

Lexicostatistical studies in Khoisan II/1: How to make a Swadesh wordlist for Proto-Tuu

The paper is the first in a planned two-part series, whose main goals are to conduct a general lexicostatistical survey of the Tuu, or South Khoisan, family of languages; to reconstruct a reliable approximation of the Swadesh wordlist for Proto-Tuu; and to clarify certain as of yet unresolved issues about the internal classification of Tuu languages. In the first part of the study, I survey the main data sources, identify the main obstacles to historical reconstruction in the Tuu domain, and make observations on some aspects of Tuu diachronic phonology. The main bulk of the paper is actually represented by the Appendix, in which I attempt to reconstruct the equivalents of the first 50 Swadesh list items for the three intermediate nodes of the Tuu family (Proto-!Ui, Proto-Nossob, and Proto-Taa).

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

Introduction

Of all the different linguistic lineages commonly united under the umbrella term of “Khoisan”, the Tuu family (originally = Dorothea Bleek’s “Southern Bushman” and Joseph Greenberg’s “Southern Khoisan”, see Güldemann 2005a) shares certain unique properties which simultaneously make it one of the most important and one of the most difficult groupings for any comparative-historical analysis of the Khoisan-speaking area. First, although the overall number of known Tuu languages is smaller than the respective number for Khoe (Glottolog, following Güldemann 2018, currently recognizes 8 different units¹ as compared to 13 for Khoe²), observed grammatical and lexical differences between these languages on the average exceed those observed between the various members of Khoe. Thus, lexicostatistical calculations show that, although the lowest observed percentages of matches within the Tuu family (e.g. 42% between !Xam and !Xóǀ) are comparable to the respective lowest percentages within Khoe (e.g. 41% between Nama and Kxoe), the internal branching of Tuu on the whole is deeper and more complicated than the internal branching of the two major subfamilies of Khoe (Khoekhoe and Kalahari Khoe; see Starostin 2013: 355, 407 for particularities). Among other things, this implies more possibilities for various important diachronic discoveries during the reconstruction of Proto-Tuu, hardly imaginable for Proto-Khoe because of the relatively young age of both of its constituent branches³.

¹ See glottolog.org/resource/languoid/id/tuuu1241.

² See glottolog.org/resource/languoid/id/khoe1241.

³ In fact, the divergence between some of the members of the Tuu family is so impressive that concerns have been voiced in the past about whether one may consider the common ancestry of all its members as an established fact (see e.g. Westphal 1962, 1971). As of today, however, there seems to be a general consensus among all specialists working on Khoisan that the observed phonetical, lexical, and grammatical correlations between all the small sub-branches of Tuu are best interpreted in terms of genetic relationship rather than contact (Güldemann 2005b). In this paper, I proceed from the assumption that this relationship has already been safely established and there is

Second, based on scrupulous phonetic documentation and phonological analysis of those Tuu languages which have survived into the modern age (namely, !Xóǝ and N|uu), this family has emerged as featuring one of the most complex sound systems in the entire Khoisan area. Thus, all known Tuu languages share no fewer than five types of click influxes, including the rare labial type θ (outside Tuu, it is only encountered in the †Hōǎ language of the Ju-†Hōǎ, or Kx'a, family), and at least some Tuu languages have more than 15 phonologically contrasting types of click accompaniments, a number unmatched by any Khoisan language outside of that particular family. Understanding the reasons which underlie this staggering complexity may provide an important insight into the evolutionary mechanisms of click systems in general, yet such an understanding is impossible to gain without a thorough diachronic study of the Tuu family as a whole.

On the negative side of things is the fact that, unlike Khoe, the Tuu family is nearly extinct. The only survivors, as has already been mentioned, are the small dialectal cluster of !Xóǝ (Taa) and N|uu, and even the latter is moribund and has, in fact, up until recently been considered completely extinct (Sands et al. 2007). All of the data that we have on the other languages come from older sources, stretching across about 150 years of ethnographic and linguistic research and widely varying in phonetic and semantic accuracy. Some of these data collections are quite comprehensive, such as Wilhelm Bleek's and Lucy Lloyd's archive on |Xam; other doculets are less lucky, being represented by ultra-short grammatical sketches and minimal wordlists. What is common for most of them, however, is the general unreliability of phonetic notation, grammatical analysis, and semantic glossing — implied by a lack of consistency between recordings of the same language by different researchers (quite often, even by the same researcher over an extended period of time) and by comparison with more recent and more accurate notations by newer and more experienced generations of scholars.

Similar problems are encountered with other Khoisan groupings as well, since data on both Ju (North Khoisan) and Khoe (Central Khoisan) languages often come from the same researchers as data on Tuu (Lucy Lloyd, Wilhelm and Dorothea Bleek, etc.). However, very few languages belonging to either of these stocks are *exclusively* represented by archaic and unreliable data; and even when they are, they usually have very close linguistic relatives with more recent and/or more accurate descriptions against which the questionable data may be cross-checked (e.g. certain Central and Southern sub-dialects of Ju against Ju|'hoan, or the extinct !Ora against its much more prominent neighbor Nama). By contrast, Tuu languages such as |Xam or ||Xegwi, while certainly not linguistic isolates per se, are still quite separate and distinct linguistic units, and cross-checking their data with, for instance, the modern phonetic and lexical descriptions of N|uu would be like trying to ensure the correctness of one's transcription of Czech or Polish by comparing it with Russian (while also having a very vague understanding of the historical phonology and lexicology of Slavic languages in general).

Subsequently, without access to more and better data (which is hardly likely, given the alleged extinction of most of those languages) our ability to properly and definitively reconstruct both the phonological system of Proto-Tuu and its lexical inventory is severely limited, and many problems will likely remain forever unresolved. Nevertheless, approximations are still possible, and any attempt to disentangle the complex web of genetic connections and areal interactions between Tuu languages and their other Khoisan neighbors is liable to shed at least some light on important diachronic processes, some of which may have even chronologically preceded the arrival of Bantu speakers into the area.

no need for special additional validation, allowing us to properly concentrate on issues of reconstruction and internal classification.

