George Starostin
National Research University Higher School of Economics, Moscow / Santa Fe Institute; gstarst1@gmail.com

Lexicostatistical studies in Khoisan II/2: Towards a more precise phylogeny for the Tuu family

The paper is the second part of an extensive study aimed at clarifying the internal relationships within the Tuu (Southern Khoisan) linguistic family of South Africa and reconstructing a reasonably accurate Swadesh wordlist for Proto-Tuu. In this section, I first investigate the issue of extensive areal contact between two languages belonging to two different subgroups of Tuu (N|ng and ǀ'Auni), filtering out potential borrowings from the former into the latter which may obscure both etymological judgement and lexicostatistical calculations. Next, lexicostatistical matrices and resulting phylogenetic trees are offered for the entire family, demonstrating that a ternary model, in which Proto-Tuu splits into three more or less equidistant branches (!Ui, Nossob, and Taa), is likely preferable to any possible binary models; arguments for a closer proximity between Nossob and Taa are analyzed and found generally inconclusive. Finally, some remarks are made about the reconstructibility of the Swadesh wordlist for Proto-Tuu and on some of its peculiar properties as compared to attested Tuu languages. The Appendix section contains the entire second half of the Swadesh wordlist (items 51–100) as reconstructed on intermediate levels and on the Proto-Tuu level.

Keywords: South Khoisan languages; Tuu languages; click languages; lexicostatistics; basic lexicon; onomasiological reconstruction.

Preliminary remarks

In the first part of the present paper, published in an earlier issue of JLR (Starostin 2021), I have given a brief description of the principal goals, methods, and problems (both technical and substantial) of conducting a detailed lexicostatistical survey of all known languages of the Tuu family; the theoretical part was then illustrated with actual comparative material from the first half of the Swadesh wordlist (items 1–50). Subsequently, this second part of the study, in addition to presenting data on the second half of the list (see the Appendix at the end of the paper), will focus on the actual analysis of the data, starting with some simple statistics and their phylogenetic interpretations, and then concentrating on more "fine-grained" manual analysis of the relevant data. The overall plan is as follows:

1) first, I shall directly address the issue of *areal contact* between the surveyed languages, most notably, concentrating on the consequences of N‖ng-ǀ'Auni bilingualism which may or may not have resulted in incorrect phylogenetic conclusions in the past;

2) next, I shall present the *lexicostatistical matrices* for the full 100-item wordlists for all languages and doculects of the survey, along with their glottochronological interpretations and a brief discussion of their problematic areas;

3) finally, identification of problems stemming from the crude statistical approach will lead to a manual re-analysis of some of the potential matches in terms of areal contacts, shared innovations, or retained archaisms, resulting in an overall grading of potential phylogenetic schemes in accordance with their respective strengths and weaknesses.

Areal contact in Tuu: the issue of isoglosses between N‖ng (N|uu) and |'Auni

In theory, any non-accidental lexical similarity between two languages that are in contact with each other today or may have plausibly been in contact with each other some time in the past may be due to lexical borrowing, regardless of whether the languages in question are genetically related. In practical terms, however, we usually evaluate such a possibility based on what could be called the "common sense scenario". Thus, if such similarities are seen as proportionally more common across basic rather than cultural vocabulary (or, at least, if they are more or less equally spread across both layers), it makes sense to assume that they stem from a common ancestor, and that special, additional arguments are needed to challenge that default assumption. (That any such lexical similarities should also display recurrent phonetic correspondences is also a vital requirement, but it is somewhat irrelevant for this purpose, since systemic borrowings between languages tend to also take place in accordance with strict correspondence patterns). Such additional arguments in favor of areal contact may come from:

phonetics — if potential cognates between languages A and B feature more than one set of recurrent correspondences, this is a clear indication that at least one of the sets takes place in more recent loans from one language to another (e.g. the case of English and French, or modern Chinese "dialects" vs. the literary Chinese language);

distribution — if language A seems to be related to language B, but language B is clearly more tightly connected to languages C and D, which, on their own, show relatively little affinity with language A, the increase in similarity between A and B is almost certainly due to secondary contact (e.g. the case of Armenian and Iranian languages, see Campbell & Poser 2008: 80);

morphology and semantics — if the formal and semantic properties of lexical items in language A consistently match only a small subset of the respective properties of lexical items in language B (e.g. the words in A are attested only as grammatically complex stems rather than simple roots, or only in what looks like secondary / figurative meanings), this can be interpreted as the result of borrowing from B rather than inheritance from a common ancestor.

When this line of reasoning is applied to the lexical data of Tuu languages, it can be seen that, on the whole, it is difficult to suspect most of them of ever having been in intense secondary contact with each other. Although phonetic correspondences between them, as has already been partially shown in the first part of the paper, sometimes give the impression of being fairly erratic (and it is not always clear how much of that is due to actual phonetic change and how much to inaccurate data transcription), it has so far been impossible to properly identify two distinct sub-sets of correspondences for any given pair of Tuu languages. Distribution-wise, as we shall clearly see below from the lexicostatistical matrices, most of the languages support a rather transparent phylogenetic structure with little out of the ordinary. As for morphology and semantics, most of the etymological work done on Tuu so far finds more evidence for morphological variation on the Proto-Tuu level, or between the different branches of Tuu, than for secondary convergence processes based on the borrowing of words in various fossilized morphological patterns from one Tuu language into another.

Thus far, the lion's share of loans in Tuu basic lexicon has been identified as stemming from non-Tuu languages — most commonly, Khoe (Traill & Nakagawa 2000), but also occasional borrowings from Bantu (especially in ‖Xegwi, see Lanham & Hallowes 1956a) or even European sources (modern N|uu is particularly liable to be influenced by Afrikaans, see Sands et al. 2007).

One notable exception to this tendency is the complicated relation between |'Auni, one of two known languages from the Nossob subgroup, and the N‖ng cluster from the !Ui subgroup. There is no need to go into detail here on the nature of the alleged N‖ng-|'Auni bilin-

gualism that resulted in the adoption of many N‖ng words into |'Auni itself; the sociolinguistic foundation for these processes was already described in early works such as Bleek 1937, and later extensively commented upon by Güldemann (2014, 2018). What interests us specifically is a procedure that would allow to realistically distinguish between "primary" (inherited) cognates and "secondary" (borrowed) cognates between these two languages; the correct solution of this puzzle would not only be instrumental in more precisely determining the phylogenetic status of the Nossob languages, but would also be of interest to any specialists beset with similar problems in other linguistic areas of the world.

First, let us draw up a complete table (Table 1) of all potential lexicostatistical (i.e. not just related, but fully equivalent in terms of semantics) matches between |'Auni and N|uu; for the latter, I list the old ‖Ng!ke forms as recorded by D. Bleek (1956, 2000), which would be likely to serve as the actual source of the borrowing, and their modern day N|uu forms as reflecting more accurate transcription.

Table 1. Lexicostatistical matches between |'Auni, "Old N‖ng" (‖Ng!ke) and Modern N|uu.

Word	**|'Auni**	**‖Ng!ke**	**N|uu**
'bird'	si=|u	|wí ~ |wi:	|qʰui-si
'blood'	‖xauu	‖xau	‖xau-ke
'claw /nail/'	‖ora-sa	‖uri-si	‖qoro-si
'come'	sa ~ sé ~ sí	si ~ se ~ seya ~ sa	sa: ~ ca:
'die'	|'ã	|'a:	|'a:
'dog'	ǂɔ̃:	!wiŋ	ǂʰun ~ ǂʰuŋ
'drink'	x'ã: ~ x'ẽ	x'a: ~ x'ã ~ x'ẽ	x'ãĩ
'ear'	ǂui	ǃwe: ~ ǃwe:-ntu ~ ǃu:-ntu	ǂui-si
'eat'	ǹ̀ ~ hà ~ hàa	ã ~ ẽ ~ ẽĩ	ʔã
'eye'	coo / c'axu	cáxu ~ ca:xem	c'axam
'fire'	|'i	|'i	|'i:
'fly'	zé	—	ze:ˢ
'give'	a	a ~ a:	ʔã:
'hair'	|ʰóo	|u ~ |ʰú	|ʰu:-ke
'hand'	|kx'a ~ |kx'an	|kx'a	|kx'a:
'head'	ǁa: / x:uu	ǁa ~ ǁa:	ǁa:
'hear'	tu: ~ tu:i	tu ~ tú ~ tu:i	ɕu:
'heart'	|'e: ~ |'ɛ:	|ai ~ |e	|e:
'horn'	‖ẽĩ	‖ãĩ	‖qʰoe-si
'I'	n ~ ŋ ~ na ~ m	ŋ ~ n	ŋ
'know'	‖xai ~ ‖x'e-ki	‖ai	‖xae
'lie'	tòa	tia ~ kia:	ɕa:
'long'	|'ǎ-si	|'a:	|'ã:
'meat'	θwe ~ θwi	θwai: ~ θwai	θoe
'mouth'	tu ~ tʰu / ǂu:	tu ~ tu:	ɕu:
'name'	|ẽ ~ |ẽn	|ẽ	ka=|ĩ
'neck'	ǂõĩ / ‖ú	!ú ~ kú	ǂqu:
'night'	‖àu ~ ‖ò	‖a ~ ‖a:	‖a:
'nose'	ǁõ	ǁu-tu	ǁu-ɕu

‘one’	ǂʼǘ ~ ǂʼǔ-u	ǁwe: ~ ǁʼwe:	ǁʼoe
‘rain’	ǁʰà:a	!ʰa ~ !a:	(ǂqau)
‘road’	!án	(tirau)	!an ~ !aŋ
‘say’	ko / ǀu	ka	ka
‘see’	ǰà: ~ ǰe	ǰa: ~ ǰe ~ ǰi: ~ ǰi	ǰa:
‘sit’	sã ~ são ~ so	so ~ so: ~ sɔ:	sũĩ
‘sleep’	θwõi	θwoiŋ ~ θwoeŋ ~ θóeŋ	θun ~ θuŋ
‘smoke’	ǁáu	(ǀwi:)	ǁoːˀ-ke
‘this’	a	a	a
‘thou’	a	a	a
‘tongue’	ǀʼãri	ǀʼẽ	ǀʼãn ~ ǀʼãĩ
‘tooth’	ǁẽĩ	ǁãĩ: ~ ǁẽ: ~ ǁẽĩ	ǁʰãĩ
‘tree’	ʘwa:a ~ ʘwa:-sa	ʘo ~ ʘo: ~ θʼʰo	θo:
‘walk /go/’	ǁʼa ~ ǁʼe ~ ǁa(:)	ǁʼa(:) ~ ǁʼai	ǁʼaʔa
‘water’	kʰá: ~ kʰáá ~ kʰái	!ʼʰa: ~ !ʰa ~ !àːˀ ~ !a: ~ ǁʰa:	!qʰa:
‘we’ (excl.)	si ~ se ~ ci	si	si
‘we’ (incl.)	i ~ e	i	i
‘woman’	ǀɛ́:	ǀʼai- ~ ǀ̰ai- ~ ǀ̰e:-	ǀe:-ki

Taken in the general context of other Tuu languages, these parallels may be divided into the following subgroups:

(1) “Pan-Tuu” roots that are present in the same basic meaning in all or most Tuu languages covering all three subgroups. Provided that within ǀʼAuni they do not feature any specific “non-ǀʼAuni” traits (e.g. an uncommon phonetic shift or morphological add-on that is more typical of N‖ng than ǀʼAuni), there is no compelling reason to treat such forms as borrowings.

These roots are: ‘bird’, ‘come’, ‘die’, ‘dog’, ‘drink’, ‘eat’, ‘fire’, ‘hair’, ‘hand’, ‘heart’, ‘horn’, ‘I’, ‘meat’, ‘name’, ‘nose’, ‘one’, ‘see’, ‘sit’, ‘sleep’, ‘thou’, ‘tongue’, ‘tree’, ‘water’, ‘we’ (both excl. and incl.).

(2) Logically close to this group are “Pan-!Ui + Pan-Nossob” roots which have no lexicostatistical (or even etymological) parallels in Taa, but are found in both ǀʼAuni and ǀHaasi. Since we have no evidence of lexical contacts between N‖ng and ǀHaasi, and since, once again, there are no specific phonetic or morphological arguments for their being borrowed into ǀʼAuni, we can essentially merge them with the first group for our purposes, also considering them most likely inherited in ǀʼAuni.

These roots are: ‘blood’, ‘hear’, ‘night’, ‘this’.

(3) Cases when ǀʼAuni and ǀHaasi contradict each other, and the ǀʼAuni form is closer to N‖ng than the ǀHaasi one. This is an apriori suspicious situation for which no universal solution is available; each case has to be evaluated on its own. These roots are:

(3a) ǀʼAuni *‖ora-sa* ‘nail’ vs. ǀHaasi *kʼa=ǂü* id. It is not likely that ǀHaasi preserves the original situation; in fact, it is even possible that Story’s semantic glossing here is inaccurate, since, strictly speaking, the form he quotes has to be literally translated as ‘fingers’ (plural prefix *kʼa=* + *ǂü* ‘finger’). Meanwhile, even if ǀʼAuni *‖ora-sa* is indeed quite similar to ‖Ng!ke *‖uri-si*, N|uu *‖oro-si*, their codas and morphological properties are different enough to reject borrowing as the likeliest solution. Since this is really a “Pan-Tuu” root as well, we prefer to treat the ǀʼAuni item as inherited;

(3b) ‘give’ — according to Bleek 1937, ǀʼAuni has a variety of synonyms here, only one of which, the verb *a*, attested in a single textual example (‘give me tobacco’), has reliable parallels

in !Ui (**a* is the main verb of giving in both |Xam and the entire N‖ng cluster). |Haasi, on the other hand, has *i*, another monovocalic root which is probably not related. Technically, this could be the result of N‖ng influence on |'Auni, so we are justified in marking this item as a potential (uncertain) borrowing;

(3c) |'Auni 'go' *‖'a ~ ‖'e ~ ‖a(:)* vs. |Haasi *ǂa* id. Etymology of the |Haasi form is unclear; the |'Auni stem, on the other hand, is the same as the "Pan-!Ui" equivalent for the same meaning (**‖'a-* or, perhaps, **‖'aʔa*, as in Modern N|uu). Given that the same root is also found in Taa in a more specialized meaning (!Xóõ *‖'âe* 'to go out hunting or gathering'), it cannot be considered a lexical innovation in Proto-!Ui, and chances of it being lost and then restored in |'Auni under N‖ng influence are low; we are fully justified in treating it as an inherited item.

(4) Cases when reflexes of a "Pan-Tuu" or a "Pan-!Ui + Pan-Nossob" root in |'Auni bear a notable resemblance specifically to N‖ng. Again, these cases have to be evaluated individually:

(4a) |'Auni *ǂ̃ui* 'ear' features the same morphological structure as N|uu *ǂ̃ui(-si)* (though, curiously enough, not the same as ‖Ng!ke *ǃ̃we:-ntu*). Even more importantly, the root shape of the word for 'ear' in |Haasi is *ǂ̃a-*, which is morphologically comparable with !Xóõ *ǂ̃ũʰ-ã̀* (assuming that *ǂ̃a-* ← **ǂ̃u-a-* with vowel contraction) — another slight indication that the |'Auni form may have been borrowed from N‖ng or, at the least, "influenced" by it in some way. Considering that other probable instances of N‖ng body part terms finding their way into |'Auni also emerge (see below), we are within our rights to mark this form as a potential (though not certified) borrowing;

(4b) somewhat more difficult is the situation with |'Auni *‖ẽĩ* 'tooth'. This form is explicitly glossed as 'tooth' only in an early source (Bleek 1929: 86), but is not found in this meaning in Bleek 1937, where the only attested meaning for it is 'horn' ('horn' and 'tooth' are transcribed as if they were homonyms in quite a few !Ui doculects, but more reliable data from ‖Xegwi and Modern N|uu show that they are, in fact, quite different roots). In any case, it very closely matches ‖Ng!ke *‖ẽĩ* 'tooth', whereas |Haasi has *k'i=‖ɛ* 'teeth' without the nasal coda. Still, the distance between these forms is not as large as in the case of 'ear'; moreover, the morphological similarity between |'Auni and N‖ng is not exclusive in this instance (e.g., |Xam also has a nasal coda), so we tentatively continue to regard the |'Auni form as inherited.

(5) "Doublets", when Bleek records two forms for |'Auni, one of which is almost always closer to N‖ng than the other. These include:

(5a) 'eye' — |'Auni *coo* vs. *c'axu*. The latter is almost identical to ‖Ng!ke *cáxu*, ǂKhomani *c'axu* etc.; the former, on the contrary, is closer to the contracted form *cxɔ*, found in |Haasi. Bleek's notes give no hints as to which of the two was the most commonly used, "neutral" form; it is permissible to simply disregard *c'axu* in the calculations and take *coo* as the inherited form, with a |'Auni-specific simplification of the cluster **cx-*, which itself appeared as a result of contraction from **c'a-x-* in Proto-Nossob;

(5b) 'head' — |'Auni *ǃ̃a:* vs. *x:uu*. The former is unquestionably the Proto-Tuu equivalent for 'head', continuing to function as such in every attested doculect of !Ui and Taa. However, in |Haasi it is not encountered at all; instead, we have *(ŋ=)xɔ* 'head' (the meaning is clearly confirmed with multiple text examples), cognate with Proto-!Ui **xu* 'face' and most likely reflecting the semantic shift 'face' → 'head'. Since the exact same form is encountered in the closely related |'Auni, it would be logical to postulate that shift on the Proto-Nossob level and analyze |'Auni *ǃ̃a:* as a re-borrowing from ‖Ng!ke (this scenario is explicitly advocated for by D. Bleek herself). We can add *ǃ̃a:* to the list of potential borrowings;

(5c) 'mouth' — |'Auni *ǂu:* vs. *tu ~ tʰu*. The situation here is similar to 'head', except that this time, **tu* is specifically Pan-!Ui, not attested in Taa. Again, |Haasi has *ǂa* 'mouth', cognate with |'Auni *ǂu:* and, further on, with Proto-Taa **ǂu-* 'mouth' (note the exact same vocalic correspon-

dence as in ‘ear’, again hinting that |Haasi *ǂa* may be historically contracted ← **ǂu-a*); |’Auni *tu* subsequently looks like a borrowing from N‖ng (also suspected by D. Bleek). Another potential borrowing;

(5d) ‘neck’ — |’Auni *ǂõĩ* vs. *‖ú*. Here, the situation is different. The former stem does indeed seem more similar to N|uu *ǂqu:* ‘neck’ than the latter, but this similarity is hindered by differences in the codas, as the |’Auni form seems likely to reflect an original morphological structure like **ǂo-iŋ* or **ǂo-ni*, whereas neither Modern N|uu *ǂqu:* nor ‖Ng!ke *!ú* ‘neck’ show any traces of nasal consonants. It is more likely here that the first form is genetically related to !Ui, and the second simply reflects a different root (it is impossible to tell the semantic difference from Bleek’s records);

(5e) ‘say’ — |’Auni *ko* vs. */u*. Both verbs are used to introduce direct speech and are found virtually interchangeable with each other in Bleek’s recordings of |’Auni texts. In |Haasi, the only attested form for ‘say’ is */wa*, obviously cognate with |’Auni */u*. The verb *ku* ‘say’ is one of the main speech verbs in N‖ng, so it is more than likely that */u* in |’Auni is the inherited term, whereas *ko* is a N‖ng verb that used to occasionally substitute it, i.e. a likely borrowing;

(5f) ‘woman’ — |’Auni */ɛ̃́:* vs. *ʆe:ki*. Both forms feature the same lexical root, but the morphological shape of the second variant is the same as in N|uu (cf. the ‖Ng!ke variants listed by Bleek: *ʔai-ti ~ ʔai-ki ~ ʆai-ti ~ ʆai-ki ~ ʆe:-ki*), whereas the first variant, showing nasality and fully matching |Haasi */ĩ* id., is especially typical of the Nossob area (more on this in the Appendix below). Most likely, the second variant is a borrowing, while the first one should be counted as an inherited term. (It is interesting that in Bleek’s texts, *ʆe:ki* ‘woman’ is encountered far more often than */ɛ̃́:* — yet there is at least one text in which the two are essentially interchangeable, so we cannot assume that in colloquial |’Auni the newer borrowing had completely displaced the original lexeme).

(6) Cases where |’Auni and N‖ng feature exclusive isoglosses vs. everybody else. There is only one of these in the 100-item wordlist, and it is almost certainly a borrowing: |’Auni *‖xai* ‘know’ = ‖Ng!ke *‖ai* id. Meanwhile, |Haasi has */üma* ‘know’, cognate with !Xóõ *ʆûmã* id. This is a case of “criss-cross” distribution, breaking up a well-established phylogenetic structure; typically, in such cases one match at best reflects inheritance and the other one should be ascribed to contact, and since there is little linguistic evidence for secondary contact between |Haasi and Taa (although geographically this would be possible), it makes more sense to treat the |’Auni form as a potential borrowing.

(7) Various unclear cases:

(7a) ‘fly’ — |’Auni *zé* and Modern N|uu *ze:ˢ* are clearly the same item, and it is highly likely that they are further connected to Taa forms such as !Xóõ *ʒãĩ̃ʰ*, Kakia *žõĩˢ ~ žweˢ*. This would seem to be a fine Proto-Tuu candidate, but the problem is that Bleek’s ‖Ng!ke has *‖óu* ‘to fly’ = |Xam *‖au ~ ‖ʰau ~ ‖ʰóu* id.; meanwhile, the verb **zVˢ* ‘fly’ seems to have a general areal distribution, perhaps amplified by its sound-symbolic nature (cf. also ǂHoan *zoeˢ* ‘fly straight’, Ju|’hoan *zõĩˢ* ‘to swarm /of bees/’, Naro *cãẽˢ* ‘to fly’ etc.). All of this raises suspicions that N|uu *ze:ˢ* may be a relatively recent innovation, and the |’Auni form could be easily borrowed from N|uu — although the genetic explanation cannot be fully ruled out either;

(7b) ‘lie’ — |’Auni *tòa* features a distinctly different coda from Proto-!Ui **ta* and Proto-Taa **tu*, but given the usual amount of vocalic variation in verbal stems, this is not enough to deny it cognacy with both of these forms; at the same time, it is notably different from N‖ng to be judged as a borrowing. We treat it as inherited from a Common Tuu root **tV-*;

(7c) ‘long’ — |’Auni */ʔã́-si* is a good match with ‖Ng!ke */ʔa:*, N|uu */ʔã:*, as well as ‖Xegwi-Z */ã* id. (although the mismatch in click effluxes with the latter is a little puzzling). Unfortunately, no equivalent for ‘long’ is recorded in |Haasi, so there is no sure way of knowing here if the

word was borrowed or not (addition of *-si* means little, since it is a highly productive adjectival suffix in |'Auni). The word may be marked as a potential, but inconclusive, borrowing due to lack of significant evidence to the contrary;

(7d) 'rain' — |'Auni *ǁʰà:a*, glossed as 'water, rain' in Bleek 1937, seems to be just a phonetic variant of *kʰá:* 'water' (← Proto-Tuu **!qʰa*); it is not clear if this polysemy may be reliably projected onto the Proto-Tuu or even the Proto-!Ui level (see RAIN in the appendix for more information), but in any case, there is hardly any sufficient reason here to suspect the word of being a borrowing from N‖ng;

(7e) 'road' — |'Auni has *!án* here, phonetically identical with Modern N|uu *!an* ~ *!aŋ* which, in turn, is cognate with ‖Xegwi *kaŋ* and can thus go back to Proto-!Ui **!an* 'path, road'. Since the word 'road, path' is not attested for |Haasi, there is no way of telling if it is inherited or borrowed. It may be noted that there are two more words with the same meaning in Bleek 1937: *ǁùru* and *ǂkx'ei*, without any recorded semantic distinctions (and without any external etymologies) — this may be a hint that at least one of them may be the inherited term, while *!án* is really a borrowing, but all of this is inconclusive. Again, for specific purposes the item may be marked as at least a potential borrowing;

(7f) 'smoke' — although |'Auni *ǁáu* is phonetically similar to N|uu *ǁoːˀ-ke*, it is even more so to |Haasi *ǁau*, implying Proto-Nossob **ǁau* and a genetic connection to the N|uu form (as well as |Xam *ǁɔ́ː*). No need to suggest borrowing in this case.

