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 has presented evidence relating Chinese to Austronesian languages [CA]. This is, in fact, an attempt to revive an old theory of Conrady, put forward in 1916. To all acquainted with the field, posing the problem in this way looks rather strange. Chinese is now generally accepted as a 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 Yeniseian languages into a single "Sino-Caucasian" macrofamily [PYR]. The theory is also not exactly new (interestingly enough, it can also be traced back to 1916 [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 reconstructions 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 (**ƛ, *ƛh, *Ł, *Łh*), based on a subdivision of the lateral *xiesheng* series into 2 types ([a]: MC *d, th, ḍ, ṭh, ś, j* — for OC **ƛ-* and the like in different contexts; [b]: MC *d, th, ś, ź, s, z, j* — for OC **l-, *sl-* in different contexts);

c) besides plain hissing consonants, a set of palatalised (or hushing) affricates is reconstructed (**ć, *ćh, *ʑ, *ʑh*), based on a subdivision of the hissing *xiesheng* series into 2 types ([a]: MC *c, ch, ʒ, s, ṣh, s, ṣ* — for OC hissing consonants; [b]: MC *c, ch, ʒ, ç, ṣh, ʒ* — for OC hushing consonants);

d) I adopt the reconstruction (proposed by J. Norman [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 [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]; Yeniseian 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, Yeniseian and North Caucasian see [PYR; HGC; Старостин 1989] (as well as the other paper of mine presented for this conference).

1. **blood** — 血 MC *xwiet,* OC **swhīt.* The word goes back to **swhīk,* with a suffixed *-k* (cf. other cases like 日 **nit* < **ni-k,* see below). This is the basic ST root for 'blood': ST **s-ʔʷīj,* cf., besides Chinese, Brm. *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 № 222 (**s-hywəy*).

The ST root can be compared with PNC (PEC) **hwĕʔnV* 'blood', reflected in Av. *han* 'meat', And. *hin,* Akhv. *hini* 'blood', Gunz. *hãj,* Gin. *ijo,* Lak *uI,* Darg. *ħi,* Ag. *iʔ* id. etc. [NCED: 496–497]. See [HGC: 18].

The basic Austronesian root for 'blood' is [VL] **[dḍ]aɣah,* [PANDYMC] **[dD]áR₁aQ₂ə,* 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. [IST: 140, 144, 161].

Cf. further PNC (PEC) **q̇HwəntV* 'elbow; knee' > Tsez. *q̇Iontu* 'knee', Lezg. *q̇Iunt* 'elbow', Lak *aInt* 'span' etc. [NCED: 925–926]; PY **g[i]d* 'elbow, joint' > Ket *uĺ-gɨt* (pl. *uĺ-gerəŋ*), Yugh *uĺ-git* 'elbow' (*uĺ* 'arm'), Ass. *kenar-xat-ken* 'elbow' (*kenar* 'arm, hand'), *pul-gat-ken* 'knee' (*pul* 'leg, foot') [PYR: 208; HGC: 25].

The basic ST root for 'bone' was apparently lost in OC. Cf. ST **ro(s)* > Tib. *rus,* Brm. *rəuh,* Lush. *ruʔ,* Kach. *n-rut*[55], Kiranti **rùt* (*?-s*), **sə̀-rut* 'bone' [IST: 39, 411]; [STC] № 6 **rus.* This can be compared with PEC **ħriçV* 'joint; cartilage' > Akhv. *raçi,* Tind. *raçal* 'joint', Ag. *jiçul* 'joint; vertebra' etc.

[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 骼 **k-rāk,* MC *kậik* ‘animal's bone’, Kach. *n-raŋ* ‘bone’, Garo *g-reŋ* etc. [IST: 133]. This root is a probable match for PNC **ɦreX̱wĕ* ‘bone’, reflected in Ing. *ṭeχk,* Av. *raX̱á,* Lak *ṯarḳ,* Tab. *jirḳi* ‘bone’ etc. [NCED: 528], PY **ʔuĺ-aʒ* ‘rib’ (Ket *uĺet,* pl. *uĺerəŋ,* Kott *ulai; *-aʒ* is a frequent suffix in body parts). Phonetic correspondences here are quite regular (NC **-X̱-* = ST **-k*, PY **-Ĺ-*). See [HGC: 18].

The basic Austronesian word for ‘bone’ is (Dempwolff) **[t]ulaŋ,* (Blust) **(tT)ulaŋ*. 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 **ʔduuk* ‘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 **q̇HwəntV,* PY **g[i]d-*.

3. **die** — 死 MC *sjí,* OC **sijʔ*. The root is a direct descendant of ST **sĭj* ‘to die’, cf. Tib. *śi,* Brm. *sij,* Lush. *thi,* Kach. *si*[33], Kiranti **si* etc. [IST: 52, 124, 436]; STC № 232 (**siy*).

The ST root can be probably connected with PNC **s̱ĭHwV* ‘to breathe; to get tired, die’: cf. Kar. *s̱uh-an-* ‘to get tired’, Lak *s̱iħ* ‘breath’, Abkh. *-p-ś(ə)-* ‘to die’, Ad. *pśə-* ‘to get tired’, etc. [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 **khʷīn.* The root contains an archaic ST derivative suffix **-n* (either collective, cf. OC 民 **min* ‘people’ < ST **mi* ‘person’, or nominalising, 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, **qhʷīj:* cf. Tib. *khji,* Brm. *khwijh,* Lush. *ui,* Kach. *gui*[31] (cf. also *čă-khjon*[33] ‘wild dog, wolf’, possibly reflecting the same suffixation as OC), etc. [IST: 42, 408, 428]; STC № 159 **kwiy*.