In this two-part paper, the next in an ongoing series on Khoisan lexicostatistics, I set up the challenge of conveying a full lexicostatistical survey of those Tuu languages which can actually be used for this purpose, as well as reconstructing Swadesh proto-wordlists for the three major linguistic clusters which constitute this family (!Ui, Nossob, and Taa) and ultimately for Proto-Tuu itself. A first attempt at Tuu lexicostatistics has already been published in Starostin 2013, along with provisional Proto-!Ui and Proto-Taa (but not Proto-Tuu) reconstructions for the 50-item “ultra-stable” subset of the Swadesh wordlist; this publication includes the revised and corrected results of that lexicostatistical survey and expands the reconstructions to include the Swadesh wordlist in its entirety.

The main bulk of both papers will constitute of appendices, containing specific comments on individual Swadesh items (due to volume limitations, the wordlist will be split in two). As for the theoretical parts, the first paper will briefly outline the data sources, the methodology, and the main issues concerning phonetic and lexical reconstruction; the second will deal with the actual internal classification of Tuu and give a brief analysis of the reconstructed proto-wordlists.

The data

Of the eight units currently listed in Glottolog as distinct Tuu languages, relatively complete Swadesh wordlists may be assembled for five, but their respective quality varies significantly depending on the age and thoroughness of the source(s). Additionally, while data from such languages as ||Ku||e, ||Kxau, and others are clearly insufficient to include them in any statistical calculations, they may still be relevant for purposes of etymological study and reconstruction. Below I list first the principal languages (and/or dialects) included in the statistical procedure, and then the list of auxiliary sources which will be consulted in the process of reconstructing wordlists for Proto-!Ui, Proto-Taa, and Proto-Tuu.

A. |Xam

Sources. This formerly widespread language became largely extinct even prior to the extensive field research of Dorothea Bleek in the first half of the 20th century; most of our knowledge on its grammar and lexicon comes from the archival records of Wilhelm Bleek and Lucy Lloyd, many of which were originally published in Bleek & Lloyd 1911 and later included into D. Bleek’s comparative dictionaries (Bleek 1929, 1956).

Dialects. Considering the overall expanse of the territories formerly populated by |Xam speakers and the fact that Bleek and Lloyd worked with a variety of informants (from Achterveld, Katkop, Strandberg, and other locations), dialectal diversity within the language must have been quite notable. However, precise differentiations are impossible without a meticulous study of the entire assembled text corpus. Lexicostatistical analysis of the data shows that there are relatively few Swadesh items transparently represented by two or more synonyms which could be thought of as representing different dialects; as for observed phonetic variation, it is not always clear when it should be ascribed to dialectal diversity or simply errors in transcription. For the purposes of the current study, we treat the entire Bleek-Lloyd corpus as a single “doculect”, while admitting that this is somewhat of a provisional simplification.

Quality. Transcription accuracy is always dubious, especially concerning the system of click accompaniments (see Traill 1995 for insightful comments on how to interpret various elements of Bleek and Lloyd’s transcription system for |Xam). Semantic glossing is frequently questionable as well, but at least in many cases it may be checked against the large assembled text corpus.

B. N||ng – N|uu

Sources. This is a large dialectal cluster which, unlike |Xam, is represented by several very distinct doculects from sources widely varying in space and time. This means that, for lexicostatistical purposes, it is possible and recommendable to build as many as three distinct wordlists: B.1 = ||Ng!ke (the dialect originally described by Dorothea Bleek; data published in Bleek 1929, 1956, and later separately in Bleek 2000), B.2 = †Khomani (the dialect originally described in Doke 1936 and Maingard 1937), B.3 = N|uu (the recently rediscovered variety spoken by several informants, with lexical data published in Crawhall 2004, Sands et al. 2007, Miller et al. 2009, Collins & Namaseb 2011; a complete Swadesh wordlist was kindly provided for the purposes of this study by Bonny Sands). For all of these dialects put together, we reserve the common name of N||ng as suggested in Güldemann 2017: 95.

Dialects. Unlike |Xam, the various attested dialects of this “macro-language” show quite a bit of lexical differentiation, though it is often difficult to tell how much of it is due to inaccurate semantic glossing, how much (especially in the case of N|uu) to very recent borrowings from other languages, and how much to gradual linguistic divergence after the original split of “Proto-N||ng”; for these reasons, as well as the relative incompleteness of the joint Doke/Maingard wordlist for †Khomani, any statistical discrepancies should be viewed with extreme caution.

Quality. Rather predictably, modern N|uu is one of the best transcribed representatives of Tuu; importantly, transcription quality in Doke 1936 and Maingard 1937 also seems superior to D. Bleek’s data (thus, both sources consistently mark the palatal click ʃ, which in most cases remains undistinguished from alveolar ! in Bleek’s records). Semantic glossing is assumed to be accurate for modern N|uu and can sometimes be checked against actual text examples for N||ng and †Khomani.

C. ||Xegwi

Sources. This language, geographically somewhat isolated from the rest of the !Ui continuum, is represented by at least three significantly different doculects, marked respectively as: (a) ||Xegwi-B — the earliest records collected by D. Bleek and published in Bleek 1929, 1956 (in her description the language is usually referred to as Batwa, a local Bantu term); (b) ||Xegwi-Z — as described by D. Ziervogel in a brief grammar sketch (Ziervogel 1955); (c) ||Xegwi-LH — as described by L. W. Lanham and D. P. Hallows in two short papers (Lanham & Hallows 1956a, 1956b).