Altogether we have thus identified 28 matches which offer no specific arguments for being treated as loans (groups 1 and 2), 9 matches where the evidence clearly points to inheritance as the most likely reason for similarity, and 9 matches where the evidence is either ambiguous or points out to borrowing as the likeliest scenario ('give', 'ear', 'fly', 'head', 'mouth', 'say', 'know', 'long', 'road'). While it may, of course, be possible that some of the judgements presented here are erroneous due to insufficient data, any particular errors would be likely to outbalance each other (i.e. undetected borrowings would be compensated for by falsely assumed borrowings), meaning that the phylogenetic results received from a dataset in which these 9 forms are marked as borrowings will probably be more reliable on the whole than results from a dataset in which they are marked as inherited. Nevertheless, for the sake of a more accurate experiment we shall apply the standard lexicostatistical procedure to both sets, and compare the results.

Tuu phylogeny based on the classic lexicostatistical method

The first step is to construct a standard lexicostatistical matrix for all compared languages. In this, we rely on the well-tested Swadesh / Starostin method (see S. Starostin 2000), which requires preliminary elimination of all borrowed items from the dataset in order to produce a "normalized" matrix and avoid potential errors in the resulting tree structure (as well as glottochronological dates). This correction is particularly essential for situations of "mass borrowings", which can drastically speed up the rate of lexical replacement; within Tuu (if we discount obvious minor impediments such as the presence of Afrikaans lexemes in Modern N|uu, etc.), such a situation is only found between N‖ng and |'Auni. However, as has already been mentioned above, for the sake of additional transparency we shall first construct the matrices according to the "maximalist" principle, i.e. pretending that (perhaps) all of the attested matches between these two languages are due to inheritance, not contact. The resulting matrix (Table 2) is as follows:

Table 2. "Maximalist" lexicostatistical matrix between Tuu languages

	ǁNg!ke	ǂKhomani	N|uu	ǁXegwi	|'Auni	|Haasi	!Xóõ	Kakia	N|uǁen
|Xam	74 %	79%	73%	59%	53%	43%	42%	40%	43%
ǁNg!ke		85%	85%	59%	59%	43%	43%	41%	43%
ǂKhomani			95%	74%	69%	53%	57%	56%	60%
N|uu				64%	59%	48%	47%	45%	48%
ǁXegwi					50%	40%	41%	41%	46%
|'Auni						72%	47%	48%	49%
|Haasi							43%	41%	45%
Xóõ								85%	77%
Kakia									78%

This matrix, in accordance with the Starling-NJ method[1], yields the phylogenetic structure of Figure 1 (the latter is also accompanied with glottochronological dates, calculated according to the Swadesh / Starostin method).

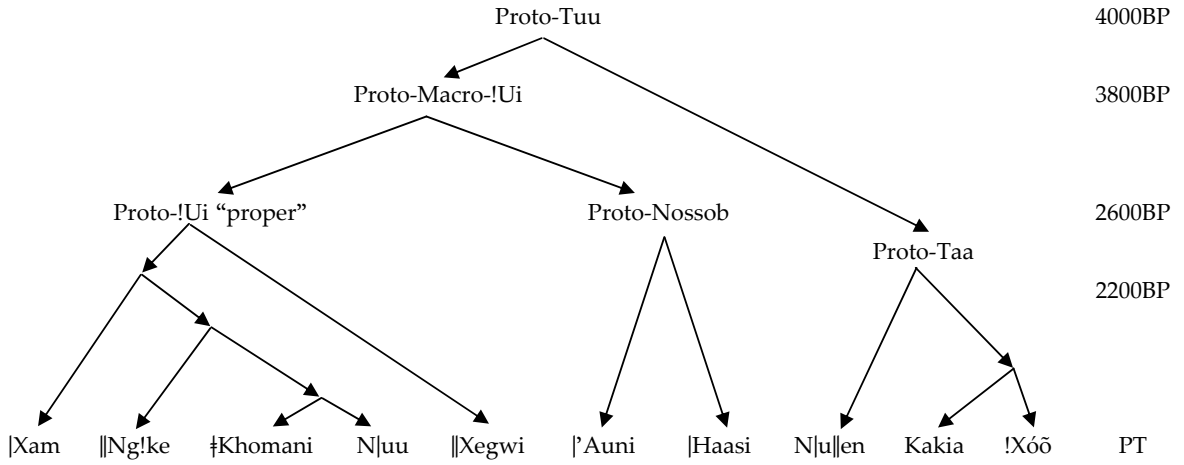

Figure 1. "Maximalist" tree for Tuu languages (= all |'Auni-Nǁng matches are counted as inherited)

Before making any comments on the matrix or on the accompanying tree, let us now present the second, "minimalist" scheme, in which the abovementioned 9 highly likely or potential borrowings from Nǁng into |'Auni are removed from the calculations altogether, reducing the total number of |'Auni lexemes counted as matches or mis-matches to 68 (Table 3).

This matrix, in accordance with the Starling-NJ method, yields the phylogenetic structure of Figure 2.

It is immediately noticeable that the biggest — and, in fact, the only — change in the phylogenetic structure concerns the position of Nossob languages: on this scheme, they are actually seen as the first branch to split off the common Tuu stem, rapidly followed by Taa, whereas Figure 1 reverses the process, putting Taa as the first branch to diverge, almost immediately followed by Proto-Nossob.

[1] The Starling-NJ method is a simple clustering method for producing phylogenetic trees, introduced by Sergei Starostin and commonly employed in the Moscow school of comparative linguistics; see Kassian 2015 for a detailed description.

Table 3. "Minimalist" lexicostatistical matrix between Tuu languages
(= 9 |'Auni-N‖ng matches are considered to be borrowings and removed from calculations)

	‖Ng!ke	ǂKhomani	N|uu	‖Xegwi	|'Auni	|Haasi	!Xóõ	Kakia	N|u‖en
|Xam	74 %	79%	73%	59%	52%	43%	42%	40%	43%
‖Ng!ke		85%	85%	59%	55%	43%	43%	41%	43%
ǂKhomani			95%	74%	64%	53%	57%	56%	60%
N|uu				64%	53%	48%	47%	45%	48%
‖Xegwi					46%	40%	41%	41%	46%
|'Auni						74%	47%	47%	48%
|Haasi							43%	41%	45%
Xóõ								85%	77%
Kakia									78%

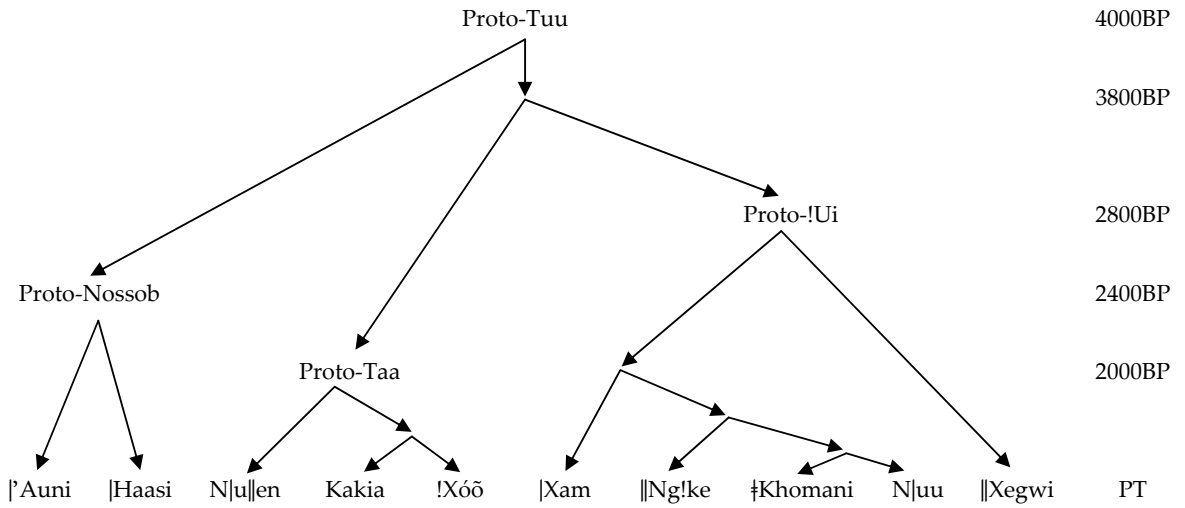

Figure 2. "Minimalist" tree for Tuu languages (= 9 |'Auni-N‖ng matches are removed from calculations).

Amusingly, neither of the two schemes agrees with Güldemann's classification, which would have Proto-!Ui as the first outlier. However, in all fairness, the glottochronological distance between the first two splits of Proto-Tuu on both schemes is so minuscule (approximately 200 years) that it lies well within the borders of statistical error. In such cases, the logical compromise is to postulate a trifurcation, reserving any definitive judgement on the exact chronological order in which it might have taken place, i.e. agree with the primary classification of Tuu as consisting of three more or less equidistant branches (Figure 3).

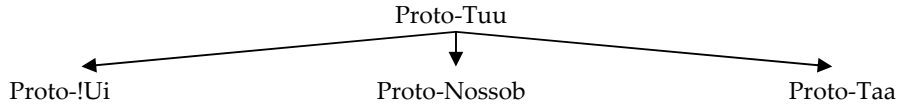

Figure 3. The likeliest evolution scenario for Proto-Tuu.

The veracity of such a scheme may be further confirmed or weakened by (a) running alternate formal methods, e.g. UPGMA or Bayesian phylogenetics, or (b) a manual analysis of

the individual isoglosses between the three branches, based on phonetic, morphological, semantic and distributional peculiarities of the compared items. Given the overall scarceness of the material, it seems far more sensible to me to prefer the latter approach. However, before proceeding to that stage, let us make some additional comments on the internal classification of all three branches.

A. *!Ui*. This is the most deeply divergent branch of all three, but largely due to the inclusion of ǁXegwi, which comes out as an obvious outlier on the !Ui tree (this agrees with Güldemann 2014: 258 listing it as an outlier). The divergence between |Xam and the entire Nǁng cluster is not so deep in comparison (approximately 2000 glotto-years according to the Swadesh-Starostin formula). Interestingly, "Modern N|uu" emerges as being much closer to the "ǂKhomani" doculects as recorded by Meingard and Doke rather than the "ǁNg!ke" as recorded by Bleek — although it has to be kept in mind that the wordlist for ǂKhomani is the most deficient one in my collection, and for that reason, match percentages between it and the other Tuu languages are inevitably skewed (cf. 59% between ǁXegwi and ǁNg!ke vs. 74% between ǁXegwi and ǂKhomani — the only explanation for such a blatant discrepancy is the severe incompleteness of the ǂKhomani list, creating the illusion of extra proximity).

Due to the impossibility of properly raising even half of the Swadesh list for the remaining attested !Ui languages (ǁKuǁe, ǁKxau, !Gã!ne, etc.), their place in this classification remains indeterminate; all that can be said with certainty is that they are probably closer to "Narrow !Ui" than to ǁXegwi, since I have not been able to discover even a single exclusive isogloss between any of them and ǁXegwi.

B. *Nossob*. As previously recognized by Güldemann, there is no reason to speak of |'Auni and |Haasi as dialects of a single language; despite all the obvious exclusive isoglosses between the two, their degree of divergence translates to almost 2,300-2,400 glotto-years (this is slightly more than for Turkic or Slavic languages, for instance). The position of "Xatia", or ǂ'Einkusi (another small dialect recorded by D. Bleek), in between |'Auni and |Haasi remains inconclusive due to critical scarceness of data.

C. *Taa*. Given the relative proximity of all attested variants of Taa, there is not much that standard 100-item wordlist-based lexicostatistics can say about their internal classification (Naumann 2014 offers a much more thorough perspective on the issue, albeit mainly based on grammatical and phonological, rather than lexical, isoglosses); however, given that at least Bleek's N|uǁen (a subdialect of Western !Xóõ) differs from Traill's Lone Tree !Xóõ (Eastern) by more than 20% of attested entries, it is clear that we are really dealing with at least several distinct *languages*; in practical terms, it means that cross-checking the lexical data of Lone Tree !Xóõ against words with the same meanings in other varieties of !Xóõ, whenever available, is a must in any historical studies of Tuu lexicology (as a particularly transparent example, cf. the situation with the 1st person pl. pronoun, where N|uǁen and Kakia agree with !Ui and Nossob languages, while the situation in !Xóõ is innovative).

Notes on the reconstructibility of the Proto-Tuu wordlist

The second most important task of the present study, in addition to constructing a phylogenetic tree for Tuu languages, was to assess the possibility of reconstructing a Swadesh wordlist for Proto-Tuu — a task equally important for the internal historical study of Tuu languages and for solving the problem of their external connections. To somewhat formalize the procedure, as well as reduce the risk of errors resulting from "homoplastic" developments (the same protolanguage item independently shifting its semantic properties to the same Swadesh

meaning in two or more distantly related languages) and make more transparent the processes of lexical replacement and semantic shift in the various Taa subgroups, we have advocated for a strict step-by-step approach, reconstructing first (whenever possible) the respective Proto-!Ui, Proto-Nossob, and Proto-Taa entries, then determining which of these can be genetically related on the basis of established phonetic correspondences (or, at the very least, regular correspondence patterns whose phonological status on the Proto-Tuu level still remains unclear, but whose recurrence can be reliably demonstrated).

Details of the reconstruction process have already been laid out in the first part of the present paper; here, I shall simply restate that a certain item is deemed to be *reconstructible* for Proto-Tuu if its reflexes are found in at least two out of three primary branches (if we accept the trifurcation scheme as the most likely one, this means that !Ui-Nossob, Taa-Nossob, and !Ui-Taa isoglosses all have comparably high chances of reflecting the situation in Proto-Tuu). Also, it is important to state that the intermediate reconstructions can technically be *pseudo-reconstructions* (i.e. a word whose regular reflexes are only found in one !Ui language and in one Taa language can perfectly well reflect the corresponding Swadesh item in Proto-Taa), but only if they do not come into conflict with actual reconstructions (i.e. if there is a "better" candidate for the Swadesh meaning, in terms of its distribution across daughter languages, it is given preference over such "minor" isoglosses). However, to the best of my knowledge, such conflicting situations are extremely rare in Taa languages.

Based on various degrees of their reconstructibility, the entries on Swadesh's 100-item wordlist can be divided into the following groups.

1. *Not reconstructible even on (most or all of) the intermediate levels*: 'bark', 'belly', 'burn', 'cloud', 'feather', 'fish', 'good', 'green', 'leaf', 'root', 'round', 'seed', 'swim', 'warm', 'white', 'yellow'.

Some of these words have to be excluded simply based on the lack of corresponding realities (e.g. 'fish', 'swim') or on their highly specialized representation in the South African environment (e.g. most of the flora-related terms, such as 'leaf', 'root', 'seed'). Others simply turn out to be highly unstable, including several adjectives ('good', 'round', 'warm') and, interestingly enough, the majority of color terms included in the Swadesh wordlist — thus, for Proto-!Ui only 'black' and 'red' can be reconstructed with a certain degree of reliability.

2. *Reconstructible on some or even all of the intermediate levels, but not reconstructible for Proto-Tuu due to lack of cognacy between the individual branches*: 'black', 'cold', 'earth', 'egg', 'foot', 'full', 'give', 'kill', 'knee', 'man', 'mountain', 'new', 'person', 'rain', 'red', 'sand', 'say', 'small', 'stand', 'star', 'stone', 'that', 'two', 'who'. It is not always clear which of these words are simply lexicostatistical mismatches (i.e. cognates exist in other branches but underwent semantic shift) and which ones are etymological mismatches (i.e. cognates simply do not exist) — the solution of this issue has to be postponed until the preparation of a full-fledged etymological corpus for Tuu languages.

3. *Reconstructible (or at least "pseudo-reconstructible", i.e. present in at least one language) for all three subgroups and cognate with each other*: 'bird', 'bite', 'claw /nail/', 'come', 'die', 'dog', 'drink', 'ear', 'eat', 'fat', 'fire', (?) 'fly', 'hair', 'hand', 'heart', 'horn', 'I', 'lie', 'meat', 'name', 'nose', 'one', 'see', 'sit', 'sleep', 'thou', 'tongue', 'tooth', 'tree', 'water', 'we', 'woman'. ('Fly' is debatable, see point 7a above where it is argued that the word may be an areal inter-family isogloss with a sound-symbolic stamp on it).

This is the single largest group of them all, providing the strongest evidence for a genetic connection between all three subbranches. It is pleasing to note that no fewer than 23 out of 32 of its constituents belong to the "ultra-stable" 50-item wordlist, suggested in Starostin 2010 and since then used as the basis for deep-level linguistic comparison in my own studies on African languages.

4. *Reconstructible based on isoglosses between !Ui and Taa*: 'all', 'ashes', (?) 'big' (problematic because of phonetic difficulties), 'breast', 'dry', 'head', 'not', 'skin', 'sun'; add 'bone', 'liver', 'louse' — three items not attested in Nossob languages at all (the other nine items have etymologically different equivalents in the Nossob branch).

5. *Reconstructible based on isoglosses between !Ui and Nossob*: 'blood', 'eye', 'night', 'smoke', 'this', 'walk /go/'; add 'hear', 'long', 'neck', 'road' if the Nossob equivalents are inherited rather than borrowed (if we decide to count them as borrowings after all, all four have to be moved to group 2).

6. *Reconstructible based on isoglosses between Taa and Nossob*: 'know', 'many', 'moon', 'mouth', 'tail', 'what'. Amusingly, if we discard the four potential borrowings from group 5, the number of exclusive !Ui-Nossob isoglosses on the Swadesh list is precisely the same as the number of Taa-Nossob isoglosses, further confirming the trifurcate division.

Although we are still a long way away from producing reliable Proto-Tuu reconstructions based on well-verified phonetic correspondences for all the constituting segments, certain useful observations on their phonetic features can be made even now. Thus, out of the 54 items reconstructible for Proto-Tuu 38 begin with click consonants (4 with labial **ʘ-*, 14 with dental **|-*, 3 with alveolar **!-*, 3 with palatal **ǂ-*, 12 with lateral **‖-*, and 2 with the enigmatic "sixth click") and 16 begin with egressive consonants (6 with alveolar sibilants / affricates, 3 with coronal stops, 2 with velar stops or affricates, 1 with nasals, 4 with initial vowels or glottal stop).

While the statistical results for non-click consonants are relatively predictable and in general agreement with the click-to-non-click rates in such languages as N|uu and !Xóõ, the distribution of different click influxes presents a big surprise, with words beginning with **|-* and **‖-* being 3 to 4 times more frequent than those beginning with **!-* or **ǂ-*. Even if we ultimately conclude that evidence for the "sixth click" is unconvincing and decide to merge it with either one of those two, the frequency rates will change only slightly. This is markedly different from, e.g., the situation in !Xóõ, where, according to Traill's dictionary, approximately 330 lexical roots have the dental click *|*, approximately 430 have the alveolar click *!*, approximately 450 have the lateral click *‖*, and approximately 350 have the palatal click *ǂ* (the numbers may be slightly different depending on which items are considered to have the same lexical root and which ones are not).

It is instructive to compare these statistics with the distribution of clicks across 42 common Tuu stems containing clicks which were tentatively reconstructed by T. Güldemann (2005: 24–28) for a general study showing the overall genetic relationship of the Tuu languages. His sample, also rooted primarily in basic lexicon but not restricting itself to Swadesh items, was an attempt to put together the most stable and widely distributed etyma, which would make any resulting statistics run on them worthy of attention. They are as follows: 3 items with labial **ʘ*, 20 items with dental **|*, 6 items with alveolar **!*, 4 items with palatal **ǂ*, 9 items with lateral **‖*. Again, we see a huge discrepancy when it comes to the dental click, although the figure for the lateral click is slightly lower than in my case (however, I would definitely change Güldemann's **!ab-* 'leopard' to a reconstruction with **‖-*, which would bring the distribution slightly closer to the one in my sample).

From a purely theoretical standpoint, this discrepancy, if its purely accidental origins are to be ruled out, can only be explained by one of two possible reasons: (a) undetected processes of click loss or click articulation shift in !Ui and Nossob languages, during which many items with alveolar and palatal clicks underwent significant phonetic change; (b) conversely, mass borrowing of lexical items from other languages into Taa (from Khoe or maybe even other, now extinct, "Khoisan" lineages) which led to an increase of items with formerly rare click influxes. Unfortunately, concrete evidence for both of these speculative hypotheses is so far lack-

ing (there are, of course, plenty of Khoe borrowings in Taa, but nowhere near enough to breach such an enormous gap). Without a doubt, some light on this problem shall be shed in the process of putting together a robust etymological corpus for Tuu, but for now this puzzle remains unresolved.

An additional look at the Nossob language issue: shared features and innovations

It would probably make sense to suggest that where pure statistics fails to properly resolve an issue (such as produce a robust binary-split tree rather than have us accept a trifurcation model), detailed manual analysis of the underlying data may potentially reveal important classificatory arguments that are inevitably lost in the cogs of crude automated algorithms. For instance, it may be important to not merely look at the *quantitative* aspects of !Ui-Nossob-Taa comparison, but also take into account the *quality* of detected matches — it is for this reason that I attempt, whenever possible, in the Appendix sections of both these papers to indicate not just the fact of cognacy between different items, but also the *degree* of cognacy, recognizing that, for instance, a certain Proto-Nossob item may be related to both Proto-!Ui and Proto-Taa, but be phonetically or morphologically closer to one of these branches rather than the other. Likewise, it would make sense to try and assess all the detected binary isoglosses (!Ui-Nossob vs. Taa, Taa-Nossob vs. !Ui) in terms of shared innovations vs. shared archaisms — this is extremely difficult in light of the overall poor state of Tuu etymology, but even a small handful of transparently resolvable cases might be quite helpful.

