NC root is not the basic one (see above on *sij 'die', *mhuk 'eye', *χuʔ 'hand', *ηhājʔ and *la 'I', *srit 'louse', *deʔ/*də, *se, *ʔi 'this', *tujʔ 'water', *du- 'who').

Proto-Yenisseian is considerably younger than Proto-North-Caucasian or Proto-Sino-Tibetan (its split can be dated at the earliest by the 1st millennium B.C.), so it has less words common with OC due to word replacement. Still we may find 9 items ('ear', 'horn', 'moon', 'new', 'one', 'tail', 'this', 'two' and 'year'), and three more have Yenisseian parallels with modified

meaning (see above on 'bone', 'die', 'I').

C. Proto-Austronesian has 4 words that can be equated with their Old Chinese counterparts:

- 1) egg **rhōn?* (< **rhōr?*): **CelúR*
- 2) give **pij-s* (< **piaj*): **bāRaj* [if **-R-* > *-0-* in PST]
- 3) one **ʔit* (< **ʔit*): **ʔisá?*
- 4) this **dā* (< **tā*): **ʔitu(h)*

One could also add № 5 (full **leŋ* < **Plāi-ŋ* : PAN **pənuh*), if we suppose a phonetic shift **-l-* > *-n-* in Proto-Austronesian. Note that not a single of these items fits into the set of correspondences presented by L. Sagart and according to his theory all of them should be rejected. There are 5 cases among the 35-word-list that can be found among the evidence presented by L. Sagart - all with different meanings in Old Chinese and Austronesian: OC **kūt* 'bone' : PAN **kukut* 'joint'; OC **rhā?* 'salt' : PAN **siraq* 'condiment'; OC **dā* 'this' : PAN **idi* 'that, there'; OC **lat* 'tongue' : PAN **dilāt* 'lick'; OC **nha?* 'thou': PAN **ina(q)* 'that, there'.

Certainly evidence like this is invalid for demonstrating genetic relationship. At the most, considering the 4 possible cognates (see above) we may think of a very deep genetic connection (on the level of Nostratic or older). In fact we may show that Old Chinese has even more possible cognates (six) with Proto-Indo-European, consider:

- 1) dog OC **khwīn* (< **qh^wīj*) : PIE **k^hwōn-*
- 2) full OC **leŋ* (< **Plāi-ŋ*): PIE **pelə-*
- 3) horn OC **krō-k* (< **Kruā*) : PIE **kr-n-*
- 4) I OC **ŋhā-* (< **ŋā*) : PIE **me-* (cf. a similar development **ŋ-* > **m-/ŋ* in Yenisseian)
- 5) moon OC **ŋot* : PIE **mēns-* (the word has a certain resemblance to the supposed Sino-Caucasian **wǎŋçǝ*, see above)
- 6) name OC **mheŋ* (< **māiŋ*) : PIE **enom-(n)*

This list can be perhaps even enlarged by pronouns (OC **dā* 'this' : PIE **to-*, with more doubt OC **ghāj* 'what' : PIE **k^wo-*).

On the other hand, the Austro-Thai hypothesis put forward by P. Benedict, seems to work very well. There are at least 12 exact matches between Austronesian and Proto-Thai (or Para-Thai), with Thai regularly using the 1st syllable:

- 1) die AN **matay* - PT **taay*
- 2) eye AN **mata* - PT **ta ~ thra*
- 3) fire AN **'apuj* - PT **pway*
- 4) give AN **bəyaj* - PT **hai*
- 5) hand AN **lima'* - PT **mya*
- 6) I AN **'aku'* - PT **kaw*
- 7) louse AN **kutu'* - PT **thraw*
- 8) moon AN **bulan* - PT **?blian*
- 9) name AN **[laʒan* - PT **?daan*
- 10) this AN **'ini'* - PT **ni/*nay*
- 11) tooth AN **[ʒipən* - PT **van ~ ven*
- 12) water AN **[dɔ]anum* - PT **na(a)m*

We may add the numerals 'one' and 'two', preserved not everywhere in Thai, but probably archaic (see above); perhaps also the roots for 'horn' (**t'uŋu'* : **no*) and 'new' (**bayu* - **hmai*), although they raise phonetic problems.

In general, the relationship between Austronesian and Thai seems to be on the same level as the relationship between Sino-Tibetan and North-Caucasian - i. e., rather distant, but discoverable.

In accordance with the rules formulated above we must now reject the Chinese-Austronesian (as well as the Chinese-Indo-European) hypothesis: the relationship between Chinese and Austronesian may exist (it is well known that no genetic hypothesis can be ultimately disproved), but on an extremely deep chronological level - deeper than Sino-Caucasian and probably even deeper than Nostratic.

So what is the meaning of the 222 lexical comparisons between Chinese and Austronesian presented in Sagart's paper? I think that for the most part we deal here with loanwords - either

from early Austronesian into Old Chinese, or vice versa. I would like to finish with the words from the paper we wrote together with I. Peiros in 1984 (where more than 30 cases like this are analyzed):

"The Sino-Tibetan protolanguage during the period before its split had been in contact with one of the branches of the Austro-Thai phylum - the Austronesian protolanguage. This contact lasted also for some time after the splitting of these protolanguages, which must account for a considerable amount of Austronesian-Old Chinese parallels."