The ST root has a precise parallel in PNC **χ̱Hwĕje* ‘dog’: cf. Btsb. *pɬu,* Av. *hoj,* And. *χʷoj,* Tsez. *ʁIʷaj,* Darg. *χIa,* Tab. *χuj,* Ad. *ɬa* etc. [NCED: 1073–1074]. See [HGC: 18].

The resemblance of the OC form to PIE **ĥ̬u̯ōn-* ‘dog’ is thus either purely fortuitous, or else reflects an extremely archaic relationship (if the Sino-Caucasian form **χ̲Hwěje* itself goes back to earlier **χ̲Hwějne,* it could be compared with Nostratic **ḳ̬ünä* ‘wolf, dog’).

The basic word for ‘dog’ in Austronesian is [VL] **’at’u’,* [PANDYMC] **Wasuʔ*. In [AT: 272] the root is compared with PT **hma* — from a supposed proto-root **[wa]kḷəwma* — which is of course pure fantasy.

5. **ear** — 耳 MC *ńɨ́,* OC **nhəʔ* (for initial **nh-* cf. Min forms: Xiamen *hi*^{6}, Chaozhou *hĩ*^{4}, Fuzhou *ŋe*^{6}, Jianou *nɛiŋ*^{6}, *nɛiŋ*^{8}). The root directly reflects ST **nă* ‘ear’ > Tib. *rna,* Brm. *nah* id., Kach. *na*^{33} ‘ear, to hear’ etc. [IST: 38, 23, 430]; STC № 453 **g-na ~ r-na*.

Further cf. PNC (PEC) **ʕwănʕV* ‘ear’ > Av. *ʕin,* Khvarsh. *ãhĩ,* Lezg. *jab,* Ud. *im-uχ* etc. [NCED: 239–240]. The compound **ʕwănʕV-diḳV*, witnessed in most Andian languages (cf. And. *hanṭiḳa,* Kar. *hanṭiḳa* etc.), links the NC and ST forms with PY **ʔɔn-gde* > **ʔɔgde* id. [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] **Calíŋaʔ*. 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 *lwə́n,* OC **rhōnʔ* (for **rh-* cf. Min forms: Xiamen *nŋ*^{6}, Chaozhou *nəŋ*^{4}, Fuzhou *lauŋ*^{6}, Jianou *sɔŋ*^{6}); the word is attested only since Late Zhou, but seems to be the original word for ‘egg’ in Chinese. A MC reading *lwə́* is also attested — which theoretically makes a reconstruction like OC **rhōrʔ* 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 Yeniseian 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 [AT: 86].

Tibeto-Burman languages have two main roots for ‘egg’: **t[u]j,* reflected in Lush. *tui,* Kach. *di*^{31}, Kiranti **ṭì* ‘egg’ etc. [IST: 48, 442; STC: 45, 135, 408]; and **ɣo(w)* (~ **ɣu*), reflected in Tib. *s-go-ŋa,* Brm. *uʔ* ‘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 **mjV̌k* (with a not quite clear vowel), cf. Tib. *mig,* Brm. *mjak,* Lush. *mit* (< **mik*), Kach. *mji?*^{31}, Kiranti **mìk* 'eye' etc. [IST: 39, 124]; STC № 402 **myak*.

In PST the meaning 'eye' is probably secondary < 'eyesight' < 'eye witness' (cf. a probable archaism in Tib. *(d)mjug* 'to show'), cf. PNC (PEC) **wĭmq̇V* 'eye witness' > Chech. *baq̇* 'true', Ing. *boq̇o* 'truth', Av. (**muq̇* >) *nuʕ* 'eye witness', Darg. *biq̇-ri* id. etc. [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 *xwáˊ*, OC **s-m[ē]jʔ*. This is certainly the reflex of the basic ST word for 'fire', ST **mēj* > Tib. *me,* Brm. *mih,* Lush. *mei,* Kach. *mji*^{31} (in compounds), Kiranti **mì* 'fire' etc. [STC: 38, 124, 139]; STC № 290 **mey*). The word, however, has no secure external parallels.

The basic Austronesian word for 'fire' is [VL] **'apuj,* [PANDYMC] **xɪapúye.* Its Para-Thai counterpart is **pway* [AT: 290].

9. **fish** — 魚 MC *ŋöˊ*, OC **ŋha* (for **ŋh-* cf. Xiamen *hi*^{2},Chaozhou *hi*^{2}). This is the basic PST 'fish' **ŋ(j)ă*, cf. Tib. *ńa,* Brm. *ŋah,* Lush. *ŋha,* Kach. *ŋa*^{55}, Kiranti **ŋə̀* etc. [IST: 36, 123, 407, 429]; STC № 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] **ʔíSkan*).

One of the few (since the mountain dwellers are not fishermen nowadays) common PNC (PEC) words for 'fish' is **χ̲wanħV,* reflected in And. *χʷami,* Tind. *χʷãˌ* Lak *ħawa,* Lezg. *ʁed* etc. [NCED: 1078]. We may suspect that PST **ŋ(j)ă* (or **ŋh(j)ă,* 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 postvelar and medial nasal; regularly a ST form like **qhʷă* 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 **Plăiŋ* (cf. Brm. *prańʔ,* Kach. *phriŋ*^{55}, Kiranti **bhlìŋ* etc.) [IST: 410]; STC № 142 **bliŋ ~ *pliŋ*).