Below I list, in alphabetical order, all the Swadesh items which satisfy the following conditions: (a) they are reconstructible for Proto-Nossob or are at least found in one of two Nossob languages; (b) they are not suspected, on the whole, of being borrowed into |'Auni from N‖ng; (c) they have reliable cognates in either !Ui, or Taa, or both of these branches; (d) they *either* have phonological or morphological features that bring them closer to !Ui or Taa, *or* they can be regarded as shared lexical innovations (rather than retentions) with one of these two branches.

1) BIRD: Proto-Nossob **si=ʗu* has the same desemanticized prefix **si-* that is frequently found in numerous dialects of Taa. Additionally, the word has a voiced efflux in common with Proto-Taa **ʗu(ˀ)-*, rather than the uvular efflux in Proto-!Ui **/qʰu-*.

2) BITE: Proto-!Ui **c'i* and Proto-Nossob **c'i* vs. Proto-Taa **siʔ(i)*. This could be qualified as a shared phonetic innovation, since it is easier to explain **siʔi → *sʔi → *c'i* than the opposite development in Taa. However, it is not very diagnostic since we also know cases where glottalic articulation of the affricate in |Haasi, as marked by R. Story, is most likely secondary (e.g. *c'i* 'to come', which has no cognates with glottalized affricates in any other Tuu languages); therefore, Proto-Nossob **c'i* could simply reflect **si-* or (*siʔ-*), and its extra similarity to Proto-!Ui could be an illusion.

3) DOG: both Proto-!Ui and Proto-Nossob feature a nasal suffix at the end of this stem (**ǂʰu-ni ~ *ǂʰu-iŋ* vs. **ǂʰɔ-ŋ*) respectively as opposed to Proto-Taa **ǂqʰa-i* without any traces of nasality. Given the existence of such !Ui variants as ‖Xegwi *ƀwe* ← **ǂ/ʰu-ai* (?), it is quite possible that different morphological variants were already present in Proto-Tuu, and it is hard to determine the relative degrees of archaism and innovation in this case.

4) EYE: even if we discard |'Auni *ca-xu* as a recent re-borrowing from N‖ng, |Haasi *cxɔ* 'eye' still transparently shows that Nossob languages have a strong isogloss with !Ui **c'a-xu*. In the first part of the paper (Starostin 2021: 123) it was already shown how this situation is better explained as a shared innovation for both groups rather than retention of a common archaism.

5) FAT: Nossob **so-a* or **so-ã* is morphologically closer to Taa **sãˀ* than to !Ui **so-ni ~ *so-iŋ*.

6) FIRE: !Ui and Nossob have **|'i* vs. Taa **|'ã*. It has been previously argued that **i* is more likely to reflect the original root vocalism, with Taa **|'ã* ← **|'i-ã*, thus, the similarity between !Ui and Nossob is probably a shared retention rather than innovation.

7) LIE: |'Auni *tòa* is closer to Proto-Taa **tu* than than Proto-!Ui **ta* (although it is likely that the base root underlying all stem variants is ultimately the same).

8) NAME: this stem's codas in !Ui (**|e* ~ **|ẽ*) and Nossob (**|eN*) are notably closer to each other than to Taa **|ã-ũ*), but it is not clear which of the two full stems is more archaic.

9) ONE: reflexes of this word, featuring the "sixth click", are notably closer between Nossob (**ǂ'u-ŋ*) and Taa (**ǂ'u-*) than between either of them and !Ui (**ɕoaʔ-/i/*). It must be noted that the very fact of the "sixth click" being consistently reflected in Nossob and Taa as palatal *ǂ* (opposed to alveolar and lateral reflexes in !Ui languages) is in itself an important phonetic isogloss, more closely binding Nossob and Taa to each other.

10) TOOTH: phonetically and morphologically, !Ui **ǁʰãĩ* is closer to Nossob **ǁe-[iŋ]* than Taa **ǁqʰã* ~ **ǁqʰan* (no signs of front vocalism in Taa codas).

11) WOMAN: paradigmatically, this stem behaves more similarly in !Ui and Nossob languages than in Taa, although this is likely to be a shared archaism (since Taa sg. **ǂâ-qáe* 'woman', lit. 'person-mother', is a transparent lexical innovation).

The list turns out to be a little more skewed in the direction of !Ui-Nossob proximity (7 cases vs. 4 cases of closer similarities between Taa and Nossob), but only in two of these cases is it possible at all to make a well-argued (though still far from waterproof) judgement about the direction of innovation, and they are, ironically, equally divided between !Ui-Nossob ('eye') and Taa-Nossob ('one'). The other lexical isoglosses between !Ui and Nossob, on one hand, and Taa and Nossob, on the other (listed under p. 5 and 6 of the previous section), also do not allow to make serious judgements about the directions of lexical replacement. In the end, it turns out that even such fine-grained manual analysis of the evidence does not allow us to properly depart from the trifurcate model of Tuu phylogeny.

One last thing that remains is to critically evaluate the specific arguments given by T. Güldemann in favor of a closer connection between Taa and Nossob and see if they are sufficient to go back to the binary-split classification and admit that the lexicostatistical data simply fail to properly reflect the true historical process. The main argumentation, laid down in Güldemann 2014, consists of two points.

1. Güldemann notes a closer affinity between the overall numeric systems of !Xóõ (*ǂ'u-* 'one', *ǂ̃Ṽ-* 'two', *ǁâe* 'three', *ǁáli* 'many') and |'Auni (*ǂ'ũ-* 'one', *ǁ̃a* ~ *ǁ̃a* 'two', *ǁ'ai-* 'three', *ǁáni* 'many'). There are, however, multiple problems with this argument. First, |Haasi — |'Auni's closest relative — seems to have only one common element with |'Auni in this paradigm (*ǂŋ́-* 'one'). Second, it remains unclear if Bleek's transcriptional variation *ǁ̃a* ~ *ǁ̃a* can really be interpreted as misheard variants of *ǂ̃a* (as suggested by Güldemann in order to justify the comparison with !Xóõ *ǂ̃Ṽ-*); note that, as a rule, Bleek transcribes the palatal click in |'Auni correctly (see DOG, MOUTH, ONE for examples).

Third, the very "numerals" *ǁ̃a* ~ *ǁ̃a* and *ǁ'ai-* are actually defined by Bleek not as numerals, but as special "particles" that precede the actual numeric stems, as in *ǁŋ ǁ̃a tis |am* 'two huts', *ǁn ǁ'ais !̃wɔna-a* 'three huts', where the identifiable numerals are *|am* 'two', *!̃wɔna* 'three' (both of them transparent borrowings from Khoe). Güldemann suggests that these morphemes may be behaving like grammaticalized markers (e.g. for dual or plural), to which language speakers are then adding actual numerals borrowed from a different language, but this is hard to prove based on the very limited number of examples in Bleek's texts.

Furthermore, even if Güldemann is correct in all of his hypotheses here, and the available evidence truly allows us to reconstruct a four-element numeric paradigm for the common an-

cestor of Taa and Nossob, it would still be impossible to present it as a shared innovation between these two groups, rather than an archaic retention from Proto-Tuu; to do this, one would need to demonstrate, at the very least, that Proto-!Ui **!'u* 'two' is more archaic in this shape and meaning than Nossob-Taa **ǂ̃V-* id. ('three' is also borrowed from Khoe across all !Ui languages, and 'many' is not properly reconstructible). Therefore, the numeric argument is at worst etymologically dubious, and at best inconclusive.

2. The second argument ultimately has more to do with typology than reconstruction: analysis of the small preserved syntactic corpus for |Haasi, as well as a few textual examples from D. Bleek's records of |'Auni, allows Güldemann to conclude that Nossob languages had an active system of grammatical agreement and lexical gender closely resembling the one described for Taa, but not attested in the !Ui language complex (or, for that matter, anywhere else in the Khoisan-speaking region). Moreover, some, if not all, of the identifiable concord markers (allowing to roughly distribute all the nominal stems of |Haasi into three different classes marked as I/E, A, and U respectively) have direct counterparts in !Xóõ.

While this may be a very strong argument in favor of reconstructing such a system for Proto-Tuu (and assuming, for the moment, that it was lost in Proto-!Ui), it is not necessarily indicative of a specific bond between Taa and Nossob languages as such. One problem is that the actual distribution of nouns across classes in |Haasi rarely matches the same distribution in !Xóõ. Furthermore, the system of markers matches Taa only partially (it is unclear, for instance, if |Haasi *A* can be equated with Taa *Ã*, for which nasalization is essential). Most importantly, perhaps, in !Xóõ class-defining morphemes are frequently seen in conjunction with the roots (e.g. Class 3 nouns often end by themselves in *-e*, which also functions as the agreement marker, etc.), and while comparative data show clearly that some of the codas in |Haasi nouns are old fossilized suffixes, they are rarely the same as in !Xóõ.

Thus, even if a word like 'dog' belongs to the same I-Class in |Haasi and !Xóõ, more important is the discrepancy in the actual shape of the stem 'dog', which is *ǂqʰà-i* in !Xóõ, but *ǂʰaŋ* in |Haasi (closer in its morphological constitution to the respective forms in !Ui, actually). As has been shown above, cases where Nossob languages show "Taa-like" morphology are statistically more or less the same as cases in which they show "!Ui-like" morphology, which once again brings us back to the trifurcate model.

Conclusions

Perhaps the most important conclusion from this study should be the realization that it is still too early to draw any definitive conclusions on the phonology and lexical constitution of Proto-Tuu, as well as the internal classification of its members — many of the judgements, assumptions, and reconstructions in this paper will yet be subject to further amendment as more data from old sources (such as the slowly emerging notebooks of D. Bleek and other researchers) and new sources (such as contemporary research on surviving Taa dialects) come to light.

Nevertheless, even the limited amount of data taken into consideration here allows to draw up a reasonably realistic picture of the evolution from Proto-Tuu to its descendants. The most important details of that picture are as follows:

(a) Proto-Tuu is, indeed, the deepest and oldest of all known "Khoisan" lineages (with the possible exception of Ju-ǂHoan, a.k.a "Kx'a"), glottochronologically dated to approximately 4000BP; this explains many of the difficulties with reconstructing a Proto-Tuu Swadesh list due to cognate loss;

(b) the disintegration of Proto-Tuu must have happened according to a trifurcate pattern, with speakers of Proto-!Ui, Proto-Nossob, and Proto-Taa parting ways at more or less the same time. Even if there was a brief period of "!Ui-Nossob" or "Taa-Nossob" unity, it did not last long enough to produce a statistically significant number of shared innovations;

(c) Proto-Tuu must have had a relatively complex, but not particularly stable, system of nominal morphology, as is seen by the numerous stem variants for the same lexical roots witnessed across (and sometimes even within) the individual branches. Reconstruction of this system is severely hampered by lack of data and inadequate grammatical descriptions for extinct languages;

(d) most curiously, distribution of click phonemes across basic lexicon morphemes reconstructible for Proto-Tuu significantly differs from usually attested distributions of click phonemes across attested languages, such as !Xóõ. This indirectly hints either at huge, hitherto not comprehensively described or understood areal influence on Tuu languages after the family's disintegration, or at our inadequate understanding of the historical development of Tuu phonology — or, possibly, both.

It is not likely that anything other than a comprehensive, systemic investigation of the etymological connections between the various Tuu languages, hopefully resulting in a proper comparative dictionary following the classic Neogrammarian principles, will allow us to move beyond these observations, which can only, for now, be used as general guidelines for further research. Whether such an endeavour will be at all possible in the near future, remains to be seen.

Notes on transcription

The transcriptional system used in this paper generally follows the transcriptional standard which is currently employed in the Global Lexicostatistical Database and is itself essentially based on IPA, but with a few important modifications.

(1) Clicks: following the system adopted in Vossen 1997, nasalized clicks are transcribed with a superscript tilde sign (*ʘ̃*, *ǀ̃*, etc.) while voiced clicks have a subscript tilde (*ʘ̬*, *ǀ̬*, etc.).

(2) Affricates: instead of IPA's digraphic combinations, single letters are used to denote alveolar (*c*, *ʒ* = IPA *ts*, *dz*) and palatal (*ɕ*, *ʑ* = IPA *tɕ*, *dʑ*) affricates.

(3) For morphological segmentation, the hyphen sign is used to separate root morphemes from suffixes (*ku-ka*, etc.), while the equation sign is used to separate roots from prefixal components (e.g. |'Auni *si=|u* 'bird', etc.).

For a more detailed description of the transcription system, including notes on transliteration of data from old sources, see Starostin 2015.

Appendix. Comparative analysis of Tuu basic lexicon (Items 51–100).

In this Appendix, I list the results of intermediate and Proto-Tuu reconstructions for the (alphabetically) second half of the Swadesh wordlist (more or less closely following the semantic specifications set out in Kassian et al. 2010). Structure of the entries follows the same guidelines as the first part, reprinted here:

(1) Name of the item, together with a formal notation of the presence / absence of lexicostatistical parallels between the three branches: e.g. [!Ui + Taa] [- Nossob] means that the reconstructions for Proto-!Ui and Proto-Taa are cognate, whereas the reconstruction for Proto-Nossob is not (this also includes pseudo-reconstructions). Sometimes, even when all three

branches reflect the same root, two out of three may be more tightly connected, for instance, sharing common morphological formations (suffixes, etc.). Such extra proximity is indicated with additional parentheses, e.g. [!Ui + [Nossob + Taa]]: it offers additional evidence for phylogenetic classification. If there are no matches whatsoever between any of the three branches, the word is marked with [-].

(2-4) Reconstructions for Proto-!Ui, Proto-Nossob, and Proto-Taa, accompanied with a list of most of the attested reflexes. If the onomasiological reconstruction is equivocal, two or more roots may be listed instead as (a), (b), etc. The ◊ sign separates listed data from comments on the reconstructions. Note that the Appendix does not necessarily list *all* the attested forms corresponding to the Swadesh items in question, but mainly those that justify the reconstruction. For complete lexicostatistical lists, the reader is advised to refer to the South Khoisan (!Ui and Taa) databases that are separately available online at the Global Lexicostatistical Database (Starostin 2011–2021).

(5) Proto-Tuu reconstruction (where it is at all possible). For reasons described above (in the "Notes on phonetic reconstruction" section), we do not systematically list Tuu protoforms, but rather use the notation "Tuu+" to indicate credible lexicostatistical isoglosses between !Ui and/or Nossob and Taa which almost certainly go back to a common Tuu protoform, and the notation "Tuu–" to indicate the lack of such isoglosses. Note that "Tuu–" also marks situations where one of the branches may have an etymological cognate in the other, but since the meanings are different, this does not qualify as a proper lexicostatistical match (e.g. BIG, etc.).

All data sources remain exactly the same as in the first part of the paper — Bleek & Lloyd 1911, Bleek 1929, Bleek 1956 for |Xam; Bleek 1956, 2000 for ǁNg!ke; Doke 1936, Maingard 1937 for the "ǂKhomani" doculects of N‖ng (ǂKho-D, ǂKho-M respectively); Miller et al. 2009, Collins & Namaseb 2011 and Sands p.c. for Modern N|uu; Ziervogel 1955 and Lanham & Hallowes 1956a, 1956b for ǁXegwi (ǁXegwi-Z and ǁXegwi-LH respectively); Bleek 1937, 1956 for |'Auni; Story 1999 for |Haasi; Traill 1994 for Lone Tree !Xóõ; Bleek 1929, 1956 for "Kakia" Taa and N|u‖en.

51. MAN [-]

- !Ui: ***ǂo ~ *ǂ'o** (ǁNg!ke *ǯo:* 'male', ǂKho-M, N|uu *ǂo:*, ǁXegwi-Z *ʎo*, ǁXegwi-LH *kwi-ʎ'o:*). ◊ Bleek's ǁNg!ke data as well as ǁXegwi samples collected by Lanham and Hallowes indicate that the primary meaning of this morpheme must have been 'male', since the meaning 'man' is actually expressed by a compound whose first part (*kwi*) means simply PERSON q.v., and it is also encountered in other compounds denoting male animals, e.g. *širi-ʎ'o:* 'male buck'. More problematic is the observed variation in click efflux articulation: the entire N‖ng cluster as well as Ziervogel's ǁXegwi shows **ǂo* (zero efflux), whereas Lanham and Hallowes record an ejective lateral affricate which reflects original **ǂ'o* (less indicative is Bleek's ǁKu‖e form *t'o* 'man', since ǁKu‖e *t'-* is also substituted for simple **ǂ-* in at least one other case, see MOON below). The "majority rule" suggests regarding **ǂo* as the original form, although it is not excluded that the variation actually reflects a separate efflux[2].

  In |Xam, the most common equivalent for 'man' is the compound form *!wi=gwai*, where *!wi* = PERSON q.v. and *gwai* means 'male'. It is unlikely that *gwai* is related to **ǂo*,

[2] It is interesting to note the existence of phonetically similar roots with the meaning 'male' in Proto-Ju (**ǯǂo*, with a voiced retroflex click) and in Khoe (**ǁo* 'male' → Nama *ǁgō-b* 'bull; tough person', Naro *ǁô* 'bull, male'): whether they may be used as evidence for more distant relationship with Tuu remains to be seen, but it should be noted that neither of the two shows a glottalized click efflux.

since random click loss is not typical of |Xam, but no other etymology can be suggested at the moment.

Interestingly enough, the plural form 'men' is suppletive in most !Ui languages, being formed from the common !Ui root ***tu**: |Xam *tu-kən*, ǁNg!ke *tu-kən ~ tu-ŋən*, N|uu *ɕu:-ke*. The only doculect in which *tu* is occasionally found in singular use is Bleek's ǁNg!ke, cf. *ŋ tu e ǁŋ a* "that man is at the hut" (Bleek 1956: 240), but since no other recorded dialect of N‖ng reflects such usage, and given that the strong association between **tu* and plurality persists into Nossob and Taa languages as well (see notes below and on PERSON), one should rather suggest either a misglossing on Bleek's part or a special back-formation in one or more old dialects of N‖ng.

- Nossob: (?) ***be** (|'Auni *be ~ bɛ*, |Haasi *bi:*). ◊ In addition to this form, common for both |'Auni and |Haasi, Bleek also lists *da ~ de* 'man, person' for |'Auni, without any clear semantic distinctions (the default word for PERSON is more likely to have been *ǂ'i ~ *ǂ'e*, see below). Archaic origin for Common Nossob **be* is dubious for phonetic reasons, since labial *b-* is not well reconstructible as an inherited phoneme for Proto-!Ui; it is quite possible that **dV* is the original word for 'man', still retained in some functions in |'Auni but already replaced in the most basic usage on the Common Nossob language by an innovation. Additionally, the suppletive plural form 'men' in |'Auni is either *tu-ke* or *tu-tu-s(i)*[3]; the first variant may be borrowed from N‖ng, but the second is more likely to be inherited.
- Taa: ***ǂa=aˢ** (!Xóõ *tâ:=à:ˢ*, Kakia *laˢ ~ la:ˢ ~ la:*). ◊ For !Xóõ, cf. also Westphal's data: ǂHūa *laʔa*, N|amani *táˢʔaˢ*, pl. *táˢʔaˢ-tu* (Westphal 1965: 139). The word is clearly a compound formation, where **ǂa* = 'person' (see notes on PERSON for issues of phonetic reconstruction) and **aˢ* = 'father'. It is possible that a more archaic root for this meaning is preserved in the suppletive plural, cf. !Xóõ *ǁxà:* 'men' (ǂHūa *ŋ́ǁǂaá* in Westphal's transcription), or in a separate root represented by N|u‖en **!ã* 'man', pl. *!ã-te* (etymological connection between this form and !Xóõ *ǁxà:* is hardly possible). But the compound form is the only isogloss between two distant nodes of the entire branch.
- Tuu-: Formally not reconstructible, but possibly ***tu**? ◊ There is no clear evidence for a separate lexical root with the specific meaning of 'male human being' on the Proto-Tuu level; most frequently, we see this meaning expressed by various compound formations which do not match each other across different lineages (e.g. **!ui-ǂo* in !Ui vs. **ǂa-aˢ* in Taa). That said, given the fact that (a) the root **tu* is reconstructible at least on the Proto-|Xam-N‖ng level specifically in the plural meaning 'men' and (b) etymologically the same root **tu* is reconstructible for Proto-Taa in the plural meaning 'people' (see PERSON), it is a distinct possibility that **tu* may have denoted specifically the 'male human being' in Proto-Tuu (either in the plural only or irrespective of number), later becoming generalized to 'people' in Proto-Taa. However, the opposite scenario (**tu* as originally 'people' is also possible); see further notes on PERSON and WHO.

52. MANY [Nossob + Taa] [- !Ui]

- !Ui: Not reconstructible. ◊ The concept is extremely unstable in !Ui: most languages have their own equivalent, e.g. (a) |Xam */kx'wai:/-ya/*, (b) ǁNg!ke *ʗai ~ ʗãĩ ~ ʗẽ:*, (c) ǂKho-M *ɕebe-ɕe* = N|uu *kebe-ke*, (d) ǁXegwi-Z *kʰyũ ~ gyeĩŋ* = ǁXegwi-LH *q'iŋ* = ǁXegwi-B *ǁxain*.

[3] See Güldemann 2002: 189 on the detailed morphological analysis of this form, which he interprets as a combination of the root *tu* with an agreement marker and a copula.

- Nossob: (a) |'Auni *ǁáni ~ ǁáĩri*; (b) |Haasi *!ɔ:-ɔ:-k'a*. ◊ Formally not reconstructible for Proto-Nossob, but at least the |'Auni form has a perfect etymological parallel in Taa. The |Haasi form (also transcribed as *!o:o:-k'a* in one example) might, perhaps, have something to do with N|uu *!xo:* 'big' (although this would imply incorrect transcription of the efflux by R. Story).
- Taa: ***ǁa-ri** (!Xóõ *ǁáli*, Kakia *ǁari ~ ǁári*, N|u‖en *ǁan-te*). ◊ The root seems to be common for all varieties of Taa; Bleek's N|u‖en form shows that the stem is probably segmentable (with morphological variants **ǁa-ri* and **ǁa-n*).
- Tuu-: Not properly reconstructible. ◊ The apparent isogloss between Taa **ǁa-ri* and |'Auni *ǁani* may very well be etymological, and count as an argument in favor of specific proximity between these two branches; however, lack of any parallels in !Ui does not allow to formally postulate it for the Proto-Tuu level.