Dialects. Judging by attested phonetic and lexical differences between the three doculects, a certain degree of dialectal diversity must have been present among ||Xegwi speakers. However, lexicostatistical discrepancies between the three sources are minimal (1–2 entries between ||Xegwi-Z and ||Xegwi-LH; slightly more between each of these and ||Xegwi-B, possibly because of less accurate semantic glossing in Bleek’s earlier records). Given the incompleteness of the sources (for ||Xegwi-Z and ||Xegwi-LH, data have to be extracted from grammar sketches and short text examples rather than actual vocabularies), it makes sense to merge them in one wordlist.

Quality. Transcription quality seems to be surprisingly adequate in the case of ||Xegwi-LH: for instance, Lanham and Hallows are among the first scholars to actually denote the presence of uvular phonemes and click accompaniments in any Khoisan language. Therefore, all data from ||Xegwi-B and ||Xegwi-Z, wherever possible, needs to be cross-checked against ||Xegwi-LH.

D. |'Auni

Sources. This language, which used to represent the westernmost spread area of Tuu, is known exclusively from records by Dorothea Bleek (Bleek 1937; lexical data also printed in Bleek 1929, 1956).

Dialects. Some dialectal variety may be identified from Bleek's records, as the equivalents for various meanings occasionally differ between the earliest ones, collected in 1911 and partially published in 1929, and the later ones, collected in 1936 and published in Bleek 1937 and Bleek 1956. It is, however, often difficult to establish whether these discrepancies (around 4–5 of them are found in items on the Swadesh list) represent true dialectal variation or inaccurate semantic glossing on the part of the researcher. Additionally, it is unclear if there are sufficient grounds to count the idiolect to which Bleek refers to as “Khatia” or “Xatia”, a very small amount of data for which were also collected by her in 1911 and published in Bleek 1956, as anything other than a minor sub-dialect of |'Auni. Finally, the occasional decision to regard |'Auni and |Haasi (see below) as dialects of a single language (e.g. in Glottolog 4.4) is hardly correct due to extremely significant lexical and grammatical differences between the two (e.g. around 20 mismatches on the Swadesh list).

Quality. Transcription quality is generally typical of D. Bleek's recordings for other Khoisan languages; external comparison raises serious doubts about the accuracy of click eflux transcription and slightly less serious about the same for click influxes.

E. |Haasi

Sources. This variety of Lower Nossob is solely known from records made by Robert Story of data from a single informant, Kabala (or Tatabesa), at the same Tweerivieren camp in 1936 where D. Bleek's data on |'Auni were collected; some of the |Haasi data were later published as part of Bleek 1956, but the complete manuscript did not officially see the light of day until Anthony Traill managed to rediscover and edit it for publication (Story 1999). Naturally, there is no dialectal variety to speak of here, but, as mentioned above, neither is there any reason to regard |Haasi as a bona fide “dialect” of |'Auni; both speech forms, as already noted by Traill in his preface to Story 1999, are more closely related to each other than to any other form of Tuu, yet both clearly have to be treated as different languages.

Quality. Although, in his own words, Story was a “complete amateur” and had no formal training in phonetics (Story 1999: 10), the overall quality of his transcription, at least at a rough glance, seems to be no worse than D. Bleek's or almost anybody else's at the time (e.g. he seems to have had a good ear for distinguishing between the palatal and alveolar clicks, with which quite a few other Khoisanologists seem to have struggled back then). The accuracy of his semantic notation can usually be confirmed by specific texts and phrases found in the manuscript. The worst problem is the scarceness of material: thus, as many as 40 Swadesh items cannot be recovered from extant data, which makes it impossible to offer reliable glottochronological datings for the moment of separation between |'Auni and |Haasi. That said, |Haasi data are of vital importance for attempting to at least partially reconstruct the basic lexicon of Proto-Nossob and, in turn, Proto-Tuu itself.

F. Taa (!Xóõ, Kakia, N|u||en)

Sources. Precisely three different varieties of Taa allow for the construction of more or less representative Swadesh wordlists. First and foremost among them is !Xóõ (more precisely, Lone

Tree !Xóõ) as represented in Anthony Traill’s now-classic and extensive dictionary of this particular dialect (Traill 1994, 2018). The other two are much older, dating back to D. Bleek’s brief research on the language of the “Masarwa” (a generic pejorative Bantu term for the San) of Kakia in 1913, and on the language of the N|u||en of Nausanabitz in 1920 (most of the data were subsequently published in Bleek 1929 and Bleek 1956). Both of these speech varieties seem to have become extinct and, to the best of my knowledge, are not directly identified with any of the still living dialectal varieties of !Xóõ (such as described, e.g., in Naumann 2014); concerning the latter, although some research has been carried out on them, no significant amounts of lexical data have been published to allow for a proper lexicostatistical comparison between them and Traill’s Lone Tree !Xóõ.

Dialects. Although all the three varieties of Taa for which it is possible to produce more or less complete Swadesh wordlists show up to about 20% of lexical discrepancies in these wordlists, which would, under normal circumstances, clearly speak of them as three different languages, the widely varying quality of recorded data does not allow to take these discrepancies at face value: Bleek herself admits that data on Kakia and N|u||en were collected in haste, and the probability of semantic and lexical inaccuracies in her records is fairly high. It is, therefore, possible that ultimately these two varieties are not nearly as distant from !Xóõ proper as are !Xóõ’s own 20 or so sub-dialects, tentatively classified in Naumann 2014 on the basis of some phonetic and grammatical isoglosses observed over the course of a general survey. In any case, at this time a detailed lexicostatistically based phylogeny of Taa languages and/or dialects is impossible due to lack of data; a tentative reconstruction of the Swadesh wordlist for Proto-Taa, based on all available evidence, is, however, somewhat within reach.