The root has no apparent Caucasian or Yeniseian parallels. Proto-Sino-Tibetan possessed a suffix **-ŋ*, appearing very frequently in adjectival roots (cf. cases like Brm. *wa* 'yellow', *waŋh* 'bright yellow' = OC 黃

ghʷāŋ* ‘yellow’ etc.). If final **-ŋ* in ST **Plăiŋ* is a historical suffix, it is very tempting to compare ST **Plă(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 **Plăi-* — 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 *jő,* OC **laʔ.* This root also means ‘to be with, together’ (written as 與) and corresponds to PST **lă* ‘to give; to take’ (cf. Lush. *la* ‘to take’, Kach. *la*[55] id., Newari *là* etc.; [IST: 162, 434]).

This word is an exact match for PEC **-ɨ̆l̥V* ‘to give’ (Chech. *al-,* Av. *Ʌ̣e-,* Lak *-ulu-,* Tsakh. *h-ile-* etc.; [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) 畀 MC *pjì,* OC **pij-s.* This is a reflex of the standard ST root meaning ‘to give’, **piaj* — cf. Tib. *sbji-n,* Brm. *pijh,* Lush. *pe,* Kiranti **bì* ‘to give’ etc. [IST: 49]; STC № 427.

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

12. **hand** — 手 MC *śə́w,* OC **Ʌhuʔ* (for aspirated **Ʌh-* 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 *lʒ́-*, not plain *ś-* in Tib.).

For PST one may reconstruct two roots for ‘arm, hand / wing’: **Ʌuʔ* (OC 手, Tib. *śog*) and **lăk* (OC 翼 **ləkˁ,* MC *jik* ‘wing’, Tib. *lag* ‘hand, arm’, Brm. *lak* etc. [IST: 138, 409, 435]; STC № 86 **g-lak*.

In PNC we also have two rather similar roots that could possibly influence each other: **HłuŁV̄* ‘sleeve; wing’ (Lezg. *luw* ‘wing’, Darg. *dulga*

'sleeve' etc.; [NCED: 589]) and **Hluɬ̆Ĕ* / **ɬ̆ulHĔ* 'arm' (Lezg. *ḳul* 'shoulder-blade', Av. *ruɬ̆* 'arm', Darg. *duluɣ* 'elbow', Kab. *bLa* 'arm' etc. [NCED: 588]).

Phonetically the correspondence is precise: PNC **HłuŁV̄* = PST **lăk*; PNC **ɬ̆ulHĔ* = PST **ɬ̆u?*. 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 **lăk* '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 **q̇wVɬ?V*, reflected in PNC **q̇w[ā]ɬ?V̆* 'hand; arm' (Ub. *q̇a-* 'hand'. Av. *q̇ʷal* 'arm', Gin. *q̇ilu* 'shoulder' etc. [NCED: 933–934]), PY **xɨre* (cf. Ket *ɨ́l*, Ar. *kar-*). In [HGC: 19] I compared this root with ST **Khʷăr* '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 *-ɬ-* : ST *-r*), and ST **Khʷăr* should be rather compared with the standard PEC word for 'hand': PEC **kwīlʔɨ* (Btsb. *ko,* Av. *kʷer,* Gunz. *koro,* Lak *kʷa,* Arch. *kul* etc. [NCED: 706-707]). In Eastern Caucasian we must obviously suppose a semantic shift (**q̇w[ā]ɬ?V̆* 'hand' > 'arm'; **kwīlʔɨ* 'fist' (or 'handful') > 'hand').]

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

13. **horn** — 角 MC *kạuk*, 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-*) > Brm. *khrəw*, Tib. *rwa* 'horn' etc.; see STC № 37 **kruw*. This is probably the original root for 'horn': another widespread root is **qʷrāŋ* (Kach. *n-ruŋ*[33], Kiranti **grãŋ* etc.; [IST: 406, 433]; STC № 85 **rwaŋ*), but it must have originally denoted 'horn as a vessel, drinking horn', cf. OC 觥 **kʷrāŋ* id.

PST **Kruā* corresponds to PNC **qwīrhV* 'horn' (cf. Lak *qi*, Ub. *qa* 'horn', Av. dial. *hʷar* 'cock's comb' etc. [NCED: 903]), PY **qɔʔ* 'horn' (Ket *qɔʔ*, Kott *hau*). See [HGC: 20] (the PNC reconstruction 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'uŋu'*, (Lee) **Suŋu*. The former is compared by P. Benedict [AT: 317]

with PT **hnook* '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 *ŋâ,* OC **ŋhājʔ* (**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,* Brm. *ŋa,* Kach. *ŋai,* Kiranti **ʔo-ŋ, *ŋa* etc. [IST: 36, 62, 123, 429]; STC № 285 **ŋay,* № 406 **ŋa.*

The root has the closest parallel in Yeniseian, 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-Yeniseian 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 **nǐ* [NCED: 855]).

The Yeniseian languages have preserved the distinction between suppletive forms **ʔaʒV* 'I' (nominative stem: PY **ʔaʒ* > 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 **ʔaw* : **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. [NCED: 1084]), while the Sino-Tibetan languages generalized the latter.

b) 予 MC *jö,* 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) **Lā̄* 'we' (1st p. pl. inclusive), cf. Akhv. *iɬ̆i* 'we' (incl.), Tsez. *eli* 'we', Rut. *jä* 'we' (incl.) etc. [NCED: 786].

The Proto-Austronesian 1st p. pron. is (Dempwolff) **'aku'*, (Dyen) **ʔakú,* very probably corresponding to PT **kaw* 'I' [AT: 203]. There exists, indeed, a local Tibeto-Burman root **k(h)a* (Tib. *kho-bo* 'I, me', Lush. *ka* 'me, my', Dhimal *ka* [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) **kítàʔ*; ‘we’ (excl.) is (Dempwolff) **kami'*, (Dyen) **kamíʔ*.