53. MEAT [!Ui + Nossob + Taa]

- !Ui: ***θoa-** (‖Ng!ke *θwai: ~ θwai*, ǂKho-M *θwoe*, N|uu *θoe*, ‖Xegwi-Z, ‖Xegwi-LH *θa:*). ◊ A stable lexical item in !Ui, except for |Xam, where the word was seemingly replaced by *ã: ~ eŋ*, a secondary nominalization of the verb 'to eat' (*ã*); the same replacement probably took place in ‖Kxau (*'a:ŋ*) and ‖Ku‖e (*õa-si*). Semantic derivation of 'meat' from 'to eat' is not uncommon in the Khoisan area (cf. Kalahari Khoe **kx'o-xu* 'meat' ← **kx'o* 'to eat /meat, hard food/'), and external data clearly show that **θoa-* is the original lexical item.
- Nossob: ***θoe** (|'Auni *θwe ~ θwi*, |Haasi *θwi:*).
- Taa: ***θa-** (!Xóõ *θàye*, Kakia *θwe*, N|u‖en *θwe: ~ θwi*).
- Tuu+: ***θ/o/a-**. ◊ An obvious isogloss between all three branches (vocalic reconstruction is uncertain, given the tendency of the labial click to labialize the following vowel in !Ui).

54. MOON [Nossob + Taa] [-!Ui]

- !Ui: ***ǂoro** (|Xam *!au!arro ~ !au!auru*, ‖Ng!ke *!orre ~ !urru ~ turro*, ǂKho-D *ǂ'ɔ̃rɔ́*, N|uu *ǂoro*, ‖Xegwi-LH *ǂolo*, ‖Xegwi-B *klolo*). ◊ Diphthongization in |Xam is likely to be secondary (cf. the same situation with NECK, NOT below); also of interest is the unique partial reduplication in this language (unless the first syllable is actually a different word, and the whole form is a compound formation rather than reduplication). Doke records a glottalic efflux for ǂKhomani, but it is not supported by data from other dialects of the N‖ng cluster.
- Nossob: ***!ʰõĩ** (|'Auni *!õĩ*, |Haasi *!ʰwi:*).
- Taa: ***!qʰan** (!Xóõ *!qʰàn*, Kakia *!xʌn*, N|u‖en *!xa:n*).
- Tuu-: Not properly reconstructible. ◊ !Ui and Taa etyma clearly have different origins and and no clear mutual parallels. However, the click onset of the Nossob word for MOON regularly corresponds to Taa (cf. DOG with a very similar correspondence), and coda differences may be explained by morphological variation (e.g. **!qʰo-i* vs. **!qʰo-an* → **!qʰan*). The case clearly counts as an etymological and lexicostatistical match.

55. MOUNTAIN [-]

- !Ui: (?) ***!ao**. ◊ In all data sources on |Xam and N‖ng, the word for MOUNTAIN (or 'hill') is the same as the word for STONE q.v. Only for ‖Xegwi the situation is different:

‖Xegwi-Z *tʰaŋ* (only quoted once in the phrase *tʰaŋ ʔe čwa* ‘black mountains’), ‖Xegwi-LH *ǂ̦u-l̃a ~ gu-l̃a* (a compound formation where the second component is probably HEAD q.v.). Both forms are somewhat problematic and have no etymological parallels in the rest of !Ui.

- Nossob: |’Auni *ǂ̦’wa*. ◊ In the phrase */’e /ʰui ge ǂ̦’wa* “we came down from the hill”. Not attested in |Haasi. Somewhat similar to the form in ‖Xegwi-LH, but the click effluxes do not match.
- Taa: ***!um** (!Xóõ *!ùʰm*, N|u‖en *!̰um*). ◊ The only equivalent in Kakia is *l̃u:n*, i.e. same word as STONE q.v.; however, !Xóõ and N|u‖en clearly have a separate lexical root for ‘mountain / hill’ (another attested meaning in !Xóõ is ‘niche for several sp. of plants, characterised by heavy sand’).
- Tuu-: Not reconstructible.

56. MOUTH [Nossob + Taa] [-!Ui]

- !Ui: ***tu** (|Xam *t:ú*, ‖Ng!ke *tu*, ǂKho-M *tu*, N|uu *ɸu:*, ‖Xegwi-LH *tu ~ t’u*). ◊ A common and stable !Ui etymon, sometimes with polysemy ‘mouth / hole’ (as in |Xam). Lanham and Hallowes note variation between *t-* and *t’-* in ‖Xegwi, but glottalic articulation is not confirmed by any of the attested external data.
- Nossob: (a) |’Auni *ǂ̦u:*; (b) |Haasi *n=ǂa*. ◊ For |’Auni, Bleek’s later records also show the form *tu ~ tʰu* ‘mouth’, noting explicitly that the form “may be ǂkhomani”. Since the earlier data, collected in 1911 and published in 1929, only gives *ǂ̦u:*, Bleek’s note is probably correct. No traces of **tu* ‘mouth’ are seen in |Haasi either; the form *nǂa* (where *n=*, as in similar body part terms, is really the pronominal possessive prefix ‘my’) is similar to |’Auni *ǂ̦u:*, but the vocalic correspondence is completely irregular and could only be resolved in morphological terms (e.g. **ǂ̦u-a → ǂa*), for which no corroborating evidence has been found so far. However, because of external parallels in Taa it may be safely assumed that the |’Auni form is more archaic and may be accepted in the status of a “pseudo-reconstruction”.
- Taa: ***ǂ̦u-** (!Xóõ *ǂ̦û-e*, pl. *ǂ̦û-m-sâ*, Kakia *!we:*, N|u‖en *ǂ̦ũẽ*). ◊ Kakia *!-* regularly reflects Proto-Taa **ǂ-* (either as a genuine phonetic development or as a regular mistranscription).
- Tuu-: Formally not reconstructible. ◊ !Ui and Taa roots are not related and have no mutual etymological parallels. However, |’Auni *ǂ̦u:* is clearly the same as the Taa form and thus, another important isogloss between the two branches.

57. NAME [[!Ui + Nossob] + Taa]

- !Ui: ***|e ~ *|ẽ** (|Xam */ẽ*, ‖Ng!ke */ẽ*, N|uu *ka=/ĩ*, ‖Kxau */ɛ:*, ‖Xegwi-LH */e:*). ◊ It is unclear if nasalization is an inherent part of the root or a remaining trace of a morphological marker (i.e. **/ẽ ← */e-/V/N*), but at least all forms are clearly related and a common !Ui ancestor is reconstructible. N|uu *ka=* is a prefix of inalienability.
- Nossob: ***|eN** (|’Auni */ẽ ~ /ẽn*, |Haasi *a=/aŋa*). ◊ The |Haasi form is also transcribed simply as *a=/ã* in the phrase “what is your name?”; it seems likely that *a=* in the vocabulary form is the 2nd p. pronominal prefix, while the second *-a* is a verbal copula. Vocalic discrepancy between |’Auni */ẽn* and |Haasi *=/aŋ-* is not easy to resolve in terms of common origin, and makes it worth considering the idea that the |’Auni form may actually have been borrowed from N|uu. On the other hand, there are a couple other cases where |’Auni *e* corresponds to |Haasi *a* (e.g. |’Auni */’e:* vs. |Haasi *=/a-* ‘heart’), so the contact scenario should not be regarded as the most probable by default.

- Taa: ***|ãũ**, pl. ***|ã** (!Xóõ */ãũ*, pl. */ã:*, Kakia */kx'ãũ*, N|u‖en */ã*). ◊ Judging by the paradigm in !Xóõ, the N|u‖en form recorded by D. Bleek was actually a plural one. Superficially, the plural form looks simpler than the singular, but its underlying morphophonological shape is actually **|ãũ-a* → **|ã* with regular contraction (exactly the same situation is observed in the case of NECK and TAIL, see below); *-a* is a frequent semi-productive plural morpheme in Taa, whereas no singulative marker like *-u* or *-ũ* can be postulated with any degree of certainty.
- Tuu+: All listed forms can be judged as cognates due to phonetic identity of the click consonant and basic root structure (**|VN*). However, reconstruction of the original quality of the root vowel is problematic; observed discrepancies cannot be explained by phonetic change alone and have to have a morphophonological explanation. The form in !Ui could be analyzed historically as **|ã-* (root) + **-i* (one of the singulative markers), which would also easily allow to equate it with the form in |Haasi. However, Taa **|ãũ* (rather than simply **|ã*) should then be symmetrically analyzed as ← **|ã-* + **-u*, despite fairly little evidence for **-u* as a separate class or number marker. Could there have been a regular development **-ãĩ* → **-ãũ* in Proto-Taa[4]? Unfortunately, there is too little comparative evidence for this sequence to make any strong conclusions. Still, this unresolved discrepancy should not be an obstacle for assuming common ancestry: vocalic correspondences between !Ui and Taa are a priori more complicated than consonantal ones.

58. NECK [!Ui + Nossob (?)] [- Taa]

- !Ui: ***ǂqu** (|Xam *!au ~ !eau ~ !ʰóu*, ‖Ng!ke *!ú ~ kú*, N|uu *ǂqu:*, ‖Kxau *ǂ'u*). ◊ A solid isogloss between |Xam and N‖ng (diphthongization in |Xam is probably secondary, as in many similar cases). Of interest is the clickless variant in ‖Ng!ke, marked by Bleek as an "occasional form". The ‖Xegwi equivalent is different: cf. ‖Xegwi-Z *ǀ'ele* vs. ‖Xegwi-LH *!eleŋ* (similar forms, but with incompatible clicks; I suspect a possible mistake in Ziervogel's records).
- Nossob: (a) |'Auni *‖u*; (b) |'Auni *ǂõĩ*. ◊ The word is not attested in |Haasi, whereas for |'Auni Bleek lists two equivalents, with only the former supported by a contextual example (*sa ko |kx'ẽsi, ho ha ‖ú ‖o* "bring beads, on my neck put them").
- Taa: ***ǂkx'ãũ** (!Xóõ *ǂkx'ãũ*, Kakia *!kx'üm*, N|u‖en *ǂũ*). ◊ Correspondences between !Xóõ and Bleek-transcribed Taa doculects are generally regular (the efflux *-kx'-* is often rendered as *-k-* /= zero/ in her N|u‖en records; *!-* regularly replaces **ǂ-* in Kakia); a minor problem is word-final *-m* in Kakia, but there are occasional other instances in which nasalization of labial vowels is rendered as a consonantal segment (e.g. *‖ũ ~ ‖um* STAND, see below).
- Tuu-: Not reconstructible. ◊ It is highly tempting to join !Ui **ǂqu* and Taa **ǂkx'ãũ* in a single etymology, but in order to do that, it would be necessary to at least demonstrate the recurrence of the click efflux correspondence between N|uu *-q-* and !Xóõ *-kx'-*, for which no fully convincing examples have been found so far. The two forms in |'Auni also present a challenge, but at least *ǂõĩ* is fully compatible with !Ui *ǂqu* under the reasonable assumption that **ǂõĩ* ← **ǂ[q]u-iŋ* (or **ǂ[q]u-ni*; cf. DOG for a similar case), where the second component is a formerly productive noun suffix.

[4] For a close-by typological parallel, cf. the curious phonetic variation between *-ãĩ ~ -ãũ ~ ã* type codas in Ju languages, e.g. Ju|'hoan *!ãĩʰ* 'tree' = !O!Kung *!ãũ* id. = Ekoka *!ãŋʰ* id.; Ju|'hoan *!ãĩ* 'neck' = !Kung *‖ãŋ ~ ‖ãũŋ* id., etc.

59. NEW [-]

- !Ui: Not reconstructible. ◊ This word is fairly well attested only in |Xam, where the most common equivalent for the meaning NEW (with polysemy: 'new / fresh / raw') is *ǂaːˢŋ*, pl. *ǂaˢǂaˢr̃r̃a* (with reduplication). Importantly, W. Bleek also records the form *ǂwe ~ ǂwẽ* in the same meaning, but textual examples show that it is always used in conjunction with 'moon', making it ineligible for lexicostatistical purposes. For ǁNg!ke, D. Bleek lists the adjectival stem *!xeː-kʸa ~ !xeː-tʸa* 'new / young', but for modern N|uu no equivalent has been recorded. The ǁXegwi-LH equivalent, attested in one phrase, is probably *ǂ'i*, with no external parallels.
- Nossob: Not attested in either |'Auni or |Haasi.
- Taa: ***ǂquˢ-** (!Xóõ *ǂquˢ-V*, Kakia *ǂxwe*). ◊ Not attested in N|u‖en. The isogloss between !Xóõ and Kakia is acceptable, even if Bleek's *-x-* for Kakia much more frequently correlates with !Xóõ aspirated *-qʰ-* rather than unaspirated uvular *-q-* (this could theoretically point to an original **ǂqʰuˢ-* rather than **ǂquˢ-*).
- Tuu+: Although there are no lexicostatistical matches between any of the three branches, it is important to note that Taa ***ǂquˢ-** corresponds rather well to |Xam *ǂwe* 'new /of moon/' (other than the lack of pharyngeal articulation in |Xam); this could imply a narrowing of the original meaning in |Xam and, consequently, a Proto-Tuu origin for the Taa stem (admittedly, this is all highly circumstantial evidence, and the conclusion is liable to change if more data are accumulated).

60. NIGHT [!Ui + Nossob] [- Taa]

- !Ui: ***ǂ̰a** (|Xam *ǂa ~ ǂaː*, ǁNg!ke *ǂa ~ ǂaː*, ǂKho-D *ǂ'āː ~ ǂàʔà*, N|uu *ǂaː*, ǁKxau *ǂaː*, ǁKu‖e *ǂa*, ǁXegwi-LH *ǂaː*). ◊ A highly stable and phonetically transparent form, although some fluctuations in click efflux articulation are observable; perhaps the ǂKho-D bisyllabic form *ǂàʔà* may be seen as a clue, although we prefer for now to follow the majority rule in our reconstruction.
- Nossob: ***ǂ̰a-** (|'Auni *ǂàu ~ ǂò*, |Haasi *ǂa-ǂa*). ◊ The diphthong in |'Auni finds no confirmation in any other sources; it has to be understood as the result of morphological contraction (← **ǂa-u*). The reasons for reduplication in |Haasi are also unclear.
- Taa: ***t̃uˢ- (**!Xóõ *t̃úeˢ*, pl. *t̃úˢ-m-tê*, Kakia *t̃òeˢ*, N|u‖en *t̃òe ~ t̃we*). ◊ A common and stable Taa root, reliably reconstructible for the proto-level.
- Tuu-: Not reconstructible. ◊ The isogloss between !Ui and Nossob is quite clear, but the Taa root is completely different, and !Ui-Nossob **ǂa* vs. Taa **t̃uˢ-* find no mutual etymological parallels.

61. NOSE [!Ui + Nossob + Taa]

- !Ui: ***t̃u** (|Xam *t̃ũ-nu*, ǁNg!ke *t̃u-tu*, ǂKho-M *t̃u-tu*, ǁKxau *t̃ú-tú*, ǁKu‖e *t̃u-tu*, N|uu *t̃u-ɕu*, ǁXegwi-Z, ǁXegwi-LH *t̃u*). ◊ The pure root is clearly preserved in ǁXegwi; most other languages reflect the extended stem **t̃u-tu*, where **-tu* is a special "anatomical" suffix, most likely derived from **tu* 'hole, mouth'. Only |Xam has a different morphological extension, although the form with **-tu* is still seen in the plural variant (*t̃ũ-t̃ũ̋ː-tu* 'noses', with root reduplication).
- Nossob: ***t̃u** (|'Auni *t̃õ ~ t̃uː*, |Haasi *t̃u*). ◊ In |'Auni, cf. also *t̃oi-tu-ke* 'nostrils'.
- Taa: **t̃uʰ-** (!Xóõ *t̃ùʰ-ɲa*, Kakia *t̃u/-ča/*, N|u‖en *t̃u-ša*). ◊ As in !Ui, the root avoids being used in pure form (although for Kakia, Bleek claims the possibility of such usage); instead,

the two most commonly encountered morphological variants are **ǀuʰ-ɲa* and **ǀuʰ-sa* (both morphemes are nominal suffixes with unclear functions).

- Tuu+: This is one of the most stable and phonetically transparent roots in the entire wordlist, clearly reconstructible as ***ǀu-** (or ***ǀuʰ-**, if breathy articulation of the vowel in !Xóõ is archaic). Of note, however, are the different models of its morphological framing in various branches; only the Nossob languages seem to prefer usage of the pure, unextended root stem.

62. NOT [!Ui + Taa] [- Nossob]

- !Ui: (?) ***‖V-** (‖Ng!ke *‖u ~ ‖ɔ́ ~ ‖e*, ǂKho-M *‖o ~ ‖e*, N|uu *‖u*, Seroa *‖au*). ◊ Basic expressions of verbal negation in !Ui languages differ highly from language to language. The only one that finds reliable etymological parallels outside of !Ui, and, consequently, the only candidate for the status of a "pseudo-reconstruction", is the morpheme reconstructible as **‖V-* on the basis of N‖ng data. According to Collins & Namaseb 2011: 10, the negation subsystem in N|uu consists of three morphemes: *‖u* (present tense), *‖am* (past tense), and *‖ae* (identificational sentences). It is possible (though not certain) that they may all contain the same root morpheme. The only other language that has a similar negative morpheme is Wuras' Seroa. There is, however, a possibility that the main negative morpheme in |Xam (recorded as *kx'au* or *kx'áu-ki*, where *-ki* is really a copula) is also etymologically related, with irregular click loss (**‖- → kx'-*) possibly due to frequent usage of the auxiliary morpheme.

  Other attested morphemes include: (a) the negative verb *ǁa ~ ǁo ~ ǁé ~ ǁi* in Bleek's records of ‖Ng!ke (e.g. *ŋ ǁo ‖ai* 'I do not know', *ŋ ǁi kieŋ* 'I do not sleep'), curiously unattested in any other doculect on N‖ng; (b) ‖Xegwi-Z, ‖Xegwi-LH *ʔa* 'not', without any etymological parallels.
- Nossob: ***ta (~ *tu)** (|'Auni *kiá ~ tiá ~ tá*, |Haasi *tʲu ~ tʲa*). ◊ Use of a CV-type morpheme with initial **t-* (its spelling as *ki- ~ ti-* in |'Auni and as *tʲ-* in |Haasi is consistent with its palatalized articulation in many !Ui languages) is common for both doculects. However, |Haasi shows two differently vocalized variants (**tu* and **ta*) whose rules of distribution are difficult to establish based on Story's examples. They may or may not have different etymological background; in any case, the only common invariant for both doculects is safely reconstructible as **ta*.
- Taa: ***‖qʰu-** (!Xóõ *‖qʰúa*, Kakia *‖wa ~ ‖k'a ~ ‖aːˀaːˀ ~ ‖ai*, N|u‖en *‖u*). ◊ In !Xóõ, the general negation morpheme *‖qʰúa* is opposed to the less frequent negative verb *‖'àː* 'not to be; malfunction'; it is possible that some of the variants in Kakia also reflect not one, but two morphemes (e.g. Kakia *‖wa* = !Xóõ *‖qʰúa*, Kakia *‖k'a* = !Xóõ *‖'àː*). We provisionally accept the !Xóõ form, with its aspirated uvular efflux, as the most archaic.
- Tuu+: It seems probable that !Ui (more precisely, N‖ng and Seroa) **‖V-** (especially its present tense stem variant **‖u*) and !Xóõ *‖qʰúa* are etymologically related, even if click efflux correspondences are not fully regular (Taa **-qʰ-* typically corresponds to !Ui **-/k/ʰ-*, see DOG, HAIR), although, on the whole, negation in Tuu languages is quite notoriously unstable, requiring a much more detailed synchronic and historical investigation.

63. ONE [!Ui + [Nossob + Taa]]

- !Ui: ***ʗoʔ-** (|Xam *!waːi ~ !w'aːi*, ‖Ng!ke *‖weː ~ ‖'weː*, ǂKho-M *‖oe*, N|uu *‖'oe*, ‖Kxau *ʔɔɛː*, ‖Ku‖e *‖kx'oa*, Seroa *‖oai*, ‖Xegwi-Z *!oa*, ‖Xegwi-LH *!waː*, ‖Xegwi-B *‖aː*). ◊ This is an interesting case where, at first, it might seem as if not all the listed forms truly belong together.

However, it may be seen that click influxes generally satisfy the conditions defined for the "sixth click" (*ʗ) of Proto-!Ui, namely, an alveolar reflex in |Xam, a lateral release in all the doculects of N‖ng, and a (rare) alveolar reflex in ‖Xegwi (although note a lateral variant in ‖Xegwi-B). A second issue is the odd variation between simple and glottalized variants of the click accompaniment, sometimes within the same dialect cluster (see data on |Xam and N‖ng). This, keeping in mind the diphthongial nature of the coda in most dialects, may be interpreted in favor of reconstructing the full !Ui stem as *ʗV?V, more precisely, perhaps as *ʗo?e or even as *ʗoa?-i, where *-i is a morphological add-on responsible for diphthongization *-oai → -oe in N‖ng and ‖Kxau. (Although this solution is imperfect in that the function of this *-i remains unclear, its suffixal nature would help explain the discrepancy with ‖Xegwi, which could be assumed to still preserve the unexpanded stem *ʗoa). The etymological decision to trace all these forms back to a single root is also confirmed by external data (see below) and systemic considerations (not a single !Ui doculect by itself presents evidence for two different inherited roots).

- Nossob: ***ǂ'uŋ** (|'Auni *ǂ'ű̃ ~ ǂ'ű̃-u*, |Haasi *ǂŋ́-k'a*). ◊ If the |'Auni and |Haasi forms belong together, this requires assuming that the latter form is either incorrectly transcribed or represents a case of irregular development from an original *ǂ'/V/ŋ, perhaps contracted due to frequent usage in combination with the copula element *-k'a*. (For an irrefutable example of irregular contraction of numerals in |Haasi, see TWO below). In any case, external evidence from Taa clearly shows that the |'Auni form is more archaic in its shape, regardless of whether it is etymologically the same as |Haasi *ǂŋ́-* or if the latter is a separate innovation.
- Taa: (?) ***ǂ'u-** (!Xóõ *ǂ'ûã*, Kakia *!k'we*, N|u‖en *!'oe*). ◊ The variable form of the stem in !Xóõ is *ǂ'u-V*, reflecting *ǂ'u-* as the original root. Morphological framing of the root in Taa differs from that in Kakia and N|u‖en, both of which rather go back to the variant *ǂ'u-e* than *ǂ'u-ã*. The fact that the palatal click in !Xóõ here corresponds strictly to the alveolar click in Kakia and N|u‖en (rather than a messy variation between *ǂ-*, *‖-*, and *!-*, typical of D. Bleek's records) is perhaps worth noting (see the same situation with TWO below), but it is yet unclear if this can be taken as evidence for reconstructing the "sixth click" on the Proto-Taa level (rather than assuming that it had merged with simple *ǂ- already before the split of Proto-Taa into its descendants).
- Tuu+: Despite some unclear moments, there is compelling evidence to think that all three branches here show reflexes of a single original root, perhaps to be reconstructed as *ʗU?U- (which would allow to explain the discrepancy in click efflux correspondences between Nossob and Taa on one hand, and !Ui on the other). Even so, Nossob and Taa forms are clearly closer to one another than !Ui in their phonological reflection of the original root.