Quality. Lone Tree !Xóõ expectedly boasts the highest quality of phonetic (and probably semantic) accuracy among all South Khoisan languages, possibly second only to N|uu (for which, however, published data are far more limited) — all due to the extensive research of Anthony Traill. Nevertheless, the huge discrepancy between the quality of Traill’s data and everything else should not lead anyone into the fallacy of conflating Traill’s Lone Tree !Xóõ with Proto-Taa itself, at least not when lexical reconstruction is involved. In terms of phonetics, there is little, if anything, that data from Kakia or N|u||en could contribute in light of Traill’s clearly superior, and extremely detailed, description of Taa phonology (comparison with Bleek’s data shows plenty of unrecognized phonetic features and a lot of mistakes in the transcription of even the basic click influxes). But from a purely lexical point of view, there is no reason to *a priori* consider the Lone Tree !Xóõ equivalent for a particular meaning as more archaic than the corresponding Kakia or N|u||en equivalent whenever the two (or three) are clearly etymologically different from each other.

G. Other !Ui languages

Data from the following languages, unquestionably identifiable as separate linguistic units belonging to the !Ui group, may and should be used for etymological purposes (including reconstruction of Proto-!Ui basic lexicon) but is generally unusable for lexicostatistical goals, making a precise identification of their respective position on the !Ui tree somewhat difficult:

- (a) ||Kxau (small grammatical sketch, a few phrases, and a very short vocabulary in Meinhof 1929; all lexical data reprinted in Bleek 1956);
- (b) ||Ku||e (a small amount of lexical data collected by D. Bleek and published in Bleek 1956);
- (c) “Seroa” and “!Gǎ!ne”, both represented by short, old, and phonetically dubious collections of lexical data by T. Arbousset, C. F. Wuras, and H. Anders (all data reprinted in Bleek 1956).

Forms from some of these doculects will occasionally be quoted below, specifically as additional etymological support for particular reconstructions, but no systematic conclusions about their historical phonology or classification details shall be drawn.

Methodology

For the sake of this paper, I proceed from the following historical assumptions:

- (a) all of the languages listed above are genetically related within a single “Tuu” family;
- (b) all of those languages may be definitively and uncontroversially divided into no fewer and no more than three separate clusters — !Ui (|Xam, N!ng, ||Xegwi); Nossob (|’Auni, |Haasi); and Taa (!Xóǎ and all of its dialects as well as Bleek’s Kákia and N|u|en), each of these representing the result of divergence from its own intermediate protolanguage.

Convincing evidence for both of these assumptions, including (partial) regular phonetic correspondences and numerous sets of lexical and grammatical isoglosses, has been presented in numerous sources, from Bleek 1956 and Westphal 1962 to more modern research (e.g. Hastings 2001, Güldemann 2005b, Starostin 2008), and alternate scenarios, such as trying to explain similarities between !Ui and Taa as a result of areal convergence (a possibility not ruled out by such notable “splitters” in the field of Khoisan studies as E. O. J. Westphal), are unlikely and generally unwarranted.

What remains much less clear is the degree of relationship of these three clusters to each other, or even of some of the individual languages within these clusters to each other. While certain elements of consensus between the various classification schemes offered by researchers do emerge, such as, e.g., the understanding that |Xam and N!ng are closer to each other than to ||Xegwi, a particularly tricky issue rests with the |’Auni-|Haasi cluster, commonly referred to today as the “Lower Nossob”, or simply “Nossob”, languages. Here at least three conflicting schemes have been put forward:

- (a) E. O. J. Westphal (1971: 381) directly groups this cluster with the Taa languages, using the term “Taa” for the entire agglomeration; furthermore, as has already been mentioned, he has forever remained skeptical about the idea of a genetic connection between Taa and !Ui;
- (b) Oswin Köhler (1981: 469) counts the Nossob languages as a part of !Ui, considering them all related to Taa (which he calls “non-!Ui”) on a deeper level; this classification scheme has generally become more popular than Westphal’s until recently;
- (c) Tom Güldemann (2014) has partially reverted to Westphal’s model, arguing for a closer affinity between Nossob and Taa while at the same time not denying that both are ultimately genetically related to !Ui. His arguments are based on a number of lexical and grammatical isoglosses, as well as a strongly supported observation that the similarities between Nossob and !Ui are exaggerated because of extensive areal contact between |’Auni and N!ng (involving elements of bilingualism).

Out of these three, Güldemann is the only author who has actually published detailed linguistic argumentation in favor of his hypothesis, which may be one reason why it is currently accepted as the default phylogenetic scheme for Tuu in Glottolog. Nevertheless, due to the scarceness and sometimes dubious quality of the data, using selective lexical and grammatical arguments in this kind of linguistic investigation (the way it is done in Güldemann 2014) may not be totally free of bias, and it would be reasonable to take a more holistic approach to the matter, if at all possible. This is why an overall lexicostatistical survey, focusing on attested core basic lexicon for all the languages involved, would be a very useful addition to Güldemann’s methods of classification; and in the event of it producing different phyloge-

netic results from Güldemann 2014, analyzing the reasons for such a discrepancy could shed new light on both the historical relations between the various Tuu languages and the methodology of phylogenetic classification as a whole.

The actual results of an initial, preliminary survey based on 100-item Swadesh wordlists for all the languages listed above have already been published in Starostin 2013: 355; they showed that, although cognacy percentages between the Nossob languages and the various !Ui languages sometimes drop to around 46–48%, they are still consistently a little higher than the average percentages between Nossob and Taa, speaking in favor of Köhler's older classification rather than Güldemann's. However, there is a way to both correct and refine those results and make them more visually transparent by shifting from direct comparison of attested languages to comparing *reconstructed* wordlists — for Proto-!Ui, Proto-Nossob, and Proto-Taa, respectively. Condensing lexical evidence from a dozen languages into the shape of evidence from just three reconstructed proto-languages would be useful in helping clear away the “chaff” of identifiably recent innovations and borrowings, and would also make it easier to focus on the analysis of specific lexical isoglosses between the three branches in order to figure out which ones may have more weight for phylogenetic classification.