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’): Brm. *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*, Brm. *siʔ* ‘to know’, Lush. *thei* ‘can, be able’, Kiranti **se-n* ‘to learn, to see’ etc. [IST: 52]; STC № 182 **syey*. It corresponds (with a regular loss of the first weak syllable) to PNC **-ămʂE* ‘to know’ (cf. Btsb. *-abʂ-*, And. *ʂin-*, Rut. *-aʂa-* ‘to know’, Darg. *umʂ-* ‘to search’, Ub. *ʂa-* ‘to know’ etc. [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)lalaʔ*. Neither of them has parallels in ST and Old Chinese.

16. **louse** — 蝨 MC *ʂit*, OC **srit*.

The word goes back to the basic ST word for ‘louse’ — ST **śrik* (Tib. *śig*, Lush. *hrik*, Kach. *ciʔ*[55] etc.; [IST: 436]; STC № 439 **ś-rik)*. ST **śrik* itself is probably a suffixed form (on the suffix **-k* see above), derived from **śar* ‘louse’ (reflected in Brm. *sanh*, Kham *śàr*, Rgyarung *sar* and Kiranti **sèr*).

ST **śar* regularly corresponds to PNC (PEC) **ʂ̄ārɨ* ‘worm’, reflected in Chech. *šēra* ‘earwig’, Lak *ʂ̄ira* ‘ascarid’, Darg. *širi*, Lezg. *šar* ‘worm’ etc. It is interesting to note that some EC languages also reflect a form **ʂ̄ār(V)-kV* (cf. God. *ʂ̄irʹka*, 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 *ŋwət*, OC **ŋot*.

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 **ŋot* can be compared with the basic word for ‘moon’ in PNC (**wămʂ̆ŏ*, cf. Chech. *butt*, Av. *moʂ*, Tsez. *buci*, Darg. *baz*, Abkh. *á-mza* etc. [NCED: 1044–1045]) and PY (**ʔVsuj*, cf. Ar. *ešuj*, *išuj*, Kott *šui*, Pump. *tuj*), with a probable reconstruction **ŋăwʂ̆ŏ* or **wăŋʂ̆ŏ* for SC.

Tibeto-Burman languages normally reflect another root, ST **slă* (lost in Chinese): cf. Tib. *zla-ba,* Brm. *laʔ,* Lush. *thla,* Kach. *šəta*[33], Kiranti **lá* etc. [IST: 57, 124, 138, 435]; STC № 144 **s-gla*). It also has a reliable NC parallel: PNC **śVlʔV* (~ *-ɬ-*, *-ɦ-*) 'light (in particular, moonlight), ray' (cf. Chech. *sa,* Darg. *šala* 'light', Kab. *-p-sə-* 'to shine', Urart. *šēl-ardə* 'moon, moon deity' etc. [NCED: 974]).

Proto-Austronesian has **bulan* 'moon', with a quite plausible parallel in PT **ʔblɨan* id. [AT: 423], and without ST cognates.

18. **name** — 名 MC *mjeŋ,* OC **mheŋ* (for **mh-* cf. Shaowu *miaŋ*[7] [Norman 1974]). This is a direct descendant of PST **măiŋ* 'name', cf. Tib. *miŋ*, *mjiŋ,* Brm. *mań,* Lush. *hmiŋ,* Kach. *mjiŋ*[33], Kiranti **mìŋ* / **nìŋ* etc. [IST: 38, 124, 134, 407, 430]; STC № 83 **r-miŋ*).

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

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

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

The most widespread Tibeto-Burman root for 'new' is **chăr* (Tib. *gsar-ma* 'new', Brm. *chanh* 'new (of moon)', Lush. *thar* 'new' etc.; [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*[1]). Chinese obviously preserves the more archaic situation, and we may safely reconstruct for Sino-Tibetan **sĭŋ* (/**sĭn*) 'new' opposed to **chăr* 'fresh'.

PST **sĭŋ* (/ **sĭn*) 'new' has a very good match in PNC **ṣ̌ăn*ʔ*V* 'new' (Chech. *ṣina,* Av. *ṣija-,* Tind. *ṣīhu-,* Rut. *ṣin-* etc. [NCED: 357–358]). Also related is Yugh *tul-ɨm* 'new' (< PY **tur-*; since, however, in PY both **-r-* and the combination **-nH-* yield *-r-*, the word may correspond either to PST **sĭŋ* (/ *-n*) or to PST **chăr*). See [PYR: 216–217; HGC: 21].

The basic Proto-Austronesian root for 'new' is [VL] **baɣu,* [PANDYMC] **baqeRɪu*, without ST parallels. P. Benedict compares PT **hmaɨ* 'new' < **q/m-báo.*

20. **nose** — 鼻 MC *bjì,* OC **bhij-s* (for **bh-* cf. Xiamen, Jianou *phi*[6], Chaozhou *phĩ*[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,* Brm. *hna,* Kiranti **nà* etc. [IST: 38, 123, 134, 430]; STC № 35 **s-na*), with no Chinese parallels and no known Caucasian or Yeniseian cognates.

The basic Proto-Austronesian word for 'nose' is [VL] **'ig'uŋ,* [PANDYMC] **qíjúŋ*, having no similar words in ST.