64. PERSON [-]

- !Ui: ***!u-i** (|Xam *!úi ~ !uí-ya*, ‖Ng!ke *!wa ~ !wi*, ǂKho-D *!wī*, N|uu *‖̃ŋ=!ui*, ‖Xegwi-Z *kwi*, ‖Xegwi-LH *kwi:*). ◊ All languages show reflexes of the exact same proto-stem, sometimes compounded with the stem *‖̃a-iŋ (→ N‖ng ‖̃ŋ) 'house'. The stem *!ui itself is morphologically complex, consisting of the original root *!u (well attested in such |Xam compounds as *!ü ǀ'a* 'girl', *!ü ǀ'a:iti* 'woman', *!u-de* 'someone', etc.) and the common nominal suffix *-i; another possible morphological variant is *!u-a*, attested in ‖Ng!ke.

A common suppletive plural for this word is reconstructible as *ǂe ~ *ǂ'e: |Xam *!ɛ ~ *!'ɛ* (Bleek), *!é ~ !k'é* (Lloyd), ǁNg!ke *!'e ~ !k'e*, N|uu *ǂe:*, ǁXegwi-Z *ǂ'e*, ǁXegwi-LH *ǂ'e*. The variation observed between click accompaniments is very similar to the one observed in reflexes of HEART q.v. and might go back to an original form like *ǂq'e*, unfortunately, without confirming parallels in Taa.

- Nossob: ***ǂe ~ *ǂ'e** (|'Auni *ǂ'i ~ ǂ'e*, |Haasi *ǂɛ*). ◊ In Bleek 1937: 218, both |'Auni forms are glossed as 'men, people', but text examples show that they can easily be used as singulatives (e.g. *ǂ'i ti ǂ'ú-u* 'one person'), and a separate plural form *ǂ'i-te* 'people' is attested as well. For |Haasi, Story lists specifically *ǂɛ* 'person' and *ǂɛɛ* (=*ǂɛʔɛ*) 'people'. Comparison with !Ui shows that the original meaning may indeed have been plural, but, unlike !Ui, the Nossob languages have neutralized original suppletion in this paradigm in favor of the plural form.
- Taa: (a) !Xóõ *tâ:*, dial. *lâ:* 'person'; (b) !Xóõ *tû:* 'people', Kakia *tu* 'person', pl. *tu-ku ~ tu-tu* 'people', N|u‖en *tu* 'person', pl. *tu-tu*. ◊ The Taa equivalent for sg. 'person' is part of a very small set of morphemes which exhibit the unusual consonantal variation *t ~ l* across dialects (another well-known example is the plural marker *-te*, dialectally encountered as *-le*). This variation is restricted to only a few morphemes, and no phonetic conditioning for a hypothetic shift **t → l* has been established; consequently, it is justifiable to see here the reflex of a special rare proto-phoneme, a reasonable phonetic interpretation of which would be a lateral affricate (**ǂ*). The original paradigm must have been sg. **ǂa*, pl. **tu* (as in !Xóõ); Kakia and N|u‖en show generalization of the plural form, with subsequent formation of new suffixal plurals (the process is structurally identical to that in Nossob languages, though the actual morphemes are different).

  It is important to note that Kakia and N|u‖en also have additional forms listed by Bleek in the meaning of 'person', notably Kakia *da* (in such bound formations as *!on-da* 'old man', *!on-da-ke* 'old woman') = N|u‖en *da* (in the example *da !'oe*, *du l̃um* 'one person, two people'); and N|u‖en *ša*, 'person', pl. *ša:re* (not found in text examples). The etymology of these forms and their semantic connection with *tu* remain unclear, but it cannot be excluded that both *da* and *ša*, or at least one of these forms, are actually dialectal reflexes of Proto-Taa **ǂa*.
- Tuu-: Not reconstructible. ◊ It is interesting to note that in both Proto-!Ui and Proto-Taa the paradigm for this word was most likely suppletive (!Ui sg. **!u-i* vs. pl. **ǂq'e*; Taa sg. **ǂa* vs. pl. **tu*), yet none of the alternants have any parallels in respective branches. The Nossob languages have their separate agenda here, but at least it is clearly closer to !Ui than to Taa (generalization of the former plural form for singulative usage).

65. RAIN [-]

- !Ui: Not reconstructible. ◊ In Starostin 2013: 378, I have attempted to reconstruct Proto-!Ui **ǂq$^{(h)}$au* 'rain' based on presumable common origins for N|uu *ǂqau*, on one hand, and |Xam *!wa: ~ !^{h}wa:*, ǁNg!ke *!^{h}a ~ !a:*, on the other. This decision now seems hasty to me: although N|uu *ǂ-* is indeed a regular correspondence for |Xam, ǁNg!ke *!-*, and even the click effluxes may be accommodated, the codas remain incompatible: N|uu *-au* should regularly correspond to *-au* in other languages (see **ǁxau* BLOOD), and there is no evidence for treating *-u* as a nominal suffix. Moreover, the fact that the earliest attested dialect of N‖ng agrees with |Xam, but not with modern N|uu, is indirect evidence for treating *ǂqau* 'rain' as an innovation (of unclear origin).

Similarly, ǁXegwi -LH *ɕʰeuŋ* 'rain', included in the same etymology, should be disqualified for at least three reasons: (a) phonetic — initial consonant should rather be *č-* or *š-* if it is to reflect Proto-!Ui **ǂq/ʰ/-*; (b) semantic — it is actually attested as the verb 'to rain' rather than noun 'rain'; (c) dialectal — for ǁXegwi-B, Bleek lists the equivalent for 'rain' as *gaa*, which is seemingly the same word as ǁXegwi-LH *gaʔa* 'sky' (← Proto-!Ui **ǃ̰aʔa*).

In the end, if there is a minimally likely candidate for RAIN in Proto-!Ui, it should be the same word as WATER q.v., since this is at least a common isogloss between |Xam and some doculects of N‖ng. But seeing how ǁXegwi rather neutralizes RAIN and SKY, and given the general lack of stability for this meaning in the Tuu area, it is preferable to leave the slot empty.

- Nossob: |'Auni *ǁʰàːa*. ◊ This is possibly the same word as WATER, even though the two are transcribed differently in Bleek's data. Not attested in |Haasi (there is a verb *ǂī́* 'to rain', but these two meanings are often lexically differentiated in Tuu).
- Taa: ***!kx'oe** (!Xóõ *!kx'ôe*, N|u‖en *!xweː*). ◊ The isogloss between !Xóõ and N|u‖en seems fairly straightforward, even if this is the only spotted case so far in which !Xóõ *-kx'-* corresponds to N|u‖en *-x-* (*-kx'-* and *-k-* are encountered more frequently in Bleek's data). The situation becomes more complex, however, if we also take into account the Kakia form *!we-ga-ǁa*, where *!we* may be the same morpheme as !Xóõ *!kx'ôe*, *-ga-* marks possession, and *ǁa* is possibly a mistranscription of the word WATER q.v.; if this analysis is correct, it would mean that 'rain' was probably not the original meaning of this word and that it could instead have denoted something like 'raincloud'. This would be consistent with observations on !Ui, where RAIN does not like to behave as a true "semantic primitive" and is typically derived from 'water' or 'sky'. Still, on a purely formal level of analysis this should not prevent us from reconstructing Proto-Taa ***!kx'oe** 'rain', perhaps with polysemy ('rain / raincloud').
- Tuu-: Not reconstructible. ◊ Based on analysis of evidence in all three branches, it is reasonable to suggest that there was no separate lexical root for RAIN at all in Proto-Tuu; whether the meaning was linked to a lexeme like WATER, SKY, or CLOUD cannot be properly ascertained.

66. RED [-]

- !Ui: ***ᶑi** (|Xam *!íː ~ !íː-ya ~ !ʰíː-ya*, ǁXegwi-LH *!e).* ◊ A solid isogloss between |Xam and ǁXegwi; preservation of *!-* in the latter is an argument in favor of the "sixth click" ***ᶑ-**. A certain problem concerns the fact that D. Bleek considers the |Xam word to be related to ǁNg!ke *!i* 'red ochre used as pigment', which is a good semantic match, but, according to the model of correspondences laid out in the first part of the paper, should have looked like **ǁi* instead. However, one should also note that ǁNg!ke *!i* 'red ochre' could actually be related not to the |Xam word, but to the word attested as ǂKho-D *ǂ'ī* and N|uu *ǂqiː* (Collins & Namaseb 2011: 14), both meaning 'red' (although in the sub-dialects of most of the modern speakers of N|uu, the equivalent for 'red' is the Khoekhoe borrowing *|kx'aba*). Since phonetic reasons make it impossible to construct an etymology which would unite |Xam, N‖ng, and ǁXegwi equivalents at the same time, we prefer to identify two etymologically different roots, **ᶑi* 'red' and **ǂqi* 'red ochre', the latter represented only within the large N‖ng cluster and probably developing the adjectival meaning 'red' in some of its daughter dialects.

- Nossob: |Haasi *cxwe-k'a*. ◊ Clearly the same root as |'Auni *coa* 'red colour, ochre', although it is unknown if the |'Auni root also had an adjectival usage (no other words for 'red' are attested in Bleek's data).
- Taa: ***|a^{h}-ŋa*** (!Xóõ */ā^{h}ŋa*, Kakia */anya*, N|u‖en *ǀàne*). ◊ The root is probably just **/a^{h}-*, since **-ŋa* is a suffixal component frequently seen in other color terms as well (see BLACK, WHITE). The N|u‖en form is only included in Bleek 1929 and is not confirmed in Bleek 1956, but seems close enough to the other two forms to merit inclusion.
- Tuu-: Not reconstructible. ◊ All three branches have their own equivalents, without any clear etymological parallels on the Common Tuu level.

67. ROAD [!Ui + Nossob (?)] [- Taa]

- !Ui: ***!aN*** (N|uu *!an ~ !aŋ*, ‖Xegwi-Z, ‖Xegwi-LH *kaŋ*). ◊ Optimal candidate for Proto-!Ui status is unequivocally pointed out by the phonetically impeccable and semantically precise isogloss between N|uu and ‖Xegwi (although the precise phonemic nature of the final nasal remains somewhat unclear). No traces of this root, however, are found in either |Xam or earlier attested forms of N|uu, all of which have their own, etymologically unclear equivalents (|Xam *!xárra* 'path', ‖Ng!ke *tirau* 'path').
- Nossob: |'Auni *!án* 'path'. ◊ Attestation of this form allows for the possibility of a Proto-Nossob **!an*, cognate with Proto-!Ui; however, no equivalents for 'road' or 'path' are attested in |Haasi, and there is nothing to disprove that the |'Auni form may just as well represent a borrowing from one of the N|uu dialects, especially since there is at least one more N|uu-|'Auni isogloss with similar phonetics and semantics: N|uu *ǁuru-ke* 'path, road, trail' = |'Auni *ǁùru* 'path'. Finally, Bleek lists yet another |'Auni word with the same meaning: *ǂkx'ei* 'road, path', which, if the other two are N|uu borrowings, could actually represent the inherited Nossob term. The situation cannot be properly resolved at our current level of knowledge.
- Taa: !Xóõ *ǂólo*. ◊ This looks like a possibly inherited term, glossed by Traill as 'path'; in the more "cultural" meaning 'road, way' !Xóõ has the Khoe borrowing *dào*, which is also the only known equivalent for this entire semantic cluster in Kakia (*dau*) and N|u‖en (also *dau*).
- Tuu-: The isogloss between Proto-!Ui and |'Auni is undeniable, but could reflect either common heritage or areal contact. Given the overall lack of stability of this concept (easily replaceable by borrowings from Khoe or words with unknown etymology), a reconstruction at the Proto-!Ui-Nossob level does not seem particularly trustworthy.

68. ROOT [-]

- !Ui: Not reconstructible. ◊ Given the specific botanic circumstances of the Tuu-speaking area, this concept is notably unstable, and it is never clear just how specific or generic attested terms are; cf. |Xam *!aui* 'root' in Bleek 1929: 71, but the same word glossed as 'wild onion' in Bleek 1956: 414. For N|uu *!^{h}abe-si* 'root', B. Sands states (p.c.) that "only one of the Eastern N|uu speakers knows this word", and suggests *ǂao-si ~ ǂãũ-si* 'root of shepherd's tree (*Boscia albitrunca*)', sometimes extended to denote 'root' in general as a more suitable term. For most of the other languages, the exact meaning 'root' is not attested at all.
- Nossob: |'Auni *ǂau-si* is glossed as 'small roots' in Bleek 1937: 219. ◊ This is clearly the same word as N|uu *ǂao-si ~ ǂãũ-si* 'root of shepherd's tree', mentioned above, and,

given the exact same morphological structure in both cases, is probably a recent borrowing into |'Auni from N|uu.
- Taa: Not reconstructible. ◊ Cf. !Xóõ *!kx'á-i*, pl. *!kx'á-ba-tê* 'woody root'; N|u‖en *ǃau-te* 'root' (a plural form). The precise semantics of both terms remains unclear.
- Tuu-: Not reconstructible, although it must be noted that N|u‖en *ǃau-te* (assuming that the transcription with *!-* is either erroneous for *ǂ-* or reflects a regular development from it) is formally compatible with N|uu *ǂao-si ~ ǂãũ-si*. Still, due to the overall difficulties with this concept, it is probably better to exclude it from analysis.

69. ROUND [-]

- !Ui: Not reconstructible. ◊ This concept is only attested for |Xam (*kːuérre-kːuérre ~ kwórre-kwórre*, a reduplicated verbal stem applicable to round objects such as 'sun' or 'egg'); there is also ‖Ng!ke *kəkɛriŋ* 'round' in Bleek 1929: 71, which may or may not be related to the |Xam item, but finds no confirmation in any other of the many later sources on the N|uu cluster.
- Nossob: Not attested for either |'Auni or |Haasi.
- Taa: Not reconstructible. ◊ Cf. !Xóõ *ʔǂnúʔm ǀʼə̂ː* 'round shaped, tubular (e.g. a branch, rod), as opposed to flat-shaped'; another synonym with almost the exact same meaning is *ʔ!nə̃ʔõ ǀʼə̂ː*.
- Tuu-: Not reconstructible, mainly due to lack of attestation in daughter languages; a rather problematic concept for the Khoisan area as a whole.

70. SAND (= EARTH) [-]

- There is not a single reliable case in which a Tuu word for 'sand' would be lexically different from the corresponding word for 'earth'; it may be safely assumed that such a differentiation did not exist on the Proto-Tuu level, either. See EARTH in the first part of the paper for more details.

71. SAY [-]

- !Ui: ***ku ~ *ka** (‖Ng!ke *ka*, ǂKho-M *ka ~ ku ~ kʼu ~ kwa ~ kɔˀ ~ kuː-wa*, N|uu *ka ~ ku*, ‖Kxau *ku*, ‖Xegwi-Z *kũ*). ◊ Variations in root vocalism are attested for this verbal root in both earlier (Maingard 1937) and newer (Collins & Namaseb 2011) sources on N|uu, seemingly without any explanation; forms such as Maingard's *kuː-wa* hint that *ka* may historically represent a contraction from **ku-a*, but this is not completely certain, though, most likely, both variants do reflect the same root (‖Xegwi-Z *kũ* also supports the idea of the labial vowel as original). In |Xam, *ka* is found glossed as 'to wish, intend, think, say' and seems to refer more frequently to mental than verbal activity; the basic equivalent for 'say' (introducing direct speech, etc.) is the morphologically complex verb *ǂákkən ~ ǂákka ~ ǂákən ~ ǂáka ~ ǂággən*, which is perhaps related to N|uu *ǂʰoa* 'to speak' and, in any case, seems to be innovative in the meaning 'to say'.
- Nossob: ***|u** (|'Auni */u*, |Haasi */wa*). ◊ For |'Auni, Bleek records two verbs that introduce direct speech: */u* and *ko*, without any clear differentiation. It may be suggested, based on the |Haasi parallel, that */u* is the inherited term, whereas *ko* is borrowed from N|uu *ku*.
- Taa: ***tV-** (?). ◊ The situation here is as follows. For !Xóõ, Traill records *téʔẽ* 'say', *tám* 'say it, mean', and *tâna* 'talk, speak'; all of these three forms could theoretically be related, going back to a single root *tV-* with various suffixal extensions, but this is im-

possible to confirm by additional examples. For N|u‖en, the only known equivalent of the meaning 'say' is *tana*, but examples of usage (Bleek 1956: 191) clearly show that the actual meaning of the verb is 'speak', just as in !Xóõ (e.g. *n kai tu oz tana te* "I hear the person speaking"). For Kakia, in addition to the same *tana* (also glossed as 'say', but in reality meaning 'speak', e.g. *si ka kumma tana-ne* "we have talked enough"), Bleek also records the verb */e*, which seems to be confirmed as 'say' by several contexts (e.g. *ŋa /e ǰum* "I say thank you"), but finds no parallels outside of this particular doculect. Subsequently, reconstruction of **tV-* 'say' for Proto-Taa is at best tentative.

- Tuu-: Not reconstructible. Even the phonetically and semantically questionable reconstructions of this item for the intermediate branches are all etymologically different (the only isogloss is between |'Auni *ko* and Proto-!Wi **ku*, but even that one, as stated above, more likely reflects areal contact than common heritage).

72. SEE [!Ui + Nossob + Taa]

- !Ui: ***ǰV** (|Xam *ǰa:* ~ *ǰã:* ~ *ǰĩ:* ~ *ǰĩ* ~ *ǰi:* ~ *ǰi*, ‖Ng!ke *ǰa:* ~ *ǰe* ~ *ǰi:* ~ *ǰi*, ǂKho-M *ǰa* ~ *ǰe* ~ *ǰẽĩ* ~ */ŋ*, N|uu *ǰa:*, ‖Kxau *ǰa:*, ‖Xegwi-Z *ǰi* ~ *ǰi-ya*, ‖Xegwi-LH *ǰi* ~ *ǰa*). ◊ Precise reasons for the vocalic fluctuations (also observed in several other verbal roots of this type) are difficult to establish, but probably reflect various mergers of the original (not properly establishable) root vowel with agreement markers or other grammatical morphemes.
- Nossob: ***ǰV** (|'Auni *ǰà:* ~ *ǰe*, |Haasi *ǰã-ǰã*). ◊ Same situation as in !Ui; the |Haasi form seems to reflect a reduplicated variant of the stem.
- Taa: ***ǰV-** (!Xóõ *ǰa̰*, Kakia *ǰa* ~ *ǰe* ~ *ǰün*, N|u‖en *ǰe:* ~ *ǰa:*). ◊ Vowel fluctuations in older sources look the same way as they do in !Ui, although, curiously, the root vocalism in Traill's !Xóõ is always *a* (including the nominal derivate *ǰa̰-sà* 'seeing', which most likely preserves the original quality).
- Tuu+: ***ǰV-**. ◊ It is better to refrain from unequivocal reconstruction of the original root vowel quality, but it is fairly clear that all reflexes in daughter languages are related and that this is one of the most historically stable elements on the wordlist (as far as Tuu languages are concerned).

73. SEED [-]

- This item is removed from consideration due to almost complete lack of attestation; the generic concept 'seed' does not seem to exist in most of the Tuu languages due to the specific realities of Kalahari flora. The only exceptions are |'Auni *c'ou* 'pips, seeds' (Bleek 1937: 207; cf. *ǂwi c'ou* "seeds of tsamma") and !Xóõ *sa̰ʔã* 'seed'. It is perhaps not coincidental that the former item is phonetically close to |'Auni *co:* ~ *c'a-xu* 'eye', while the latter is only tonally and grammatically different from !Xóõ *sàʔã* 'face'; both similarities suggest possible secondary origins for this term. Regardless of this observation, nothing is reconstructible here.

74. SIT [!Ui + Nossob + Taa]

- !Ui: ***so-** (|Xam *s:'o* ~ *š:o*, ‖Ng!ke *so* ~ *so:* ~ *sɔ:*, ǂKho-M *sũĩ* ~ *swẽĩ* ~ *swẽˢ*, N|uu *so:* ~ *sũĩ*, ‖Xegwi-Z *šo-ge*). ◊ Data from |Xam and N|uu allows to reconstruct an old contrast between the stative verb 'to sit = to be sitting', expressed by the pure root **so*, and the dynamic action verb 'to sit down', expressed by the suffixal extension **so-iŋ* (assimilated to **su-iŋ* ~ **sũ-ĩ* in most dialects of N|uu, but cf. *s:oé:ŋ* 'to sit down' in W. Bleek's

|Xam records with a different direction of assimilation). Semantic glossing is not always accurate across various sources, but the contrast between stative and non-stative has at least been expressly documented for N|uu (Collins & Namaseb 2011: 20). Ziervogel's transcription *šoge* for ǁXegwi most likely reflects the same stem as *šoɢaʔane* 'sit and wait form me' in Lanham & Hallowes 1956: 116, analyzed as *šo* 'sit' + *ɢaʔa* 'wait' + *ne* 'me'; Bleek simply lists *šo:* as the primary equivalent for ǁXegwi.

There is also a common !Ui suppletive plural stem for this verb: |Xam *!ʰau:* ~ *!ʰáu-ʒwa* 'to sit' = ǁNg!ke *!ãũ* = N|uu *!qʰãũ* id. (← Proto-!Ui **!qʰau-*).

- Nossob: |'Auni *sã* ~ *são* ~ *sõ* ~ *so* 'to sit; to sit down, put down, set, stay'. ◊ Not clear if this item, easily relatable to !Ui **so-*, is etymologically the same as |Haasi *c'i* 'to sit'; the main problem here would not be so much the ejective affricate as the completely unexpected vocalism. In any case, the |'Auni forms, due to their own specific vocalic properties, are probably authentic rather than borrowed from N|uu (regarding variants such as *sã*, the correspondence between |'Auni *a* and !Ui *o* is quite current; variants such as *sõ* ~ *so* may either preserve the original vocalism in specific grammatical or phonetic contexts, or could indeed be influenced by N|uu).