The general methodology for reconstructing proto-wordlists of the Swadesh type was already described in detail in several of this author's previous publications (Starostin 2013: 153–183, Starostin 2016) and, from a substantial point of view, needs no major modifications when applied to available Tuu material. Most of the specific challenges encountered along the way are of a technical nature — namely, scarceness and phonetic / semantic inaccuracies in the source data. These can sometimes be neutralized through careful scrutiny, but on the whole, of course, it should be well understood that the presented results are only as good as the data that currently support them, and are liable to change with each new significant publication of an additional data source (although, unfortunately, this is not likely for most of the languages involved in this study).

An important tripartite distinction could be introduced between *reconstructions*, *pseudo-reconstructions*, and *zero reconstructions* for each of the Swadesh items within each of the three subgroups. For the wordlist appendix below, the following rules are observed.

(a) A *reconstruction*, marked with an asterisk, is generated when cognates are attested in at least two separate doculects which do not represent close sub-dialects of a single language. In the case of !Ui, this means that the word has to be encountered at least in two out of three main clusters (|Xam; ||Ng!ke – †Khomani – N|uu; ||Xegwi), or, failing that, at least in one of them + one or more supporting languages whose data are not eligible for lexicostatistics (e.g. an isogloss between |Xam and ||Ku!e, or between N|uu and ||Kxau). Technically speaking, since |Xam and N!ng are closer to each other than to ||Xegwi, this does not allow to formally equate a “Proto-|Xam- N!ng” reconstruction with “Proto-!Ui” in the absence of a clear cognate in ||Xegwi; however, considering the scarceness of ||Xegwi data, we do not really have the luxury of downplaying |Xam – N!ng isoglosses, and for the sake of this particular phylogenetic study it seems reasonable to go along with a slightly broader understanding of “Proto-!Ui”.

Accordingly, in the case of Nossob languages “Proto-Nossob” is understood as the common invariant of cognates in |'Auni and |Haasi; in the case of Taa “Proto-Taa” is understood as an isogloss between !Xóõ and either Kakia or N|u!en (or all three).

(b) *Pseudo-reconstructions* can sometimes be substituted for actual reconstructions for both lexicostatistical and etymological purposes. Thus, if out of all the languages belonging to one of the three main subgroups of Taa, the Swadesh item in question is only attested in *one* language, and the form itself is not transparently identifiable as a recent morphological derivation or borrowing, there is a more-than-zero chance that it might actually be a direct reflex of the proto-

item (a very common situation for Nossob languages, where available data on !Auni are much more extensive than data on !Haasi, see ASHES, BARK, BELLY etc. below); naturally, this chance is increased even further if the form has credible external cognates in any of the other two branches.

If there are two or more non-cognate forms for the same equivalent in different languages and it is impossible to make a sound judgement on which one is the lexicostatistical archaism and which ones are the innovations, it is permissible to count them all as “technically synonymous” pseudo-reconstructions (see, e.g., BELLY or BIG in the !Ui list below), in the sense that each of them has a comparable chance of having expressed the required Swadesh meaning in the proto-language (this is more credible than the idea of “absolute” synonymy in the proto-language, with each daughter language retaining only one of the several earlier synonyms). Again, discovery of a potential cognate for one of these “technical synonyms” on the external level of comparison drastically increases its chances and almost (but not quite) raises the item’s status from pseudo-reconstruction to actual reconstruction.

(c) Finally, *zero reconstructions* — implying, among other things, that this particular item has to be excluded from lexicostatistical calculations — appear whenever the required item is either not found at all in any of the languages, or, if found in any of them, is transparently identifiable as a recent innovation or borrowing. In the list below, there are very few genuine zero reconstructions, since most of the Swadesh items are found to have some sort of equivalent in at least some of the discussed languages; the biggest problem is with a very small bunch of concepts whose “near-universality” does not properly apply to Tuu realities (e.g. FISH, notably absent in the area, or LEAF, seemingly a difficult concept for Tuu speakers which is usually expressed by borrowings).

Regarding the highest level of reconstruction (Proto-Tuu), we consider any Swadesh item to be formally reconstructible for Proto-Tuu if it is reconstructible in the exact same Swadesh meaning for both Proto-!Ui *and* Proto-Taa. The lower level reconstructions may be pseudo-reconstructions, i.e. an isogloss between !Xam and !Xóõ (or even an isogloss between !Xam and the far less reliably attested Kakia or N|u|len on the other end) may be taken as strong evidence for a Proto-Tuu reconstruction, unless there are additional obstacles to this interpretation (e.g. both forms may be easily interpreted as recent borrowings from a Khoe source). The Nossob languages, with their phylogenetic status not yet clearly resolved, are currently not very telling: it is extremely important to spot all the exclusive !Ui-Nossob and Taa-Nossob isoglosses, yet directly equating them with Proto-Tuu is impossible before the final conclusions on their position on the genealogical tree of Tuu languages.

With all possible Proto-!Ui, Proto-Nossob, Proto-Taa, and ultimately Proto-Tuu reconstructions on hand, the natural advantage is that it shall be much easier to not only calculate the distances between the specific branches, but also to analyze the possible classification alternatives in terms of individual shared archaisms and innovations, reducing the overall evidence to a small, but objectively attained, number of truly diagnostic etymologies. These results will be presented in the second part of the paper.

Notes on phonetic reconstruction in Tuu

Considering how much emphasis has been placed (and will continue to be placed) on the word “reconstruction” in this paper, some clarifications must be made about how we actually understand this term when applied to Tuu data. At the present state of our knowledge about Tuu languages as a whole, it is extremely difficult, if not downright impossible, to rigorously and rigidly apply the classic Neogrammarian methodology in order to elicit fully regular pho-

netic correspondences between the phonemic systems of these languages — mainly due to the relative scarceness of data from most of them, and to the generally poor transcription quality of those languages which are indeed represented by relatively large corpora (like |Xam or !'Auni). There is plenty of phonetic similarity between them, and there are enough recurring patterns of correspondences to usually (though not always) recognize etymological cognates, but a highly detailed system of correspondences which would fully cover all the subsystems of the complicated Tuu phonologies (click influxes, click effluxes, non-click consonants, vowels and their secondary features, tones, etc.) and reduce them to a parsimonious and typologically credible Proto-Tuu inventory at best requires a much huger research effort than is currently possible, and at worst may turn out to be objectively unreachable.