21. **one** — — MC *ʔjit,* OC **ʔit.* This is one of several common Sino-Tibetan roots for 'one', reflected also in Brm. *ać* 'a unit, one', Kanauri *id* 'one' [STC: 94] < PST **ʔĭt.*

Since PST **-t* regularly corresponds to NC affricates, it seems quite justified to compare the root with PNC **cHə̆* (< **Hə̆cə̆*) 'one' (cf. Chech. *cha?,* Av. *co,* Khvarsh. *has,* Lak *ca,* Lezg. *sa,* Ub. *za* etc. [NCED: 323–324]) and PY **χu-sa* 'one' (Ket *qūś,* Kott *hūča,* 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] **ʔesáʔ, *ʔisáʔ* 'one' (within Para-Thai similar forms can be found in Laqua *tiə,* Kelao *si, tsi* [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 **rhāʔ* (for **rh-* cf. Min forms: Xiamen *lɔ⁶,* Fuzhou *lo⁶*). The word has no ST, NC or Yeniseian parallels, and L. Sagart [CA: 22] may be right in comparing it with PAN (Dempwolff) **t'iḷa[h],* (Dyen) **siraq* 'condiment, salt' (P. Benedict in [AT: 369] connects this root with PT **kliə*).

b) 鹽 MC *jem,* OC **lam.* The closest parallel is Brm. *jamh* 'powder, salt-petre' allowing to reconstruct PST **jə̆m.* 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 *ɣɐ̈m;* for unaspirated **gr̥-* cf. Min forms: Xiamen *kiam²,* Fuzhou *keŋ²* etc.) 'salt, salty'. See [IST: 436]; STC № 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 *ʒâ,* OC **ʒ(h)āj*. The root also has a quite reliable Sino-Tibetan etymology: cf. Tib. *chwa* 'salt', Brm. *ćhah* id., Kanauri *cha* etc. < PST **Cāj* (/ **Cuāj*) [IST: 124, 504]; STC № 214 (**tsa*).

PST **C(u)āj* can be directly compared with the basic NC root for 'salt', **ʦ̱wĕnhV* (reflected in Av. *ʦ̱am,* Akhv. *ʦ̱ani,* Kar. *ʦ̱ãji*, Gunz. *cɔ̃,* Lak *ʦ̣ʷu* 'salt',

Abkh. *a-ṗá* ‘salty’ etc. [NCED: 71, 372]). For the development of **-nh-* > *-Ø* (*-j*) see the other paper of mine presented for this conference. In PY cf. perhaps **sin-* in Kott *śin-ćēt* ‘salt’.

The basic Proto-Austronesian root for ‘salt’ is [VL] **ʼatʼin,* [PANDYMC] **ʼasɪíLə,* without obvious Chinese or ST parallels.

23. **stone** — 石 MC *źek,* 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 **ƛəŋ* / **ƛək,* reflected in Brm. *kjauk* (PLB **k-luŋH* / **k-lukH*), Lush. *luŋ,* Kach. *n-luŋ*[31], Kiranti **lùŋ* ‘stone’ etc. [IST: 24, 69, 434]; STC № 88 **r-luŋ*. The probable Chinese reflex of this root is 碭 MC *dàŋ,* OC **Ł(h)āŋ-s* ‘veined stone’. PST **ƛəŋ* has a plausible NC parallel: PEC **ƛ̆ănχwV* ‘cobble-stone’, cf. Gin. *ƛ̣iχʷin* ‘cobble-stone’, Akhv. *ƛ̣aχa* (dial. *ṭanχa*) ‘ruins’ etc. [NCED: 774]. L. Sagart compares OC 碭 with Proto-Austronesian **bəlaŋ* ‘striped’, but the word means actually ‘spotted’ and should be probably kept apart.

The basic PAN root for ‘stone’ is [VL] **batuʼ,* [PANDYMC] **baṭúʔə;* 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 **nĭj* ‘sun, day’, cf. Tib. *ńi,* Brm. *nij*, Lush. *ni,* Kach. *šə-ni*[55] etc. [IST: 37, 429]; STC № 81 (**niy*). The word has no known parallels outside Sino-Tibetan.

Proto-Austronesian has two words for ‘sun’: [VL] **ʼa(n)dav, *ha(ŋ́)gʼav*, [PANDYMC] **qajaw, *qaLjaw*; [VL] **vaɣiʼ*, [PANDYMC] **waRih.* The former is connected by P. Benedict [AT: 402] with PT **ʔdaaw* ‘star’ or **[t]hraaw* ‘open sky’; the latter — with PT **wan* ‘day’ [AT: 266].

25. **tail** — 尾 MC *mwɨ́j,* OC **məjʔ*. The word goes back directly to the basic PST word for ‘tail’, **măj:* cf. Brm. *mrih* (< **r-mi-h*), Lush. *mei,* Kach. *mai*[31], Kiranti **mé* etc. [IST: 38, 408, 430]; STC № 282 **r-may*.

ST **măj* regularly corresponds to the basic PNC (PEC) root for ‘tail’: PEC **mēʁV̆* (cf. Btsb. *muʁ,* Av. *maʁ,* Tsez. *maħi,* Lak *maʁ* etc. [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 *huːt*

(pl. *hūŕəŋ*), Yugh *fu:t* (pl. *fūdiŋ*), Kott *fugai* / *phugai,* Ar. *phugaj* ‘tail’. See [HGC: 22].

The Proto-Austronesian word for ‘tail’ is [VL] **’ikuɣ,* [PANDYMC] **w4ikuR1,* 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 **ćhej*ʔ = Tib. *će-s* ‘so, thus’ (PST **ʒ́hěj*). Further etymology of the root is unclear.

b) 斯 MC *sje,* OC **se.* The ST root is probably **šĕ* or **šĭ,* occurring with various suffixes: cf. Kach. *ši*33 ‘he, she, it’, *ši-ŋ*31 ‘thus, in this way’, Brm. *sań* ‘this, that’, *səw*ʔ ‘this’, *su* ‘he, she’, Lush. *so* ‘that’, Tib. *śo-s* ‘the other’.