  Both |'Auni and |Haasi have additional synonyms for this concept, e.g. |'Auni *!ãũ* 'to sit, to squat'; |Haasi *!xi-k'i* 'to sit'. The former is clearly the same item as Proto-!Ui **!qʰau-* 'to sit' (of many); whether this is an inherited Nossob term or a borrowing from N|uu is impossible to determine, both solutions are more or less equiprobable. For the |Haasi term, no diagnostic contexts are available to ascertain its true meaning or usage.
- Taa: ***cʰu** (!Xóõ *cʰû:*, Kakia *ču*, N|u‖en *šu* ~ *ču*). ◊ The precise nature of the affricate remains to be ascertained; cf. the stunning number of variants for this root listed in Maingard 1958 (*c'o* ~ *č'o* ~ *čə* ~ *č'ou* ~ *su* ~ *šu* ~ *čʰo*). Unlike !Ui, there seems to be no separate morphological variant for the non-stative verb 'to sit down'. There is, however, also a suppletive plural stem: !Xóõ *!'ã́:* = N|u‖en *!'a:* 'to sit' (of many), reconstructible as Proto-Taa **!'a-*.
- Tuu+: ***c(ʰ)o** is reliably reconstructible, despite some phonetic uncertainties (aspiration, etc.), as the common Tuu morpheme for the meaning 'sit' based on data from all three branches. The suppletive plural stems (Proto-!Ui **!qʰau-* and Proto-Taa **!'a-*), despite a matching click influx, are on the whole unreconcilable with each other; note that, because of |'Auni *!ãũ*, Nossob and !Ui are closer to each other in this respect than to Tuu.

75. SKIN [!Ui + Taa] [- Nossob]

- !Ui: ***Tuŋ** (|Xam *t:ũŋ*, ǁNg!ke *tũ* ~ *twã* ~ *diõ*, ǂKho-M *ʒo*, N|uu *ʒũ:*, ǁXegwi-LH *tuŋ* ~ *tũ:*). ◊ Of note in this case is the unexpected voiced articulation of the initial consonant in modern N|uu, as well as at least some of the older dialects (palatalization **t-*/**d-* → **ɸ-*/**ʒ-* in them, on the other hand, is quite regular). A few other cases of such fluctuations are known, but this is the most transparent one; it is possible that the voicing reflects influence of some additional feature that cannot be recovered from available data. Cf. the situation in Taa for a potential clue.
- Nossob: ***ǁ(')U**: |'Auni *ǁ'ù:*, |Haasi *ǁɔ*. ◊ Although Bleek's and Story's data contradict each other on the precise nature of the click efflux, this is hardly a valid reason to deny the common origin of both items. If the word is related to |Xam *ǁɔ* 'outer skin', 'shed skin' with its voiced efflux, this makes the reconstruction **ǁU* more probable than the one with glottalization.
- Taa: ***tuˀm** (!Xóõ *tùˀm*, Kakia *t'üm*, N|u‖en *t'ùm*). ◊ It is likely that glottalized *t'-*, attested in Bleek's transcriptions for Kakia and N|u‖en, is a perception error for *t-* + pharyn-

gealized vowel, as in Traill's transcription of !Xóõ. Alternately, it is possible that glottalization was in fact primary, and that !Xóõ underwent a sporadic (or even regular) change from **t'um* to **tuˀm*.

- Tuu+: !Ui **Tuŋ* and Taa **tuˀm* (**t'um*?) are clearly relatable; coda correspondences here are the same as in BREAST and LIVER, implying either different models of suffixation or a phonetic scenario (**-ŋ* → **-m* in Taa as a regular development?). It also seems as if the strange voicing in N|uu and the fluctuation between *t'u-* and *tuˀ-* in Taa may be correlated, but until more examples of such correlation are found, it is difficult to jump to conclusions. In any case, this is yet another instance where !Ui and Taa agree vs. the Nossob languages, which have innovated a different term (perhaps semantically extended from more specialized usage).

76. SLEEP [!Ui + Nossob + Taa]

- !Ui: ***θu- ~ *θi- (**|Xam *θu̯oin*, ‖Ng!ke *θwoiŋ ~ θwoeŋ ~ θóeŋ*, ǂKho-M *θ'wõ ~ θ'wonna*, N|uu *θun ~ θuŋ*, ‖Kxau *θan*, ‖Xegwi-Z *θi*). ◊ The only thing certain in this reconstruction is the initial consonant; root vocalism may have been **i*, as seen in Ziervogel's "future tense" stem for ‖Xegwi (he also lists *θi-ɲe* as "present tense" and *θi-ɲa* as "past tense" stems), with assimilation to the labial click everywhere else, but this is not obvious. Most of the stems end in nasal consonants, but, again, ‖Xegwi *θi* shows that they may all be suffixal in origin.
- Nossob: ***θV-i-** (|'Auni *θwõi*, |Haasi *θwa-ai*).
- Taa: ***θV-n-** (!Xóõ *θân*, Kakia *θwõĩ ~ θwoin*, N|u‖en *θwoin*). ◊ Cf. also !Xóõ *θûm* 'sleep (n.)'.
- Tuu+: This is a very stable root, but variation in the coda across all known languages is too extensive to allow for an unequivocal reconstruction other than **θV-*. It is certain that at least some of the original morphological variants ended in a nasal (**θV/-/n*), but whether it was really an integral part of the root or a suffixal extension is hard to determine.

77. SMALL [-]

- !Ui: ***ǂ'eni** (|Xam *ǂ'ɛŋŋi ~ ǂ'ènniŋ*, ‖Ng!ke *ǂ'ĩ*, N|uu *ǂ'ĩ*, ‖Xegwi-Z *ǂ'ine*, ‖Xegwi-LH *ǂ'ini*). ◊ This adjectival stem is preserved in its bisyllabic shape in |Xam and ‖Xegwi (where the development of palatal click into a lateral affricate is regular; LH *ǂ'ini*, with glottalic articulation, is probably a more accurate transcription than Ziervogel's *ǂ'ine*), but contracts to *ǂ'ĩ* in N|uu. Of note is the explicit marking of the palatal click articulation in |Xam, since typically Proto-!Ui **ǂ* is marked in both W. Bleek's and L. Lloyd's records as *!*; this may have something to do with the "expressive" nature of the word.

  Additionally, in the meaning 'small' some sources list reflexes of a common !Ui morpheme that starts with a labial click, e.g. ‖Ng!ke *θwain-ki*, ǂKho-M *θ'kõ*, ǂKho-D *θó-nē*, ‖Xegwi-LH *θa-ri*. This corresponds to |Xam *-θwa ~ θu̯á*, a diminutive morpheme usually found in conjunction with words denoting 'children' or 'young of birds / animals', and to modern N|uu *θũ* with more or less similar usage. It is quite possible that this is the original equivalent for the unbound adjective 'small', but that already in Proto-!Ui its usage had become more restricted, whereas **ǂ'eni* had already acquired more productive functions.
- Nossob: (a) |'Auni *ǁai*; (b) |Haasi *nʸái-si*. ◊ The |'Auni form is not very reliable, as it is attested only in Bleek 1929: 76 (an early source); in any case, it can hardly be the same as |Haasi *nʸái-si*, which is fairly unique in itself (beginning with a palatal nasal). Unclear situation on the whole.

- Taa: (a) !Xóõ *ǀ'ûi* (suppletive pl.: *ǀq'án-tá*); (b) Kakia *ǀǰona*; (c) N|u‖en *ǀari*. ◊ The concept is clearly unstable in the Taa branch as well; all three doculects show different equivalents. Note the presence of the diminutive formant *Θà:* in !Xóõ, e.g. *Θàye-Θà:* 'animal' (= 'meat-small').
- Tuu-: While the diminutive morpheme **ΘV-* is clearly archaic and reconstructible at the Proto-Tuu level, the same cannot be said about the unbound adjective 'small', expressed by different equivalents at the level of each subgroup (and even notoriously unstable within most of them).

78. SMOKE [!Ui + Nossob] [- Taa]

- !Ui: ***‖o^{ʕ}** (|Xam *‖ǰ:*, ǂKho-D *‖'ɔʔō-ké*, N|uu *‖o:^{ʕ}-ke*). ◊ Although the |Xam form is attested only scantily, the isogloss between it and the N|uu cluster allows to make a reliable reconstruction at least on the "narrow !Ui" level. The only other well-attested form is ‖Xegwi-LH *k^{h}aʔa-zi*, whose origins are hard to ascertain (according to Lanham & Hallowes, the suffix *-zi* frequently marks verbal derivatives or borrowings from Bantu languages, but *k^{h}aʔa* is hardly identifiable as either a verbal stem or a Bantu borrowing).
- Nossob: ***‖au** (|'Auni *‖áu*, |Haasi *‖au*). ◊ The |'Auni form is listed as a noun in Bleek's dictionary; for |Haasi, Story indicates that *‖au* is found in both nominal and verbal usage, although the only textual example is within the context 'I smoke'.
- Taa: (a) !Xóõ *ckx'âye*; (b) Kakia *ǁ̃a:lu*. ◊ Not attested in N|u‖en. The form in !Xóõ is notoriously similar to Proto-Khoe **c'án(ì)* 'smoke' (Vossen 1997: 476), especially since it is possible that the former was phonetically realized as **ckx'án-*; however, direct borrowing from a relatively recent Khoe source is excluded for phonetic reasons.
- Tuu+: Comparison between Proto-!Ui **‖o^{ʕ}* and Proto-Nossob **‖au* 'smoke' is distributionally and semantically solid; phonetics-wise, the correlation between **-o* and **-au* may raise questions, but is not completely unprecedented (at least Nossob **-au* vs. !Ui **-oe* is recurrent, cf. |Xam *‖xwe:* vs. |'Auni *‖xau* 'cold', |Xam *‖kx'oe* vs. |'Auni *‖kx'au* 'back'). The parallel may be accepted as a lexicostatistical match between the two branches, whereas the situation in Taa is different.

79. STAND [-]

- !Ui: (a) |Xam *!^{h}e: ~ !^{h}é*; (b) ‖Ng!ke *ǁ̃a*, ‖Kxau *ŋa*; (c) N|uu *l̃a: ~ l̃a'a*; (d) ‖Xegwi-Z *!'oʔo-ge*. ◊ The situation here is complicated not just because of the relative instability of the concept, but also due to insufficiently accurate semantic glosses in most available sources. Thus, while forms in group (b) are clearly the same verb as |Xam *ǁ̃a* 'to be (located), stay', available examples do not make it clear if ‖Ng!ke truly extends the usage of this verb to contexts with the meaning 'stand' (in a vertical position), or if it is actually the same as in |Xam. In light of this, no particular reconstruction can be reliably offered for Proto-!Ui.
- Nossob: (a) |'Auni *!'ǎ́*; (b) |Haasi *t̃wa*. ◊ Despite some phonetic similarity, the two forms are hardly reconcilable with each other (lack of glottal release and extra labialization in |Haasi would be unexplainable in a common etymology). It should be noted that Collins & Namaseb also record an alternate variant *!'ana* 'to stand' for N|uu; if so, the |'Auni form could be suspected of having been borrowed from N|uu (with secondary contraction).
- Taa: ***‖^{h}ũ** (!Xóõ *‖^{h}û:*, Kakia *‖ũ ~ ‖õ ~ ‖^{h}ũ ~ ‖um*, N|u‖en *‖ũ ~ ‖'^{h}u ~ ǂ'^{h}u*).
- Tuu-: Not reconstructible. ◊ The concept seems to be fairly stable in the Taa cluster, but not in !Ui or Nossob languages.

80. STAR [-]

- !Ui: (a) ***‖kx'oaˀ-** (‖Ng!ke *‖waiˀ-sa ~ ‖kx'we:-sa*, ǂKho-D *‖'wāi-ɕē* pl., N|uu *‖kx'oeˀ-si*, ‖Ku‖e *‖an-te* pl., ‖Kxau *‖'ɔan-si*); (b) |Xam *|u̯áˀ-ttən*, ‖Xegwi-Z *|ou-ni* pl. ◊ The most commonly encountered root for 'star' in !Ui is spread throughout the N|uu cluster and is further confirmed by entries for ‖Ku‖e and ‖Kxau; the common invariant in all these forms is the root **‖kx'oa-* (probably with vowel pharyngealization, expressly attested in Bleek's data on ‖Ng!ke as well as modern N|uu), which is found in conjunction with different suffixes (N|uu **‖kx'oaˀ-i* → modern *‖kx'oeˀ*; ‖Ku‖e + ‖Kxau **‖kx'oa-n*) and additional markers of singularity or plurality.

  To this root is opposed the entry in |Xam *|u̯áˀ-ttən* (cf. the reduplicated plural: *|uaˀ-|uáˀ-ttən*), which finds a distributionally surprising parallel in ‖Xegwi-Z (*|ou-ni*). The latter form is somewhat suspicious as to its morphological constitution, and finds no support in alternate sources for ‖Xegwi (Bleek has an etymologically unclear *kale* 'star' in its place); also of note is the seemingly full homonymy of the root morpheme in |Xam with *|u̯aˀ-* 'cloud' (although any semantic connection between 'star' and 'cloud' would be decidedly non-trivial). Still, phonetically and semantically the match between |Xam and ‖Xegwi is impeccable, allowing to reconstruct **|uaˀ- ~ *|auˀ-* (vocalism metathesis in ‖Xegwi?) as an alternate candidate.
- Nossob: |'Auni *!'ʰa:*. ◊ Not attested in |Haasi.
- Taa: ***‖ona** (!Xóõ *‖ōna*, Kakia *‖wana-te ~ ‖wana-te* pl., N|u‖en *‖'ana-te* pl.). ◊ This could originally have been a plural form, given the abundance of sg. **CVn* / pl. **CVn-a* paradigms in !Xóõ (with subsequent formation of a new, more productive plural by means of the formant *-te*).
- Tuu-: It is extremely tempting to compare Proto-!Ui **‖kx'oaˀ-* and Proto-Taa **‖ona*, but the match would clearly be problematic due to the non-recurrent nature of the efflux correspondence. Until additional supporting evidence surfaces, we prefer to keep these items separate, and assume that no common etymon can be reliably reconstructed for Proto-Tuu.

81. STONE [-]

- !Ui: ***!ao** (|Xam *!au ~ !óu*, ‖Ng!ke *!au ~ !áu*, N|uu *!ao*, ‖Kxau *!ao*, ‖Xegwi-LH *ɕ'eo*). ◊ Inclusion of the ‖Xegwi form is somewhat problematic: *ɕ'eo* should normally go back to **k'ao* ← **!'ao*, yet there are no signs of a glottalized efflux in either |Xam or N|uu. On the other hand, glottalization is indirectly supported by the curiously shaped ‖Ku‖e form *dʔɔ* 'rock', recorded by D. Bleek (*d-* is the normal ‖Ku‖e reflex for **!-*, cf. *dõa* 'tortoise' = |Xam *!óë* id., *dɔaxu* 'sky' = |Xam *!waxu* id., *dwene* 'three' = |Xam *!wanna* id., etc.); one may wonder if this does not reflect an original **!aʔo*, contracted to **!'ao* in some of the daughter languages. More problematic is that ‖Xegwi also has variants of the word 'rock, stone' with initial affricates or fricatives: cf. ‖Xegwi-LH *ǯwe*, ‖Xegwi-Z *žeu*, ‖Xegwi-B *že*, *žu:*. Lanham & Hallowes mention that "this is not the common Bushman word for 'stone', and it was obtained from one group of informants only", and propose borrowing from Sotho *li:=ǯwe*. This explanation is not unquestionable, but the alternate solution (suggesting an irregular dialectal development **!-* → *ǯ-*?) is certainly less preferable.
- Nossob: (a) |'Auni *‖kx'ɔ:*; (b) |Haasi *!òè*. ◊ The former item is only found in the early source of Bleek 1929 and is quite dubious. The form in |Haasi is at least well supported by text examples.

- Taa: ***ǁu-** (!Xóõ *ǁū-le*, pl. *ǁū-n*, Kakia *ǁü-le ~ ǁǁü-le*). ◊ The Kakia variant with the lateral click is probably just a typo. N|u‖en has a different entry — *!'um* 'stone', which may or may not be a phonetic variant or a mistranscription of the same word as *!um* 'mountain' (see above).
- Tuu-: The main !Ui and Taa equivalents for 'stone' are clearly different. It is possible that |Haasi *!òè* and !Ui **!ao* (**!aʔo*?) represent the same root, but the correspondence between the diphthongs is unique, requiring us to set up some complicated morphological scenario to account for it; we prefer to treat these two items as unrelated for the moment.

82. SUN [!Ui + Taa] [- Nossob]

- !Ui: ***‖'oN-** (|Xam *‖kx'õĩŋ ~ ‖ó̃ĩŋ ~ ‖'õĩː ~ ‖'õĩːŋ ~ ‖'ú́ĩ́ŋ*, ‖Ng!ke *‖'õẽ ~ ‖'õĩ ~ ‖'õĩŋ ~ ‖õẽ ~ ‖õĩ ~ ‖õĩn*, ǂKho-M *‖'ũĩ̃*, N|uu *‖'ũĩ̃*, ‖Kxau *‖'oːeː*, ‖Xegwi-LH *‖umi*, ‖Xegwi-B *‖õĩ̃ ~ ‖'uːn*). ◊ The ‖Xegwi-LH form suggests an early Proto-!Ui reconstruction **‖uni* (→ ‖Xegwi *‖umi* with assimilation), which would be then contracted to **‖ũĩ̃* after the separation of ‖Xegwi. Things are, however, more complex, since (a) ‖Xegwi itself also has the variants *‖õĩ* and *‖'uːn*, recorded by Bleek and (b) data from |Xam and N|uu show unpredictable fluctuations between **‖ũĩ̃* and **‖'ũĩ̃*, such as usually stem from complex interactions between the original click efflux and secondary features of the vocalism. One solution (which, arguably, agrees best with external evidence) would be to reconstruct the root as **‖'oN-* (original root vocalism is more likely to have been **o* than **u*, as there would be no reason for |Xam forms like *‖uiŋ* to develop into *‖oiŋ*), and explain most of the non-trivial developments by its merging with various nominal suffixes, e.g. **‖'oN-iŋ* → **‖(')õĩ(ŋ) ~ *‖(')ũĩ̃(ŋ)*, etc. Alternate scenarios are, of course, possible and will be more fully explored in the future.
- Nossob: ***|'V-** (|'Auni */ɛ́ ~ /ɛn*, |Haasi */'i*). ◊ The difference in click effluxes is puzzling; even more puzzling is the phonetic proximity of both forms to the Common Nossob (and Tuu) equivalent for 'fire' (|'Auni */'i*, |Haasi */i* — with *reverse* correspondences!). It is tempting to think that 'sun' in Nossob could be derived from 'fire' with some additional vocalic suffix, later contracting with the root and leading to a shift in the click efflux articulation, but a credible historic scenario is hard to come by; in any case, the word has no etymology that would be separate from 'fire'.
- Taa: ***‖'an** (!Xóõ *‖'ân*, Kakia *‖'ʌn*, N|u‖en *‖'ʌn ~ ‖'ẽː*).
- Tuu+: If the Proto-!Ui root is indeed reconstructible as **‖'oN-* (which is at least one of the possible scenarios), it is perfectly compatible with Proto-Taa **‖'an* (the same vocalic correspondence is seen in BIG, FAT and quite a few items outside the Swadesh list). The Nossob forms clearly do not belong here and, if the connection with FIRE is indeed valid, are to be treated as innovative.

83. SWIM [-]

- Not reconstructible. The only Tuu language for which this verb is attested is |Xam (*θxuː* 'swim'). The word may, in fact, be archaic ('swimming' must have been a common reality for speakers of |Xam and, quite possibly, of Proto-Tuu as well), but there is no comparative evidence whatsoever to prove that.

84. TAIL [Nossob + Taa] [- !Ui]

- !Ui: ***!^{h}Vi** (|Xam *!^{h}wí*, ‖Ng!ke *!ei*, ǂKho-D *!ãĩ^{ʕ}*, N|uu *!^{h}ai*, ‖Xegwi-Z *k^{h}i*). ◊ Reconstruction of the original vocalism is problematic here; *-w-* in |Xam may be secondary (along with

numerous other cases of "epenthetic *-w-*" found in this language), but fluctuation between diphthongic *-ai* and monophthongic *-i* has not been explained. However, **!^{h}-* is reliably reconstructible based on data from |Xam, N|uu, and ǁXegwi (could the odd pharyngealization in Doke's *!āiˁ* be a mistake for aspiration?).

- Nossob: (a) |'Auni *ǂwi*; (b) |Haasi *i=|á:-a*. ◊ |Haasi *i=* is probably a possessive prefix. External data (see Taa below) show that the |'Auni form is likely an innovation.
- Taa: ***|ãũ**, pl. ***|ã** (!Xóõ *|ãũ*, pl. *|ã:*, Kakia *|ãũ ~ ǀãũ*, N|uǁen *ǀ̣ãũ*). ◊ The !Xóõ form is homonymous with NAME (see above) and shares exactly the same grammatical characteristics; the original root may have been **|au-* or simply **|a-* (the external parallel in |Haasi favors the latter choice).
- Tuu+: The undeniable isogloss between Taa and |Haasi *=|a:-* suggests the reconstruction of **|a-* as the original root for TAIL, which Proto-!Ui replaces with an innovation of unknown origin.

85. THAT [-]

- !Ui: (a) |Xam *|e: ~ |e ~ |e:-á*; (b) ǁNg!ke *á*; (c) ǁNg!ke *ǁ̃e=á ~ ǁŋ-á*, N|uu *ǁ̃a: ~ kea*; (d) ǁXegwi-Z *ʔe=ta*, ǁXegwi-LH *ʔe=na ~ ʔe=la ~ ʔe=ta*. ◊ Descriptions of deictic pronoun systems for most !Ui languages are highly inadequate, and textual examples are almost always ambiguous. On the whole, at least for "Narrow !Ui" (without ǁXegwi) it would make sense to reconstruct **a* as a "general" deictic stem, most commonly used to denote proximal deixis (see THIS), whereas distal deixis must have been denoted by using it as a base for various spatial particles — such as *|e* in |Xam and *ǁ̃e* in N|uu; the latter, as far as modern N|uu is concerned, undergoes irregular phonetic development in both its prepositional (*ǁ̃e-a → ke-a* with click loss) and its postpositional form (*ǁ̃e-a → ǁ̃a:* with vocalic contraction), cf. *ǂo: ǁ̃a:* 'that man' vs. *kea ǂo:* id. in Collins & Namaseb 2011: 35–36. However, it is unclear which of these variants — if any — is more archaic than the other; moreover, concern can be raised over their similarity with semantically identic morphemes in Khoe (*ǁ̃a-* 'that', **ʖ̃e-* 'this', see Vossen 1997: 377), possibly implying areal interference. The situation is also clearly different in ǁXegwi, where the "general" deictic stem *ʔe=* is combined with different morphemes (*=na, =la, =ta*) to form different (and semantically ambiguous) deictic pronouns. Keeping all this in mind, it is perhaps best to refrain at the moment from attempting to identify the principal morpheme(s) responsible for denoting the idea of distal deixis in Proto-!Ui.
- Nossob: (a) |'Auni *ha ~ he ~ hi*; (b) |Haasi *cɔ:-a*. ◊ Fluctuation in |'Auni may be due to contraction with various nominal class markers. Nothing is properly reconstructible for Proto-Nossob, since the attested morphemes are clearly different (moreover, textual evidence to support accurate semantic glossing is pretty much nonexistent).
- Taa: (?) ***tV-** (!Xóõ *tVʔV*, Kakia *ta-le*, N|uǁen *ti*). ◊ All contexts for Kakia and N|uǁen are highly dubious; as for !Xóõ, *tVʔV*, like **a* in !Ui, is more of a "generic" deictic stem than specifically 'this' or 'that' — in order to form distal deixis stems, it is usually extended with different additional morphemes (nominal stem *tVʔV=BVʔV*; adjectival or verbal stem *tV(ʔV:)-yà kV* 'there, that /proximate/'; adjectival or verbal stem *tVʔV:-sà kV* 'there, that /remote/').
- Tuu-: All data clearly show that expression of distal deixis widely fluctuates even within the small subgroups, let alone in between them. Nothing is properly reconstructible.