Nevertheless, even at this stage it is possible to operate on the level of what might be called “lax” reconstructions, along lines already suggested for Tuu languages in Starostin 2008, 2013. What this means is separating the phonological units of Tuu into categories which are found, based on comparative evidence, to be generally both more stable from a historical perspective and more consistently transcribed from a notational perspective — and those which seem to be more fluent over time, as well as less easily defined by inexperienced field workers. “Lax” reconstructions might then latch on to the more reliably established correspondences for the first category, while offering reasonable approximations (for instance, bluntly based on the majority rule) for the second. Such half-way reconstructions are always amendable if more high quality data come along or additional recurring patterns are confirmed statistically, but even without this they can still serve as proper historical evidence, provided that at least a certain “sound skeleton” has been recovered for them based on Neogrammarian-type correspondences.

According to my observations, the generally stable parts of phonological inventories in Tuu can be defined as (a) click influxes; (b) non-click consonants, especially in word-initial position; and, to a slightly lesser degree, (c) main root vowels (not including vocalic codas, correspondences between which are often chaotic, possibly because they represent variable morphological add-ons). The least stable parts, in addition to vocalic codas, are tones (if only because prosody is not marked consistently and reliably in any of the older sources) and click effluxes — which often show tremendous variation not just between different languages, but even between closely related dialects or sub-dialects of the same language. Below I adduce several important notes on each of these subseries, additionally referring the reader to my earlier and more detailed, but also sometimes outdated, observations on the comparative phonology of Tuu as published in Starostin 2008.

A. *Click influxes*. Correspondences between these segments are more often than not regular and trivial (one-to-one), but there are some important exceptions. The principal correlations are listed in Table 1; for some extra details (largely irrelevant when applied exclusively to the 100-item wordlist) see Starostin 2008: 365–370.

	Xam	N!ng	Xegwi	!'Auni	!Xóǒ
*θ	θ	θ	θ	θ	θ
*					
*ǰ	!	ǰ	λ / š	ǰ	ǰ
*!	!	!	∅	!	!
*					
*Ǿ	!		!	ǰ	ǰ

Table 1. Principal correspondences between click influxes across major Tuu languages

Notes.

(1) Labial click (*ǀ): see Starostin 2008: 366 on several examples where labial clicks in Taa may correspond to lateral clicks in !Ui, perhaps indicating secondary labialization. It is still unclear whether this correspondence is truly regular or if all the listed examples are just accidental resemblances; in any case, none of them are relevant to the data subset of the 100-item wordlist.

(2) Dental click (*ǃ): see *ibid.* on such specific correspondences as Taa *ǃq(?) = !Ui *c(?) and Taa *ǃn = !Ui *d. Examples for these are somewhat more reliable than for (1), but, once again, they are only encountered outside the Swadesh wordlist.

(3) Palatal click (*ǂ). This is the least stable of all click influxes in Taa, and it deserves more detailed commentary. First, in such extinct languages as !Xam and (maybe) some of the dialects of N!ng, such as the Bleek-transcribed ||N!ng!ke, it seems to have merged with the alveolar click (*ǂ → !), see below examples such as DOG, EAR, EGG etc.). The reason why I suspect it must have been a real diachronic development rather than a simple transcriptional error is that there are quite a few entries in !Xam which have been transcribed, both by Wilhelm Bleek and Lucy Lloyd, with an initial ǂ (cf. !Xam ǂenn 'to know', ǂa 'to kick', ǂxoa 'elephant' etc.), but many, if not most, of them look like relatively recent borrowings from a Khoe source⁴. This would imply that after the original palatal click had shifted to a different manner of articulation (perhaps merging with the alveolar click or becoming so close to it as to become indistinguishable for the early scholars of Khoisan⁵), it may have been reintroduced into the language/s/ along with lexical loans from their Khoe neighbors.

Second, in ||Xegwi the palatal click undergoes a unique development, shifting toward a non-click lateral affricate articulation. The regular development seems to be *ǂ → ǂ̣ (see DOG, EAR, EGG below), but occasionally post-alveolar fricative reflexes (č, š) are observed as well; this seems to happen when the click has a uvular efflux (cf. N!uu ǂqōē 'short' = ||Xegwi-Z čwe id.; N!uu ǂq^hoe 'wind' = ||Xegwi-LH šwe: id.). Unfortunately, scarceness of available ||Xegwi data prevents us from being able to fully describe the picture here, which must have been typologically somewhat similar to the well-studied behavior of palatal clicks in Eastern Kalahari Khoe languages (Vossen 1997: 285–288).

(4) Alveolar click and lateral click (*ǁ, *ǂ). Both of these are typically quite stable, but the alveolar click undergoes seemingly regular deletion in ||Xegwi as well (*ǂui 'person' → ||Xegwi *kwi*, etc.), again, parallel to similar developments in Kalahari Khoe.

(5) The “sixth click influx” (provisionally marked as *ǂ for lack of a better idea⁶). This reflects the unusual, but seemingly recurrent correspondence “!Xam ! : N!ng || : ||Xegwi ! : (?) Nosob ǂ : Taa ǂ”, established on the data of several basic items on the Swadesh list (BONE, ONE, RED, also FOOT in !Ui) as well as additional basic lexicon (e.g. the root for ‘female breast / milk’, listed in Starostin 2008: 368). The evidence for this extra influx is not overwhelming, but too strong to be brushed away as a mix of accidental lookalikes and incorrect transcriptions; in particular, given the regular deletion of the plain alveolar click in ||Xegwi, it is the only way to account for those cases in which ||Xegwi lexical items still feature the alveolar click (and cannot

⁴ Bonny Sands suggests that the loans may have come specifically from Korana (Sands 2014: 13).