The root can be probably compared with PEC **źwV,* reflected in Tsez. *žo* ‘this’, Khin. *sä* ‘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*) [NCED: 1087–1088].

c) 是 MC *ʒ́é,* OC **de*ʔ ‘this is’, 時 MC *ʒ́ɨ,* OC **də* ‘this’ (cf. also 之 MC *ćɨ,* OC **tə* ‘him, her, it’, functioning as an object). The ST root is **tĕ* or **tă,* also occurring with different suffixes: cf. Tib. *do* ‘this’, *de* ‘that’, Brm. *thəw* ‘this’, Lush. *ti*ʔ ‘that’, Kach. *dai*33 ‘this, that’ etc. [IST: 133]; STC № 21 **day*.

The root can be directly compared with PNC **dV-* ‘that, this’: Chech. *dʕā,* Av. *do-* ‘that’, Tsez. *je-da* ‘this’, Tab. *du-* ‘that’, Abkh. *-da* ‘a deictic stem (that, there)’ etc. [NCED: 404–405]. The Yeniseian parallel is PY **tu-* ‘this’ (Ket *tu-,* Yugh *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 *ʔji,* OC **ʔij* < PST **ʔĭ,* cf. Brm. *ʔi,* Lush. *i* ‘this’.

The pronoun **ʔi* is well known in NC languages: Chech. *i-,* Tsez. *je-,* Darg. *i-,* Lezg. *i,* Ub. *jə-* ‘this’ etc. [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 **šĕ,* PNC **źwV*), **to-* ‘this, that’ (cf. PST **tĕ* / **tă,* PNC **dV,* PY **tu-*), **e-*/**i-* ‘this, a deictic stem’ (cf. PST **ʔĭ,* PNC **ʔi*). 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] **ʔiníʔ* (compared by P. Benedict in [AT: 408] with PT **ni*/*nay* ‘this’);

b) [VL] *'*i[t]u'*, [PANDYMC] **ʔitu(h)* (compared in [AT: 406] with PT **(n)tu* 'they, that'). It is certainly possible to compare **ʔi-*, occurring in both of these pronouns, with PST **ʔi-* and **-tu(h)* — with PST **tĕ* / **tə̆.* 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 **dəʔ,* and PAN **ati* 'there' (distant) for OC 之 **təʔ.* What we said above, however, is true also for these pronominal stems.)

27. **thou** — 爾 MC *ńé,* OC **nhejʔ* (for **nh-* cf. the archaic form in Jianou, *ni*[8]). OC also has 汝 **nhaʔ* (MC *ńǒ*) '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 **nhāʔ* < **nhāŋʔ* 'your' (the final **-āŋʔ* is not allowed in OC, while *xie-sheng* points to **-ŋ* in the series). Cf. Brm. *naŋ,* Lush. *naŋ,* Kach. *naŋ*[33] 'thou, you' etc. [IST: 38, 118, 410, 430]; STC № 407 **naŋ*).

The PST 2d person pronoun **nă* / **nă̆ŋ* 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 **nhaʔ* with PAN **ina(q)* 'that, there' can not be taken seriously).

The archaic paradigm of the suppletive 2d person pronoun must have been **u̯V* (direct stem, nominative) : **K(w)V* (oblique stem), cf. PY **ʔaw* (direct stem) : **ku-* (possessive stem, marker of the 2d p. object). In PNC we have **u̯ō* : **ʁwV* 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', Brm. *kwaj* 'you', *kha-ŋ* id., Gurung *kjàŋ* 'thou'. But the original **u̯V-* vanished completely, being replaced by **nă* 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 / Yeniseian) **KʷV-* — 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!) **u̯V,* but the matter still needs further investigation.

28. **tongue** — 舌 MC *źet,* OC **lat.* The closest parallel is the Kach. form *šiŋ-let*[31] 'tongue', pointing to PST **(s)-lăt* / **(s)-lĕt*. The Chinese and Kachin form can hardly be taken away from the rest of Sino-Tibetan forms, pointing to **s-lăj* / **m-lăj* : cf. Tib. *lće,* Brm. *hlja,* Lush. *lei* (PKC **m-lei*), Dimasa *salai,*

Rawang *phəlɛ*, Newari *me ~ mye* etc. [IST: 40, 121, 124, 136, 166, 434]; STC № 231 (**m-lay ~ *s-lay*).

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

Some questions may be raised by the correspondence **-ć-* : **-j* (normally **-t* would be expected, and, indeed, the Chinese and Kachin form possibly reflect the more regular variant **mlăt ~ *slăt*). This was probably conditioned by the position of **-c̣-* within the cluster **-lc̣-*, and in addition there exists another extremely similar case: PNC **mĕlʒ́:V* 'place, ground' (Chech. *mott*, Tsez. *moči*, Arch. *maǯ* etc.) corresponding to PST **m-ləj* 'earth, ground' (Brm. *mrij*, Lush. *lei*, Nung *məli* etc. [IST: 24, 184]; STC № 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 *ćhí*, OC **thəʔ*. So far no satisfactory etymology was proposed for this root. *Xie-sheng* (phonetic 止 **təʔ*) seems to point to OC **thəʔ*; Min dialects, however, have a quite unexpected velar initial (Xiamen, Chaozhou, Fuzhou *khi*^{3}). 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^{w}a* (Tib. *so*, Brm. *swah*, Lush. *ha*, Kach. *wa*^{33} etc. [IST: 53, 124, 135, 407, 430]; STC № 437 **s-wa*). It has, however, no Chinese reflex and no reliable external parallels (cf. perhaps Proto-Lezghian **säχ^{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] **ŋipən*, very probably corresponding to PT **van ~ *ven* id. [AT: 412], but having no cognates in ST or Chinese.