86. THIS [!Ui + Nossob] [- Taa]

- !Ui: ***a** (|Xam *a ~ a:*, ǁNg!ke *a*, N|uu *a*). ◊ This is the "generic" deictic pronoun which can, on its own, express proximate deixis at least in |Xam and in N|uu. In |Xam, this monovocalic pronoun sometimes changes to *e:* (most likely, reflecting contraction with a nominal class marker), but in N|uu, it seems to be the only equivalent for THIS. Whether it has anything to do with *ʔe=* in ǁXegwi-Z *ʔe=la* 'this' (= ǁXegwi-LH *ʔi=la ~ ʔe=la* 'this, that') is unclear.
- Nossob: (a) |'Auni *a*; (b) |Haasi *gʸa-ŋ*. ◊ The |'Auni form is clearly the same as the !Ui pronoun. Story's *gʸa-ŋ* 'this' for |Haasi is, however, quite mysterious (in any case, truly diagnostic contexts with an adjectival 'this' are not attested in Story's manuscript).
- Taa: ***tV-** (!Xóõ *tVʔV: ~ tV:ʔV ~ tánʔǹ*, Kakia *ti*). ◊ The simple stem in !Xóõ is also used as the basis for other deictic pronouns (see THAT).
- Tuu+: The isogloss between |Xam, N|uu, and |'Auni speaks in favor of **a* as the likeliest of all deictic pronominal stems to go back to the Proto-Tuu level. There is, however, a sharp divide in this respect between !Ui and Nossob, on one hand, and Taa, on the other, where the principal "general" deictic stem is **tV-*, without any parallels in the other two branches.

87. THOU [!Ui + Nossob + Taa]

- !Ui: ***a** (|Xam *a ~ a-á*, ǁNg!ke *a*, ǂKho-M *a*, N|uu *a*, ǁXegwi-Z *ʔa ~ ʔa-ŋ*, ǁXegwi-LH *ʔa-ʔe ~ a-ʔe*). ◊ The basic root shape, stripped of all additional markers (such as emphatic particles, etc.), is always **a*.
- Nossob: ***a** (|'Auni *a*, |Haasi *gʸà:=a*). ◊ The |Haasi form is listed with a prefixal emphatic particle.
- Taa: ***a** (!Xóõ *ãʰ*, Kakia *a*, N|u‖en *a ~ a-a*). ◊ Breathiness of the vowel in !Xóõ may be an original feature, in which case the reconstruction has to be amended to **aʰ*.
- Tuu+: ***a** is clearly reconstructible as the common, ubiquitously preserved root morpheme for the 2nd p. sg. pronoun in Proto-Tuu.

88. TONGUE [!Ui + Nossob + Taa]

- !Ui: ***|'ani** (|Xam *|'énni ~ |'érr̃i*, ǁNg!ke *|'ẽ*, pl. *|'e:n-yən*, ǂKho-M *|'an*, ǁKxau *|'anan-si*, N|uu *|'ãn ~ |'ãĩ*, ǁXegwi-B *|'ẽ*. ◊ The original bisyllabic stem shape is arguably best preserved in |Xam (*|'énni*, with vocalic assimilation) and in ǁKxau (where the strange form *|'anan-si* looks like a secondary singulative from a plural form, i.e. ← **|'ani-Vn*$_{PL}$*-si*$_{SG}$). In most other languages intervocalic *-n-* is lenited and reduced to nasalization of the vowel.
- Nossob: |'Auni *|'ãri*. ◊ Not attested in |Haasi, but cf. Xatia *|a: ~ |ãˢa* (Bleek 1956: 268; marking of pharyngealization is curious, but worth taking into consideration because of the Taa parallel). The |'Auni form looks inherited rather than borrowed from N|uu.
- Taa: ***ʔ|naˢn** (!Xóõ *ʔ|nàˢn*, pl. *ʔ|nàˢn-a ~ ʔ|nã̀:ˢ*, Kakia *|̃a:n*, N|u‖en *|'a:ni*). ◊ The N|u‖en form is slightly suspicious due to lack of nasality in the click efflux; perhaps this is really SVIa (Krönlein's N|usan, a dialect of |Xam) rather than SVI (N|u‖en)?
- Tuu+: All listed forms clearly belong together, although it is hard to say if **|'-* or **ʔ|n-* has to be reconstructed for Proto-Tuu (the first scenario would imply that nasality in Taa is secondary, probably through the influence of the nasal coda) due to lack of additional data.

89. TOOTH [[!Ui + Nossob] + Taa]

- !Ui: ***‖ʰai(-N)** (|Xam *‖ʰẽːi*, pl. *‖ʰe‖ʰẽĩ*, ‖Ng!ke *‖ãĩː ~ ‖ẽː ~ ‖ẽĩ*, pl. *‖eŋəŋ ~ ‖ẽĩŋ ~ ‖ẽ‖ẽ*, ǂKho-M *‖ẽĩ ~ ‖ẽĩ-si*, N|uu *‖ʰãĩ*, ‖Ku‖e *kx'e* pl., ‖Xegwi-Z, ‖Xegwi-LH *‖ʰi*, pl. *‖ʰi-ŋ*). ◊ The click onset is safely reconstructed as **‖ʰ-* based on the joint evidence of |Xam, modern N|uu, and ‖Xegwi. The coda presents more difficulties, with a unique correspondence series; however, ‖Xegwi offers a clue, allowing to assume **‖ʰai-* (→ ‖Xegwi *‖ʰi* in a regular manner) as the original singulative root and **‖ʰai-ŋ* as the old plural form, which became generalized as the singular in |Xam and N|uu and from which more innovative plural forms were later formed by various productive means.
- Nossob: (?) ***‖e-** (|'Auni *‖ẽĩ*, |Haasi *k'i=‖ɛ*). ◊ The |'Auni form, attested only in the early source Bleek 1929: 86, raises some doubts (it looks too suspiciously close to N|uu to be reliably recognized as inherited), but the |Haasi form is undeniably archaic, reflecting the original stem without nasality (just as in ‖Xegwi; *k'i=* is the productive prefix of plurality).
- Taa: ***‖qʰaN** (!Xóõ *‖qʰãː*, Kakia *‖xũ*, pl. *‖xaːni*, N|u‖en *‖ʻʌn-te* pl.). ◊ Nasality is always a part of the stem here, but (a) it is not clear if it is more appropriate to reconstruct **-ã* or **-an*, (b) it is highly probable that it was a class marker anyway (the word belongs to class 2 in !Xóõ, whose regular concord marker is *-ã*).
- Tuu+: ***‖qʰa-** is the likeliest reconstruction of the original root shape underlying all the attested reflexes in all three branches. Of note is that !Ui (at least ‖Xegwi) and Nossob (at least |Haasi) agree in reflecting the morphological shape **‖qʰa-i*, as opposed to Taa **‖qʰa-N*.

90. TREE [!Ui + Nossob + Taa]

- !Ui: ***ʘo** (|Xam *ʘ'ʰo*, ‖Ng!ke *ʘo ~ ʘoː ~ ʘ'ʰo*, ǂKho-M *ʘo* 'wood', N|uu *ʘoː* 'wood', ‖Kxau *ʘoː*, ‖Xegwi-Z *ʘo ~ ʘʰo ~ ʘʰoŋ*, ‖Xegwi-LH *ʘ̃òː-zì* 'tree', *ʘ̃oː* 'wood'). ◊ Attested variation between click effluxes is quite flabbergasting here; keeping in mind that the correlated Taa parallel begins with **ʔʘn-*, it makes sense to suggest that here, too, the original efflux was more complex than the simple velar release attested, e.g., in modern N|uu, but uniqueness of the correspondence series makes it difficult to propose anything with certainty. From a semantic / lexicostatistical perspective, it is important to note that in some languages, the stem is only glossed with the meaning 'wood', most notably modern N|uu, where, according to most sources, the common equivalent for 'growing tree' (or, specifically, for 'shepherd's bush /Boscia albitrunca/') is *ǂ'ʰiː ~ ǂqʰiː*. However, both internal and external data show that there is no reason not to project the common 'tree / wood' polysemy onto the Proto-!Ui level.
- Nossob: (a) ***ʘ̬o-** (|'Auni *ʘ̬waːa ~ ʘ̬waː-sa* 'wood, stick, tree', |Haasi *ʘ̬öi* 'wood, stick'); (b) |Haasi *ǂʰai* 'tree'. ◊ The situation in |Haasi seems to be more or less the same as in N|uu, either reflecting a contact scenario or the result of independent ("homoplastic") development. Interestingly, for |'Auni Bleek does not report any similar dichotomy between 'wood' and 'tree', despite it being in far more obvious contact with N|uu than |Haasi.
- Taa: ***ʔʘ̃a-** (!Xóõ *ʔʘnàye*, pl. *ʔʘnã̀ː*, Kakia *ʘ̃oe: ~ ʘ̃oi*, N|u‖en *ʘ'aː*). ◊ !Xóõ, most likely, preserves the original preglottalized nasal efflux, simplified (or mistranscribed) in the other two doculects.
- Tuu+: The original root for 'tree / wood' is probably to be reconstructed as **ʔʘ̃a-* or **ʔʘ̃o-*; the complexity of the click efflux would account for the variety of reflexes (sometimes real and sometimes the results of transcriptional inaccuracy) in all daughter languages other than !Xóõ.

91. TWO [-]

- !Ui: ***!uʔ-** (|Xam *!úː ~ !'uː*, ‖Ng!ke *!u ~ !'ú*, ǂKho-M *!'u*, N|uu *!'uː*, ‖Kxau *!'uː*, ‖Ku‖e *!'u*, ‖Xegwi-Z *kʰyũː*, ‖Xegwi-LH *k'uː ~ ɸ'uː*, ‖Xegwi-B *‖u ~ ‖'u*). ◊ Most of the attested reflexes would speak in favor of simply reconstructing **!'u* 'two' for Proto-!Ui. However, it is impossible to discount the recurrent fluctuation between simple velar and glottalized click effluxes in |Xam, ‖Ng!ke, and possibly ‖Xegwi as well; such correlated fluctuations are a rarity in old records and almost certainly indicate more complexity within the protoform. Provisionally, this is accounted for by the reconstruction **!uʔ-* (perhaps **!uʔu*?) with glottalization defined on the vowel rather than on the initial click; a contracted variant of this stem could easily result in a common development to **!'u* in the majority of attested dialects. Note that this is almost the same situation as in the case of the numeral ONE q.v.; this would lead to suggest that, perhaps, **-ʔV-* might have been some special morpheme employed in the formation of numerals (*both* numerals, as it seems likely that Proto-!Ui lacked separate lexemes for numerals higher than 'one' and 'two').
- Nossob: (a) |'Auni *|am*; (b) |Haasi *s=‖aː-maː*. ◊ The |'Auni form is a transparent borrowing from a Khoekhoe source. The |Haasi form is morphologically complex; initial *s=* is correctly identified by Güldemann (2002: 193) as a contracted form of the copulative element *si-* (as encountered in |'Auni *si |am*, etc.), but his proposal to identify *‖aːma(ː)* as a single stem may be questioned. In any case, it makes sense to suggest that |Haasi preserves the original Common Nossob stem for 'two', although it finds no etymological parallels in either !Ui or Taa.
- Taa: (?) ***ǂ̃um** (!Xóõ *ǂ̃ûm*, Kakia *ǃ̃um ~ ǁ̃um*, N|u‖en *ǃ̃um*). ◊ Correspondences here are almost the same as for ONE q.v., meaning that palatal **ǂ-* is the best, but not the only, bet for reconstruction (lack of transcriptional variants with *ǂ-* for Bleek's two doculects is puzzling).
- Tuu-: All three branches show separate equivalents for this numeral. Despite some phonetic similarity, Proto-!Ui **!uʔ-* and Proto-Taa **ǂ̃um* can hardly belong together (this would be possible if *!-* in most !Wi reflexes could be traced back to **ǂ-*, but modern N|uu and ‖Xegwi unambiguously speak in favor of original **!-*).

92. WALK (= GO)[5] [!Ui + Nossob] [- Taa]

- !Ui: ***‖'a** (|Xam *‖'a(ː) ~ ‖'aŋ ~ ‖'é*, ‖Ng!ke *‖'a ~ ‖'aː ~ ‖'ai*, ǂKho-M *‖a ~ ‖'a*, N|uu *‖'aʔa*, ‖Kxau *‖'a ~ ‖'aː ~ ‖'aŋ ~ ‖'a-i*, ‖Xegwi-Z *‖a ~ ka ~ ga*). ◊ Lack of glottalic articulation in ‖Xegwi-Z is somewhat puzzling, but the presence of variants *ka, ga* with click loss show that the word, in general, seems to be subject to irregular phonetic developments. Note that in various sources, the meaning 'go' is sometimes expressed by a different verb: |Xam *tàiˢ ~ tǎẽˢ*, ǂKho-M *tãĩ*, N|uu (W) *ʒaːˢn ~ ɸaːn*, (E) *ʒãĩˢ*, ‖Xegwi-Z *t'ãʔã-ne*, ‖Xegwi-LH *t'aʔa ~ t'aʔan*, reflecting Proto-!Ui **taˢ-* with different suffixes. General analysis, however, shows that the probable meaning for this verb in all !Ui languages is really 'walk' rather than 'go'.
- Nossob: (a) |'Auni *‖'a ~ ‖'e ~ ‖a ~ ‖aa*; (b) |Haasi *ǂa*. ◊ The |Haasi form must be an innovation, since the |'Auni form clearly belongs with !Ui. Cf. also |'Auni *tãĩ ~ tai ~ taãĩ*

[5] In accordance with the traditional practice of the Moscow School of comparative linguistics and the currently accepted standards in the Global Lexicostatistical Database, Swadesh's concept of 'walk' is replaced with 'go' (due to the latter's typically higher stability across the world's languages).

'to walk, to go' = |Haasi *tʸá-ai* 'to go'; external data suggest that this is really 'to walk' rather than 'to go'.

- Taa: ***sa** (!Xóõ *sâ:*, Kakia *ša*, N|u‖en *sa ~ ša*).
- Tuu+: |'Auni clearly aligns itself with !Ui here (and there are no significant arguments to assume that all of the attested forms are borrowed from N|uu), as opposed to Taa, where a possible etymological parallel to !Ui-Nossob **‖'a-* could be the !Xóõ verb *‖'a-* (*‖'âe*) with the specific meaning 'to go out hunting and/or gathering'. It is interesting to note that while Taa (or at least !Xóõ) seems to have a lexicalized opposition of **si* 'to come' vs. **sa* 'to go', !Ui and Nossob languages show free variation between these two stems, both in the meaning 'come'; not clear if this is coincidence or correlation.

93. WARM (HOT) [-]

- !Ui: (a) |Xam *k:áˢo ~ k:au:ˢ ~ k:auˢ-k:áuˢ* 'warm'; (b) ǂKho-M *hã:-i* 'hot (of sun, etc.)', ǂKho-D *há:ʔī* 'warm', N|uu *ha:ˢ ~ ha:ˢ-i* 'warm, hot (of weather)'; (c) ‖Ng!ke *‖oˢna*ˢ *~ ‖onà* 'hot (of sun)'; (d) ‖Xegwi-LH *kʰuru* 'warm'. ◊ Nothing is properly reconstructible here on the Proto-!Ui level, not just because every language (if not every doculect) seems to have a different equivalent, but also because semantic accuracy of the glossing usually leaves a lot to be desired, with the meanings 'warm' and 'hot' hopelessly entangled with each other.
- Nossob: Not attested in either |'Auni or |Haasi.
- Taa: (a) !Xóõ *kûbi* 'be hot, warm (e.g. sand, food, water)'; (b) Kakia *θwi* 'hot (of sun)'; (c) N|u‖en *‖k'u:*. ◊ Not reconstructible for Proto-Taa; same problems as with !Ui.
- Tuu-: The concepts of 'hot' and 'warm' are poorly documented and generally unstable, which requires us to exclude them from comparison.

94. WATER [!Ui + Nossob + Taa]

- !Ui: ***!qʰa** (|Xam *!wa: ~ !wá ~ !wã*, ‖Ng!ke *!ʰa: ~ !ʰa ~ !à:ˢ ~ !a: ~ ‖ʰa:*, ǂKho-M *!ʰa*, ǂKho-D *!ʰà*, N|uu *!qʰa:*, ‖Xegwi-Z *kʰa:*, ‖Xegwi-LH *qʰa:*, ‖Xegwi-B *‖ʰa:*). ◊ Most of the transcriptional variation in older sources probably reflects attempts to transcribe the initial click *!qʰ-*, explicitly attested in modern N|uu and evolving with perfect regularity to *qʰ-* in ‖Xegwi-LH (Ziervogel mistakenly transcribes the initial uvular as velar, whereas D. Bleek perceives it as a lateral click).
- Nossob: ***kʰa** (***qʰa**?) (|'Auni *kʰá: ~ kʰáá ~ kʰái*, |Haasi *kɑ̀*). ◊ It is important to mention another variant for |'Auni: *‖ʰà:a*, glossed as 'water, rain' (see RAIN). Since there is no serious evidence to suggest the existence of two etymologically different roots for these concepts in |'Auni, Bleek's spelling with a lateral click, just as it does in the case of ‖Xegwi, may actually reflect an initial *qʰ-*, in which case the Proto-Nossob reconstruction should be amended to **qʰa* (and a subset of uvular phonemes should be assumed for the protolanguage). Note also the peculiar transcription of the |Haasi form as *kɑ̀*, with a rare vowel that A. Traill identifies as a "clear low front vowel" (Story 1999: 15); one might wonder if this is in any way related to the supposedly uvular articulation of the preceding consonant.
- Taa: ***!qʰa** (!Xóõ *!qʰà:*, Kakia *!ʰá ~ !ʰa: ~ !xa:*, N|u‖en *!ʰa*).
- Tuu+: All the forms are obviously related, and matching data from !Xóõ and modern N|uu allow to reconstruct Proto-Tuu ***!qʰa** 'water' beyond reasonable doubt. Click loss in Nossob is quite exceptional in this case, but it must be noted that relatively few

items with **!-* in Proto-!Ui and Proto-Taa have reliably identifiable cognates in Nossob languages, and it cannot be stated with certainty that click loss is not regular here (particularly in conjunction with the uvular efflux).

95. WE [!Ui + Nossob + Taa]

- !Ui: (a) ***si** excl. (|Xam *s:i ~ s:i-s:i*, ǁNg!ke *si*, ǂKho-M *si ~ sa*, N|uu *si*, ǁKxau *si*, ǁKuǁe *si*); (b) ***i** incl. (|Xam *i*, ǁNg!ke *i*, ǂKho-M *i*, N|uu *i*, ǁKxau *ʔi*, ǁKuǁe *i*; ǁXegwi-Z *ʔi*, ǁXegwi-LH *ʔi-ʔe*). ◊ All "Narrow !Ui" languages show a clear-cut dichotomy between exclusive **si* and inclusive **i*; however, all attested doculects of ǁXegwi only feature *i* as the default 1st p. pl. pronoun with no regard to clusivity. External data confirm that the situation in ǁXegwi has to be treated as innovative.
- Nossob: (a) ***si** excl. (|'Auni *si ~ se ~ ci*, |Haasi *ci*); (b) ***i** incl. (|'Auni *i ~ e*, |Haasi *i*). ◊ For |'Auni, Bleek explicitly states the same exclusive / inclusive dichotomy as for !Ui languages. In Story's |Haasi manuscript, the difference between *ci* and *i* is never explained, but both forms are encountered in different contexts (cf. *ci ɑ̀ k'i=θwi:* 'we eat meat' vs. *i c'au* $k^{y}\varepsilon$ 'we milk them'), and it is highly likely that the situation here was exactly the same as in |'Auni.
- Taa: (a) ***si** excl. (Kakia *ši ~ šia ~ ša ~ si*, N|u‖en *si ~ si-sa*; cf. !Xóõ *īsî* 'we'); (b) ***i** incl. (Kakia *i*, N|u‖en *i*; !Xóõ $\bar{\imath}^{h}$ 'we' gen.). ◊ This is a rare case when data from older, generally less reliable sources come across as more important than data from Traill's well-curated description of !Xóõ: Bleek's Kakia and N|u‖en show the same dichotomy between exclusive and inclusive pronouns as !Ui and Nossob languages, whereas Traill's !Xóõ (the "Lone Tree" variety) shows no signs of it; instead, we find two forms, $\bar{\imath}^{h}$ and *īsî*, listed as synonymous. Of these, *īsî* almost looks like a collocation of **i* + **si*, though one might reasonably doubt the chances of such an odd formation (a general 'we' consisting of 'we incl.' + 'we excl.'?). In any case, bisyllabic *īsî* finds no parallels outside of Lone Tree !Xóõ and must be regarded as a likely innovation. Note that Maingard (1958: 106), in his own description of !Xóõ, finds exactly the same dichotomy as in the Bleek-described varieties of Taa (i.e. *i* 'we incl.' vs. *si ~ ši* 'we excl.').
- Tuu+: All three branches rather unequivocally suggest the necessity of reconstructing ***si** 'we excl.' vs. ***i** 'we incl.' for Proto-Tuu. This opposition is neutralized in ǁXegwi (generalizing the variant *i* for both purposes) and in one or more dialects of Taa, but remains stable everywhere else. Note that it is hard not to suspect a potential link between **si* 'we (excl.)' and the morphemic contrast between prefixal **sa=* 'we (incl.)' and **si=* 'we (excl.)' in Proto-Khoekhoe (Vossen 1997: 234); however, grammatical contrast between exclusive and inclusive markers is only typical of the Khoekhoe branch of Khoe, and is not formally reconstructible to the Proto-Khoe level, which would rather speak in favor of old Tuu influence on Khoekhoe than vice versa (provided this is not just a case of accidental resemblance).

96. WHAT [Nossob + Taa] [- !Ui]

- !Ui: (?) ***de**. ◊ It is not clear if **de* can be reliably reconstructed on the Proto-!Ui level specifically in the pronominal function of 'what?' rather than as just a general interrogative morpheme. It is in this latter function that it is explicitly encountered in |Xam: *c'a-de* 'what?', where *c'a* = 'thing', while *de* on its own is also encountered in other interrogative functions, e.g. 'where?', etc.