⁵ In this respect, it may be instructive to recall a typological parallel in which the original Ju (North Khoisan) palatal click *ǂ has shifted to a retroflex articulation (!) in Ekoka !Xun (König & Heine 2001: 22–23), already after the original retroflex click *ǂ! had merged with lateral *ǂ in that same dialect. Could something of the sort actually have taken place in some of the now-extinct Tuu languages?

⁶ The symbol ǂ is actually borrowed from Clement Doke’s ingenious, but forgotten alphabet for click consonants, where it was reserved for the unvoiced alveolar click (now commonly marked as !).

be explained away as borrowings). Postulation of a phonologically distinct sixth click influx for Proto-!Ui and Proto-Tuu would make these protolanguages typologically unique (no living or attested extinct Khoisan language has more than five), but not theoretically impossible; more work on available material is necessary to understand whether the observed correspondence should be truly traced back to a separate phonological contrast or whether it may be explained by a conditioned split.

B. *Click effluxes*. Very few Tuu languages can be said to have adequate descriptions of their complicated click accompaniment systems. The best ones have arguably been produced by Traill for !Xóǒ (up to 19 different effluxes per influx), Miller et al. for N|uu (up to 10 different effluxes per influx), and by Lanham and Hallows for ||Xegwi (up to 7 different effluxes per influx). Even these descriptions may not be completely accurate and finalized in terms of recognized contrasts, and observed correspondences between different languages are by no means trivial.

Our current “lax” strategy on the matter is simple: for Proto-!Ui and Proto-Taa, unless there is a very strong individual argument about the secondary nature of these effluxes, we provisionally accept the efflux in N|uu and in !Xóǒ (respectively) as representing the proto-state — simply because any discrepancy between these languages and the earlier described ones may be theoretically attributed to incorrect transcriptions in older sources (where the same word may very often be found transcribed in multiple variants with different click effluxes). If this tactical decision somehow contradicts the majority rule, i.e., for instance, the N|uu click efflux is not the same as the efflux in the majority of other !Ui reflexes, such a situation deserves detailed individual analysis⁷.

C. *Non-click consonants*. A staggeringly low percentage of either Proto-!Ui or Proto-Taa Swadesh items are reliably reconstructible with a word-initial non-click consonant (approximately 14–15 items on the Proto-!Ui wordlist and 18–20 items for Proto-Taa), which goes to show how thoroughly integral click phonemes are to these languages (for comparison, the corresponding number for Proto-Ju, even though Ju languages have the second most complex inventory of click phonemes after Tuu, is no fewer than 35 items out of 100). This does not mean that the Proto-Taa system of non-click consonants was necessarily modest — Traill lists more than 40 such consonants for !Xóǒ, of which only very few can be reliably proven as secondary — but it does mean that the issue of an accurate reconstruction of this sub-system for Proto-Taa is not particularly relevant for our current task.

Phonemes encountered in basic lexicon items include *t- (HEAR, LIE), *k- (ALL), *s- (BITE, COME, FAT), the ejective velar affricate *kʰ- (DRINK), and the alveolar affricates *ʃ- (FLY) and *c'- (EYE), though for these last two phonemes evidence is more marginal and problematic. Correspondences for the others are largely trivial (arguably the most serious phonetic change is from *t- to palatal *c'- in N|ng), though see notes on BITE for a possible affricativization scenario for *s- in certain contexts. Not a single complex consonantal cluster, such as *tkʰ', etc., is reconstructible for this particular subset of the basic lexicon in any of the daughter branches of Tuu.

⁷ It should be kept in mind that click efflux articulation in Tuu, as well as other Khoisan languages, may occasionally be correlated with secondary features of vowel articulation, such as nasalization, pharyngealization, glottalization, and breathiness — both “genuinely” (if vocalic articulation exerts assimilative influence on the efflux, or vice versa) and “virtually” (if, in one of the less than accurate sources, a vocalic feature is transcriptionally mistaken for a back closure release, or vice versa). Unfortunately, secondary vocalic features are quite inconsistently marked in older sources.

D. *Vowels and codas*. Reconstruction of the Proto-!Ui, Proto-Taa, and especially Proto-Tuu systems of vowels and vocalic/consonantal syllabic codas is extremely difficult due to huge amounts of variation, which should be attributed not only to phonetic change (or pseudo-phonetic change, reflecting inaccurate transcription) but also to morphological variation, as the exact same nominal, adjectival, or verbal root may frequently be encountered in different languages (or even within the same language) in combination with different suffixal components – noun class markers, agreement morphemes, or various other clitical elements fused with the root and not recognized as separate morphemes.

The *main vowels* in Tuu languages, as follows from reliable modern data on Nluu and !Xóð, are typically restricted to three unrestricted phonemic units (*a, o, u*), occurring freely and frequently after any consonants; and two highly restricted units (front vowels *e, i*), whose occurrences after click phonemes are exceedingly rare, but who are somewhat more frequently met after non-click phonemes. The original picture may have been more complicated, as there are numerous cases in which the vowel *a* in Taa corresponds to either *e* or *o* in !Ui languages (see examples in Starostin 2008: 372); it is still unclear if such situations reflect additional original phonemes (such as **ε* and **ɔ*) or the results of phonetic contraction of different morphological variants (for a good example, see notes on FIRE below).

The precise inventory of Proto-Tuu *codas* (i.e. second morae of nominal and verbal word forms, which are often morphologically detachable even on the synchronic level, or may be shown to have been fossilized through external comparison) cannot be determined at the moment; on the whole, relatively few bimoraic sequences may be reliably reconstructed by comparing !Ui, Nossob, and Taa data. Given the fact that only !Xóð yields itself relatively well to detailed morphophonological analysis (in Nluu, most of the old derivational morphemes seem to have lost their productivity, and data on all other languages are antiquated and unreliable), reconstruction of nominal and verbal morphological elements for Proto-Tuu may turn out to be an even more challenging task than the reconstruction of its click system. Consequently, in the current paper, the emphasis is always on checking whether a bisegmental (initial click or non-click consonant + main vowel) sequence may be identified as the original root morpheme for Proto-!Ui, Proto-Nossob, Proto-Taa, and, ultimately, Proto-Tuu: by default, discrepancies between codas are provisionally written off as reflecting morphological variation, either already present on the Proto-Tuu level or arising independently in one or more branches after the split of the proto-language.