30. **two** — 二 MC *ńì*, OC **nij-s*. The word is a very clear descendant of PST **K-nĭj(s)*, cf. Tib. *gńis*, Brm. *hnać*, Lush. *hniʔ*, Kach. *ni*^{33}, Dimasa *gi-ni*, Kiranti **ní(k)* 'two' etc. [IST: 37, 135, 411, 429]; STC № 4 **g-nis*.

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

The PY form for 'two' is **xɨna* (Ket *ɨ̄n,* Yugh *ɨn,* Kott *īna,* Ar. *kina,* Pump. *hine-aŋ*). This is probably an exact match for PST **K-nĭj(s)*, although the loss of **-s* is unclear (we would rather expect **xɨnas*).

The comparison of PST **K-nĭj(s)*, PY **xɨna* and PNC **nă̆ws̬i* suggests that the first two forms reflect some archaic compound. Indeed, another widespread NC root for 'two' is **q̇Hwā̆* which can correspond to PY **xɨ-* 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] **ḍuva',* [PANDYMC] **DewS₃a?;* a very similar form is Proto-Li **draw* [AT: 211]. The root may be very archaic (cf., e. g., the similar-looking PIE **du̯ō(u)* 'two'), but has nothing in common with the PST and OC numeral 'two'.

31. **water** — 水 MC *śwí.* This word is usually reconstructed with a lateral initial (cf., e. g., in [DEZC: 570]: **hljuəj?;* in our reconstruction it would be **sluj?* or **ɬuj?*). However, there is no *xie-sheng* 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 **ćwí* < 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 *cɨ*[1], Fuzhou *cü*[1]; 叔 MC *śük* 'father's younger brother' < **tikʷ,* cf. Xiamen *cik*[7], Chaozhou *cek*[7] etc. — all containing dental stops in *xie-sheng* 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. [IST: 48, 442]; STC № 55; [STC: 45, 134].

A very probable parallel is PNC (PEC) **-V̄ṭwV* 'to pour, soak' (cf. Btsb. *-oṗṭ-* 'to soak', Av. *ṭe-,* Tsez. *eṭʷ-,* Lak *-uṭi-* 'to pour' etc. [NCED: 1034–1035]).

In PAN two basic words for 'water' are reconstructed: [VL] **[dḍ]anum,* [PANDYMC] **DaNúme* (with a quite interesting parallel in PT **na(a)m*

'water' [AT: 420]); [VL] **vajəɣ*, [PANDYMC] **wahiR* (with a not quite secure parallel in Proto-Li **ya* 'river' [AT: 420]). Neither of them reveals any resemblance to the Chinese word. L. Sagart [CA: 44] compares it with PAN [VL] **'aluɣ*, [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 *ɣâ*, 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 *ɣo*, OC **ghā* 'how, what': Xiamen *ɔ²*, Chaozhou *ou²*).

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 **x̲V* 'who, what' (Av. *x̲i-* 'who (obl. stem)', Inkh. *λu* 'who, what', Tab. *fu* 'what', Ad. *xa-t* 'who' etc. [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 *cui²*, Fuzhou *sui²*). 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 **tŭ* 'who, which' (also having some paradigmatic modifications, notably **tŭ-j*), cf. Tib. *du* 'how many', Lush. *tu* 'who; whose'; Kach. *gəde³¹* 'how many', Brm. *a-ti* 'which'.

PST **tŭ* can be compared with PNC **ʔāādV* '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. [NCED: 244–245]). In PNC, however, this is not one of the basic interrogative pronominal roots: it is rather a specialization of **ʔāādV* '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 **tŭ* has replaced the original **su* 'who' (Tib. *su*, Brm. *a-su*, Kiranti **su* etc. [IST: 135]). The latter corresponds to PNC **s̲aj* 'what', PY **ʔas-* / **sV-* id. [HGC: 23].

It is interesting to note the semantic and phonetic proximity of Proto-Austronesian [VL] **[t']a[j]i[']*, [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 *xie-sheng* 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. [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 is thus rather hypothetical. In any case, OC **prəm* has replaced the basic ST root for 'wind', **lij* (reflected in Brm. *lij,* Lush. *thli* 'wind', Kach. *buŋ-li*[55] 'breeze' etc. ([IST: 195]; STC № 454 **g-liy*) and corresponding to PNC **λwəɬʔV* 'wind', PY **ʔuĺ-* 'whirlwind' [HGC: 30]).

The root for 'wind' in Proto-Austronesian is [VL] **ʼaŋin,* [PANDYMC] **h₂áŋin*, without any ST or OC parallels.

35. **year** — 年 MC *nien,* OC **nhīn* (for **nh-* cf. archaic Min forms: Shaowu *nin*[7], Chaozhou *hĩ*[2]). Also means 'harvest' in OC, but this meaning is secondary: cf. the basic ST root for 'year', **snīŋ* (Brm. *hnać* < **hnik* — for this development see above, Kach. *šəniŋ*[33], etc. [IST: 38, 410]; STC № 368 **s-niŋ*).

The root combines the meanings 'year' and 'old, ancient' (cf. Tib. *rńiŋ* — 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* > **snīŋ,* with a usual in these cases shift of the final resonant).