In N|uu, the situation is as follows: (a) for ǁNg!ke, Bleek lists the complex forms *dʸi-si ~ gi-si ~ kí-si ~ ʗi-si*, all of which are hard to reconcile with each other, but at least the first one definitely goes back to **di-si* ← **de-si* with assimilation, further cognate with the general interrogative *dʸe* 'where?' = |Xam *de*; (b) for ǂKhomani, Maingard (1937: 247) only lists *ʒi-si* 'what?' = Bleek's *dʸi-si*; (c) for modern N|uu, Collins & Namaseb (2011: 63) list two forms, *ɕui* (Sands quotes this as *ɕũĩ* with nasalization) and *ʒi-si*, as synonymous. The former would seem to be an innovation and may have developed out of *ɕu:* 'who?' (see below) with additional suffixation.

Notably, for ǁKxau Meinhof (1929: 169) lists *dɛ: ~ den* as the default equivalents for 'what', while the question 'where?' is actually expressed by the combination of this morpheme with other words, e.g. *ǁxa dɛ:* 'where?'. This is perhaps the strongest, if still not entirely sufficient, argument for reconstructing 'what?' as the original meaning for ***de**.

More problematic is the situation in ǁXegwi, where the only known form is ǁXegwi-LH *tʰĩ:* 'what?'. Even if the final vowel is assumed to be a suffixal extension (perhaps the same **-ĩ* or **-iŋ* as in N|uu *ɕũ-ĩ*?), phonetic realisation of the initial consonant as *tʰ-* rather than *d-* is surprising. On the other hand, we really do not know the regular reflex of Proto-!Ui **d-* in this language, so it is permissible to tentatively accept this stem as a genuine cognate.

- Nossob: |Haasi */ʰa ~ /i*. ◊ No information on |'Auni. Vocalic fluctuation in |Haasi is not explained, but may be of the same nature as in Taa (see below).
- Taa: !Xóõ */V ... èʰ*. ◊ A combination of the general interrogative particle */V* and the 3rd person singular / Class 3 harmonic pronoun *èʰ*. For Kakia and N|uǁen, no reliable data are available.
- Tuu+: On a purely formal basis, the isogloss between |Haasi */ʰa ~ /i* 'what?' and !Xóõ */V ... èʰ* id. allows to reconstruct ***/V*** 'what?' as the optimal candidate for Proto-Tuu. It must, however, be kept in mind that in !Xóõ, this is a general interrogative particle rather than the pronoun 'who?' proper; admittedly, the same concern may be raised over the status of Proto-!Ui ***de**. On the whole, the subsystem of !Ui interrogatives is clearly unstable and easily lends itself to various models of restructuring.

97. WHITE [-]

- !Ui: (?) ***!ui** (|Xam *!úi-tən ~ !úi-ta*; ǁKxau *!ui* 'white /of horse/'). ◊ This color term is highly unstable; most languages have their own equivalents, sometimes transparently borrowed (ǂKho-M *!'uɾi-ya*, N|uu *!'uri-a* ← Khoekhoe *!'uri* 'white'), sometimes without any etymological connections (ǁXegwi-Z *ša*, ǁXegwi-LH *ša:*). On an interesting note, ǁNg!ke */ɔ:wa* 'white' = |Xam */o:ˀwa ~ /kx'ó:ˀwa*, found only in W. Bleek's records and glossed as 'pale' or 'red' (Bleek 1956: 321, 339); if the latter glossing is not completely fortuitous, this might be the same word as Proto-Khoekhoe **/kx'aba* 'red', thus, yet another areal borrowing. The only word which looks potentially archaic is |Xam *!úi-tən ~ !úi-ta* 'white', further corroborated by its discovery in Meinhof's ǁKxau records; a rather weak link, but formally acceptable.
- Nossob: |Haasi *ǁa:*. ◊ Not attested in |'Auni.
- Taa: (a) !Xóõ *ʗ̃ú-ŋa* 'white', *kâ=ʗ̃úi-sà* 'whiteness'; (b) Kakia *ǁxwá*; (c) N|uǁen *!ari*. ◊ The N|uǁen item is clearly the same as !Xóõ *!aˀɽi* 'white' in Maingard 1958: 102, phonetically and semantically glossed as *!áˀli* 'whitish and shiny (silver, light grey, gold)' in Traill 1994: 75. The form in Kakia has no parallels. The bare root in !Xóõ is *ʗ̃ú-* (*-ŋa* is the same

adjectival suffix that is also seen in color terms such as BLACK and RED q.v.), but its reconstructibility for Proto-Taa is uncertain without reliable external cognates.

- Tuu: Unclear. It is highly tempting to trace Proto-!Ui (more accurately, |Xam-‖Kxau) **!ui* and !Xóõ *ǃú-* to the same source, but the click effluxes contradict each other; there are no signs of nasality in !Ui and no ways to explain its secondary origin in Taa. Given the overall lack of stability for this concept, it is perhaps best not to stretch the evidence here and leave all the slots empty.

98. WHO [-]

- !Ui: (?) ***tu** (‖Ng!ke *tú-e*, ǂKho-M *ɸʰu-xai*, N|uu *ɸu*, ‖Kxau *tu*, ‖Xegwi-Z *to:*, ‖Xegwi-LH *towa*). ◊ In most sources on N|uu, the interrogative pronoun *tu ~ ɸu* is always encountered only in strict conjunction with the general interrogative morpheme *xai ~ xae*, which typically accompanies other types of questions as well. This probably means that, from a historical perspective, *tu* can hardly be judged as an interrogative stem, and, in fact, its phonetic equivalence with Proto-!Ui **tu* 'men' (see notes on MAN above) cannot be a coincidence, especially considering that the same derivational model is also found in other Khoisan languages (e.g. Ju|'hoan *ha-žoe* 'who?' ← *ha* 'interr. morpheme' + *žu* 'person', etc.). On the other hand, the likelihood of **tu* being used at least as a significant *part* of the interrogative formation in Proto-!Ui increases with the addition of ‖Xegwi (Z) *to:*, (LH) *towa* (probably ← **tu-wa* or **tu-a* with a second component which could also go back to a general interrogative particle). In ‖Kxau, according to Meinhof, 'who?' is simply *tu*; theoretically, this could also be a contraction or morphological simplification from an earlier complex construction in which *tu* was only the first part.

  In stark contrast with this, the main equivalent for 'who?' in |Xam is *!u-de*, where *!u* = 'person' (see above) and *de* is either 'what?' or a general interrogative morpheme. This is a rather transparent derivation from the point of view of |Xam proper, and since it finds no correlations in other !Ui languages, it would be logical to regard it as a recent innovation.
- Nossob: (?) |'Auni *sa*, |Haasi *ci:*. ◊ For |'Auni, the only actually attested form is *sa-ka* 'whose?' (Bleek 1937: 197), where *-ka* is the productive possessive marker. It is not even clear if the |'Auni and |Haasi forms are related, although the correspondence *s-* : *c-* is quite regular; whether |'Auni *sa-ka* is assimilated from **si-ka*, or |Haasi *ci:* is the result of contraction (**sa-i* ?) remains a matter of pure speculation. In any case, both languages present evidence for some sort of **sV*-type morpheme as the principal carrier of the required meaning 'who?'.
- Taa: !Xóõ */V ... èʰ*. ◊ According to Traill's description, there is no difference in !Xóõ between the animate 'who?' and the inanimate 'what?'. Separate forms for 'who?' are not attested at all in Bleek's published records on Kakia and N|u‖en; it is, therefore, impossible to make any judgements on Proto-Taa.
- Tuu-: Nothing is reconstructible; all three branches have different morphemic strategies of expressing the required meaning, and all of them are just as volatile as in the case of 'what?'.

99. WOMAN [[!Ui + Nossob] + Taa]

- !Ui: ***|'a- ~ *|a-** (|Xam sg. */'a:i-ti*, pl. */á:-gən*, ‖Ng!ke sg. */'ai-ti ~ /'ai-ki ~ ǀai-ti ~ ǀai-ki ~ ǀe:-ki*, pl. */a-gən ~ /a:-gən*, ǂKho-M */'ai-ɸe ~ /'ei-ɸi ~ /ei-ki*, ǂKho-D */ēiɸī*, N|uu sg. */e:-ki*, pl. */a:-ke*, ‖Ku‖e */'a:-ti*, ‖Kxau sg. */a-ti ~ /a-u*, pl. */a:-kn*, ‖Xegwi-Z */a-zi*, ‖Xegwi-B */'a:-ze*). ◊ The plural stem of

this item is actually easier to reconstruct than the singular, since most languages unambiguously agree upon **|a-* as the root morpheme, followed by a nominal marker (**|a-kən* or **|a-ke*). The singular stem, however, shows far more variety. Perhaps the most archaic situation is preserved in |Xam, where the singular form may be analyzed as */'a-* (main lexical root, also seen by itself in such constructions as *!wi |'a* 'girl', lit. 'person-woman') + *-i* (a frequent nominal marker) + *-ti* (a rare singulative marker). If so, the paradigm has to be treated as originally suppletive, with **|'a-* 'woman /sg./' and **|a-* 'women /pl./' reflecting etymologically different roots whose phonetic similarity is accidental (or, if they are related, reflects some ultra-archaic phonetic alternation going beyond Proto-!Ui).

Such a clear contrast, however, is seen exclusively in |Xam, where transcriptions made by both W. Bleek and L. Lloyd consistently show */a-* for the plural form and */'ai-* for singular. Already in D. Bleek's transcriptions of ǁNg!ke we see variation between forms like */'ai-ti ~ |'ai-ki*, on one hand (which closely match |Xam), and *|ai-ti ~ |ai-ki ~ |e:-ki*, on the other, with a voiced velar efflux instead of glottalization; the same variation is seen in Maingard's transcriptions of ǂKhomani, whereas in modern N|uu the only attested variant is */e:-ki*, with the same simple velar efflux as in the plural form (however, the old contrast is still seen in the vocalism, with sg. */e:-ki* going back to earlier **|ai-ti*, while pl. */a:-ke* preserves the original pl. stem **|a-*). Other !Ui languages seem to behave randomly — thus, ǁKu‖e */'a:-ti* is recorded with a glottalized efflux, while Meinhof's ǁKxau */a-ti* also shows simple velar release.

For now, the most economical explanation for this confusion is to treat the |Xam situation as original and consider all the singular forms with /- rather than /'- in daughter languages as reformed by analogy with the suppletive plural. This still does not explain, e.g., the voiced release in such ǁNg!ke forms as *|ai-ki*, but the alternate solution — to provisionally set up a special click efflux in Proto-!Ui which would be reflected as this kind of alternation — would need plenty of additional evidence to be properly corroborated.

- Nossob: sg. ***|ẽ**, pl. ***|ã** (|'Auni sg. */ɛ̃:*, pl. */an*, |Haasi sg. */ĩ*, pl. *k'a=|ã*). ◊ In light of |'Auni pl. */an*, the reconstruction of the pl. form should perhaps be amended to **|a-n*, but it cannot be excluded that this transcription merely inaccurately conveys vowel nasalization. The paradigm is formally suppletive, but at least the click effluxes are always the same this time around. Comparison with !Ui evidence suggests that **|ẽ ← *|a-i(N)*. Note also the synonymous form *|e:-ki* 'wife, woman' in |'Auni — this is, in all likelihood, a borrowing from N|uu, resulting in yet another etymological "doublet" in |'Auni.
- Taa: sg. ***ǂâ-qáe**, pl. ***|aˢ-N** (!Xóõ sg. *tâ:=qáe*, pl. */ã:ˢ*, Kakia sg. *la=kai ~ la=kái*, pl. *la=ke ~ la=ké*, N|u‖en sg. */an*, pl. *|ãĩ*). ◊ The original situation is probably best preserved in !Xóõ, which features a suppletive paradigm: a special lexical root for the meaning is found in the plural (*|à:ˢ*), while sg. 'woman' is expressed with a compound formation (*tâ:* 'person' + *qáe* 'mother / female'). Exactly the same situation is observed, for instance, in the ǂHūa dialect as described in Westphal 1965: 139: sg. *la-qˢaé* vs. pl. *ŋ|ãʔa*. Other dialects of Taa, however, prefer to generalize the paradigm one way or the other — thus, Kakia reforms the plural form based on the singular, similar to (though not quite the same as) N|amani sg. *ta-qˢaé*, pl. *ta-qˢaé-tu* (also quoted from Westphal 1965: 139); in N|u‖en, on the other hand, it is the plural stem that seemingly becomes generalized[6].

[6] Additionally, cf. more recent material collected by Christfried Naumann (2011: 21): West !Xóõ sg. *sí=|ãˢn* 'woman', pl. *sí=|àˢã* vs. !Ama sg. *táá=qáé* 'woman', pl. *sí=|àˢã* 'women' (with the stem further enlarged with the frequent Taa nominal prefix *si=*).

- Tuu+: Oddly enough, the plural stem of this paradigm is more reliably reconstructible for Proto-Tuu than the singular — !Ui **/a-*, Nossob **/ã* and Taa **/aˀ-N* all belong together, although the absolute lack of pharyngealization in !Ui is somewhat puzzling. The singular stem, on the other hand, is either transparently innovative (Taa **ǂâ-qáe* 'mother-person'), derivable from the plural via morphology (Nossob **/ẽ* ← **/ã-i*), or questionable (does Proto-!Ui **/'ai-/ti/* historically contain the same root as **/a-*, or is this true suppletivism at work?). Importantly, Nossob languages clearly align themselves closer to !Ui than to Taa in this case; the alternation *e* ~ *i* in sg. vs. *a* in pl. perfectly correlates with the alternation *ai* sg. vs. *a* pl. in !Ui, reflecting either a shared !Ui-Nossob innovation or an archaic morphological model, lost in all Taa idioms.

100. YELLOW [-]

- This word is excluded from comparison, since it is very rarely attested, and most of the actual attestations are semantically ambiguous (e.g. ǁNg!ke */àˀla*; |Haasi *!a:*; !Xóõ has at least three possible equivalents — *ǂqʰúi* 'yellow, Naples yellow, yellow ochre, very light blue-grey, light green-grey', *ǂ̃àiˤʰ* 'yellow, chrome yellow, cadmium yellow, gold coloured', *ǁáʔu* 'yellow, chrome orange, terra cotta' — none of which have reliable external parallels).

References

Bleek, Dorothea F. 1929. *Comparative vocabularies of Bushman languages*. Cambridge University Press.

Bleek, Dorothea F. 1937. Grammatical Notes and Texts in the |Auni Language. In: J. D. R. Jones (ed.). *Bushmen of the Southern Kalahari*: 195–220. Johannesburg: Witwatersrand University Press.

Bleek, Dorothea F. 1956. *A Bushman Dictionary*. *American Oriental Series, vol. 41*. New Haven, Connecticut: American Oriental Society.

Bleek, Dorothea F. 2000. *The ǁN!ke or Bushmen of Griqualand West (ed. by Tom Güldemann). Khoisan Forum Working Paper No. 15*. Köln: University of Cologne.

Bleek, Wilhelm H. I., Lucy Lloyd. 1911. *Specimens of Bushman Folklore*. London: George Allen & Co.

Campbell, Lyle, William J. Poser. 2008. *Language Classification: History and Method*. Cambridge University Press.

Collins, Chris, Levi Namaseb. 2011. *A Grammatical Sketch of N|uuki with Stories*. Köln: Rüdiger Köppe Verlag.

Doke, Clement M. 1936. An outline of ǂKhomani Bushman phonetics. *Bantu studies* 10: 433–461.

Güldemann, Tom. 2002. Using older Khoisan sources: quantifier expressions in Lower Nosop varieties of Tuu. *South African Journal of African Languages* 2002(3): 187–196.

Güldemann, Tom. 2005. Tuu as a language family. In: Tom Güldemann (ed.). *Studies in Tuu (Southern Khoisan). University of Leipzig Papers on Africa, Languages and Literatures 23*: 11–30. Leipzig: Institut für Afrikanistik, University of Leipzig.

Güldemann, Tom. 2014. The Lower Nossob varieties of Tuu: !Ui, Taa or neither? In: Tom Güldemann, Anne-Maria Fehn (eds.). *Beyond 'Khoisan': Historical relations in the Kalahari Basin*: 257–282. Amsterdam: John Benjamins Publishing Company.

Güldemann, Tom. 2018. Historical linguistics and genealogical language classification in Africa. In: Tom Güldemann (ed.). *African Languages and Linguistics*: 58–444. Berlin: DeGruyter Mouton.

Kassian, Alexei. 2015. Towards a Formal Genealogical Classification of the Lezgian Languages (North Caucasus): Testing Various Phylogenetic Methods on Lexical Data. *PLoS ONE* 10(2): e0116950.

Kassian, Alexei, George Starostin, Anna Dybo, Vasiliy Chernov. 2010. The Swadesh wordlist. An attempt at semantic specification. *Journal of Language Relationship* 4: 46–89.

Lanham, Leonard, D. P. Hallowes. 1956a. Linguistic relationships and contacts expressed in the vocabulary of Eastern Bushman. *African studies (Johannesburg)* 15(1): 45–48.

Lanham, Leonard, D. P. Hallowes. 1956b. An outline of the structure of eastern Bushman. *African studies (Johannesburg)* 15(3): 97–118.

Maingard, Louis F. 1937. The ǂKhomani dialect of Bushman: its morphology and other characteristics. In: J. D. Rheinallt Jones, Clement M. Doke (eds.). *Bushmen of the southern Kalahari*: 237–275. Johannesburg: Witwatersrand University Press.

Maingard, Louis F. 1958. Three Bushman Languages. Part II: The Third Bushman Language. *African studies (Johannesburg)* 17: 100–115.

Meinhof, Carl. 1929. Versuch einer grammatischen Skizze einer Buschmannsprache. *Zeitschrift für Eingeborenen-Sprachen* 19: 161–188.

Miller, Amanda, Johanna Brugman, Bonny Sands, Levi Namaseb, Mats Exter, Chris Collins. 2009. Differences in airstream and posterior place of articulation among N|uu clicks. *Journal of the International Phonetic Association* 39(2): 129–161.

Naumann, Christfried. 2011. *A preliminary classification of Taa dialects*. Presentation given at: 20th International Conference on Historical Linguistics (ICHLXX).

Naumann, Christfried. 2014. Towards a genealogical classification of Taa dialects. In: Tom Güldemann, Anne-Maria Fehn (eds.). *Beyond 'Khoisan': Historical relations in the Kalahari Basin*: 283–302. Amsterdam: John Benjamins Publishing Company.

Sands, Bonny, Amanda Miller, Johanna Brugman. 2007. The Lexicon in Language Attrition: The Case of N|uu. In: Doris L. Payne, Jaime Peña (eds.). *Selected Proceedings of the 37th Annual Conference on African Linguistics*: 55–65. Somerville: Cascadilla Proceedings Project.

Starostin, George. 2015. *Annotated Swadesh wordlists for the !Wi group (Peripheral Khoisan family)*. Ms. Available online at: starling.rinet.ru/new100/kwi.pdf.

Starostin, George. 2021. Lexicostatistical Studies in Khoisan II/1: How to make a Swadesh wordlist for Proto-Tuu. *Journal of Language Relationship* 19/1: 99–135.

Starostin, George (ed.). 2011–2021. *The Global Lexicostatistical Database*. Moscow: Higher School of Economics / Santa Fe, NM: Santa Fe Institute. Available online at: http://starling.rinet.ru/new100/.

Starostin, Georgij. 2013. *Jazyki Afriki: opyt postrojenija leksikostatisticheskoj klassifikacii. Tom I: Metodologija. Kojsanskije jazyki*. Moscow: Jazyki slav'anskoj kul'tury.

Starostin, Sergei. 2000. Comparative-historical linguistics and lexicostatistics. In: Colin Renfrew, April McMahon, Larry Trask (eds.). *Time depth in historical linguistics*. *Vol. 1*: 223–265. Cambridge: The McDonald Institute for Archaeological Research.

Story, Robert. 1999. *K'u|^{h}a:si manuscript (MS collections of the Ki|^{h}azi dialect of Bushman, 1937)*. *Ed. by Anthony Traill. Khoisan Forum Working Paper No. 13*. Köln: University of Cologne.

Traill, Anthony. 1994. *A !Xóõ dictionary*. *Quellen zur Khoisan-Forschung/Research in Khoisan studies, Bd 9*. Köln: Rüdiger Köppe Verlag.

Traill, Anthony, Hiroshi Nakagawa. 2000. A historical !Xóõ-|Gui contact zone: linguistic and other relations. In: Herman M. Batibo, Joe Tsonope (eds.). *The state of Khoesan languages in Botswana*: 1–17. Gaborone: Basarwa Languages Project.

Vossen, Rainer. 1997. *Die Khoe-Sprachen: Ein Beitrag zur Erforschung der Sprachgeschichte Afrikas*. Köln: Rüdiger Köppe Verlag.

Westphal, Ernst O. J. 1965. Linguistic research in SWA and Angola. In: *Die ethnischen Gruppen Südwestafrikas (Wissenschaftliche Forschung in Südwestafrika, Bd 3)*: 125–144. Windhoek: Südwestafrikanische Wissenschaftliche Gesellschaft.

Ziervogel, Dirk. 1955. Notes on the language of the eastern Transvaal Bushmen. In: E. F. Potgieter (ed.). *The disappearing Bushmen of Lake Chrissie. Hidding-Currie publications of the University of South Africa, no. 1*: 34–64. Pretoria: J L van Schaik.

Г. С. Старостин. *Л*ексикостатистические исследования по койсанским языкам II/2: к уточнению филогенетической классификации языков семьи ту

Статья представляет собой вторую часть обширного исследования, основные цели которого — уточнение внутренней классификации южноафриканской языковой семьи ту (= южнокойсанской) и реконструкция максимально достоверного 100-словного списка Сводеша для пра-ту языкового состояния. Первый раздел статьи посвящен проблеме

ареальных контактов между двумя языками, относящимися к разным ветвям ту (нǁнг и ǀ'ауни), и отсеиванию вероятных заимствований из первого во второй для повышения точности результатов лексикостатистических подсчетов. За этим следуют собственно лексикостатистические матрицы и основанные на них схемы классификации для всей семьи, наглядно свидетельствующие в пользу исходно тернарного (а не бинарного) распада семьи на три ветви (!ви, носсоб и та); аргументы, ранее приводимые в поддержку более близких связей ветвей носсоб и та, признаются недостаточными. Статья также содержит ряд наблюдений над результатами реконструкции стословника для пра-ту состояния и отмечает некоторые любопытные особенности, отличающие лексикон пра-ту от его современных потомков. В Приложении приводятся подробные комментарии относительно возможности реконструкции на промежуточных уровнях (пра-!ви, пра-носсоб, пра-та) и на уровне пра-ту конкретных элементов из второй половины списка Сводеша.

Ключевые слова: южнокойсанские языки; языки ту; щелчковые фонемы; лексикостатистика; базисная лексика; ономасиологическая реконструкция.