The root has a very probable Sino-Caucasian etymology: cf. PNC **śwänɨ̌* 'year' (Av. *son,* Lak *šin,* Rut. *sän,* Ub. *šʷa* etc. [NCED: 975–976]); PY **sin-* 'old, ancient' (Ket *śin,* Yugh *sin*), **s[i]-Ga* 'year' (with a temporal suffix **-Ga*; cf. Ket *śī,* Yugh *sī,* Kott *šēga,* Ar. *šhej,* 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-ɤhīt* :	**s-ʔʷīj*
2) ‘die’	**sij* :	**sǐj*
3) ‘dog’	**khʷī-n* :	**qhʷīj*
4) ‘ear’	**nhəʔ* :	**nă*
5) ‘eye’	**mhŭk* :	**mjV̆k*
6) ‘fire’	**s-m[ē]jʔ* :	**mēj*
7) ‘fish’	**ŋha* :	**ŋ(j)ă*
8) ‘full’	**leŋ* :	**Plăiŋ*
9) ‘give’	**pij-s* :	**piaj*
10) ‘horn’	**krō-k* :	**Kruā*
11) ‘I’	**ŋhājʔ* :	**ŋā(j)*
12) ‘louse’	**srit* :	**śrik*
13) ‘name’	**mheŋ* :	**măiŋ*
14) ‘new’	**sin* :	**sĭŋ (/-ŋ)*
15) ‘one’	**ʔit* :	**ʔĭt*
16) ‘salt’	**ʒ(h)āj* :	**C(w)āj*
17) ‘sun’	**ni-t* :	**nĭj*
18) ‘tail’	**măjʔ* :	**măj*
19) ‘this’	**də,*deʔ* :	**tă*
20) ‘thou’	**nhaʔ* :	**na-*
21) ‘tongue’	**lat* :	**mlăj / *mlăt*
22) ‘two’	**nij-s* :	**K-ni(s)*
23) ‘water’	**tujʔ* :	**tŭj*
24) ‘year’	**nhīn* :	**snīŋ*

We can perhaps also add interrogative pronouns (25: ‘what’ **ghā-* : **qhā-*; 26: ‘who’ **du-* : **tŭ-*), 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’; **laʔ* ‘give’ : **lă* ‘take’; **ʎuʔ* ‘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 Yeniseian languages. Old Chinese has at least 13 precise phonetic and semantic matches with PNC in the 35-wordlist:

1) ‘blood’ **s-whī-t* (< **s-ʔʷīj*) : **hwĕʔnV*
2) ‘dog’ **khʷī-n* (< **qhʷīj*) : **χ̱ħwĕjV*
3) ‘ear’ **nhəʔ* (< **nă*) : **ʕwănʕV*
4) ‘give’ **laʔ* (< **lă*) : **-ĭɬ̱V*
5) ‘horn’ **krō-k* (< **Kruā*) : **qwīrhV*
6) ‘moon’ **ŋot* : **wə̆mc̣ŏ*
7) ‘new’ **sin* (< **sĭn /-ŋ*) : **c̣ă̆nʔV*
8) ‘one’ **ʔit* (< **ʔĭt*) : **Hə̆cə̆* (/ **cHə̆*)
9) ‘salt’ **ʒ(h)āj* (< **C(w)āj*) : **c̣wĕnhV*
10) ‘tail’ **mə̆jʔ* (< **mə̆j*) : **mēʁV*
11) ‘tongue’ **lat* (< **mlăt*) : **mĕlc̣ĭ*
12) ‘two’ **nij-s* (< **K-ni(s)*) : **nă̆wɬ̱ši*
13) ‘year’ **nhīn* (< **snī-ŋ*) : **śwänɨ̆*

We can perhaps add 14 (‘fish’ **ŋha* : PNC **χ̱wanħV*), although there are some phonetic doubts (see above), and the interrogative pronoun (‘what’ **gʰā-* : PNC **χ̱V*), 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’, **ʎuʔ* ‘hand’, **ŋhājʔ* and **la* ‘I’, **srit* ‘louse’, **deʔ*/**də, *se, *ʔi* ‘this’, **tujʔ* ‘water’, **du-* ‘who’).

Proto-Yeniseian 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 Yeniseian 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’ **rhōnʔ* (< **rhōrʔ*) : **CelúR*
2) ‘give’ **pij-s* (< **piaj*) : **bəRaj* (if **-R-* > *-Ø-* in PST)
3) ‘one’ **ʔit* (< **ʔit*) : **ʔisáʔ*
4) ‘this’ **də* (< **tə̆*) : **ʔitu(h)*

One could also add № 5 (‘full’ **leŋ* < **Plăi-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-wordlist 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 **rhāʔ* 'salt' : PAN **siraq* 'condiment'; OC **də* 'this' : PAN **idi* 'that, there'; OC **lat* 'tongue' : PAN **dilat* 'lick'; OC **nhaʔ* '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* (< **qhʷīj*) : PIE **ḱwōn-*

2) 'full' OC **leŋ* (< **Plăi-ŋ*): PIE **pelə-*

3) 'horn' OC **krō-k* (< **Kruā*) : PIE **ḱr-n-*

4) 'I' OC **ŋhā-* (< **ŋā*) : PIE **me-* (cf. a similar development **ŋ-* > **m-/-ŋ* in Yeniseian)

5) 'moon' OC **ŋot* : PIE **mēns-* (the word has a certain resemblance to the supposed Sino-Caucasian **wəŋçŏ,* see above)

6) 'name' OC **mheŋ* (< **măiŋ*) : PIE **enom-(n)*

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

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 **bəɣaj* —	PT **haɨ*
5) 'hand'	AN **lima'* —	PT **mya*
6) 'I'	AN **'aku'* —	PT **kaw*
7) 'louse'	AN **kutu'* —	PT **thraw*
8) 'moon'	AN **bulan* —	PT **ʔblɨan*
9) 'name'	AN **[']aǵan* —	PT **ʔdaan*
10) 'this'	AN **'ini'* —	PT **ni* / **nay*
11) 'tooth'	AN **[']ipən* —	PT **van ~ ven*
12) 'water'	AN **[dḍ]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’ (**baɣu* : **hmaɨ*), 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.”