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 ($\tilde{\theta}$, \tilde{l} , etc.) while voiced clicks have a subscript tilde (θ , l , etc.).

(2) Affricates: instead of IPA's digraphic combinations, single letters are used to denote alveolar (*c, ʒ*) and palatal (*ç, ʝ*) 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 1–50)

In this Appendix, I list the results of intermediate and Proto-Tuu reconstructions for the first (alphabetically) 50 items on the Swadesh wordlist (more or less closely following the semantic specifications set out in Kassian et al. 2010). Each entry is structured as follows:

(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 \diamond sign separates listed data from comments on the reconstructions⁸. 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.).

1. ALL [!Ui + Taa] [-Nossob]

- !Ui: ***ku** (|Xam *ku*; ||Ng!ke *kwa*; Seroa *ku*). \diamond Attestation in ||Ng!ke is somewhat dubious (the word is only found in the earlier source Bleek 1929, not in Bleek 1956), but the |Xam entry is hardly questionable. Isolated equivalents (a) in N|uu: *huni-ki* (= #Kho-M *huni-ǂe*); (b) in ||Xegwi-LH: $\zeta'i \sim \zeta'i$ (only found as part of composite pronominal stems *i- $\zeta'i$* ‘we all’, *u- $\zeta'i$* ~ *u- $\zeta'i$* ‘you all’).
- Nossob: |’Auni *bà* (?). \diamond Cf. the example in Bleek 1956: 13: *tuku bà su !^hǂbati* “men shall all return”. Not clear if this semantic glossing should be trusted, especially given that the word *||ani* is also occasionally glossed as ‘all’, e.g. *ku totos ||ani* “all the people” (on the other hand, the primary meaning of *||ani* is probably ‘many’, cf. below).
- Taa: ***kU-ka**^f (!Xóǎ *kô: kàǎ*, Kakia *ku-ka*^f ~ *ku-ka*). \diamond Clearly a compound, but it is hard to delineate the individual meaning and function of each component. For N|u|en, the

⁸ These comments are sometimes identical with notes on specific items and reconstructions which have already been published as part of the !Ui and Taa databases at the Global Lexicostatistical Database (Starostin 2011–2021). However, the present paper also adds new details and observations that are relevant for reconstruction purposes, while at the same time omitting a large amount of synchronic information on the actual South Khoisan forms which may be found in the database notes.

only attested equivalent is *||árrri*, the same word as ‘many’ q.v.; it is impossible to tell if both meanings truly merged in the same word or if this is just a case of inaccurate semantic glossing.

- Tuu+: A clear isogloss between at least |Xam (+ Seroa) and Taa, allowing to reconstruct ***ku** as a basic Tuu morpheme for ‘all’. Specific evolution of this meaning in various daughter languages and its correlation with the related meaning ‘many’ may be obscured by inadequate glossing and insufficient contextual data.

2. ASHES [!Ui + Taa] [-Nossob]

- !Ui: ***!qui** (|Xam *!úi ~ !úí*, †Kho-D *!wí*, N|uu *!qui*). ◇ Not attested in ||Xegwi or any of the minor sources. Perhaps phonetically identical with |Xam *!ku:i* ‘to burn, smart, pain’ (Bleek 1956: 449), but without data from other sources it is premature to suggest semantic derivation (may simply be a case of homophony or close phonetic similarity).
- Nossob: |’Auni *!ʰana*. ◇ Phonetically similar to !Xóõ *||q’âna* ‘dirt, rubbish, rust’, but the click influx correspondence would be unprecedented (unless the |’Auni form is inaccurate).
- Taa: (a) !Xóõ *ƒòa*, Kafia *||wa;* (b) N|u||en *!’wi*. ◇ Technically, the form in !Xóõ is more reliable than item (b), and its distribution is confirmed by the parallel in Kafia (with a mistranscribed lateral click, cf. ‘bone’, etc., below). However, the word is also phonetically identical with the widely distributed Proto-Kalahari Khoe root **ƒoa* ‘ashes’ (Vossen 1997: 417), and the lack of etymological parallels in !Ui strongly suggests that we are simply dealing with one of !Xóõ’s many borrowings from the neighboring |Gui. In this light, the form attested by D. Bleek in N|u||en looks more trustworthy as a potential archaism.
- Tuu+: The isogloss between !Ui and N|u||en (West Taa) strongly suggests Proto-Tuu ***!qui** ‘ashes’, replaced in !Xóõ by a borrowing and not found in the Nossob subgroup.

3. BARK [-]

- !Ui: Not reconstructible. ◇ The word is attested consistently only within the N|uu cluster, where all forms are identifiable as borrowed from Kalahari Khoe (cf. PKK **||x’ũ* ‘bark’ in Vossen 1997: 421): †Kho-D *||x’ũŋ*, N|uu *||x’ũ:-si*. The only other known form is ||Ng!ke *!o;*, somewhat dubious because it is not backed by any textual examples.
- Nossob: |’Auni *||õ:*. ◇ This could actually be the same word as †Kho-D *||x’ũŋ*, etc., i.e. also a borrowing from Kalahari Khoe, although Bleek’s transcription of the click efflux (zero instead of expected *-x’-*) would seem to contradict this.
- Taa: (a) !Xóõ *gú-le* (pl. *gú-n*), Kafia *gu-le;* (b) N|u||en *!um*. ◇ The !Xóõ form has obvious parallels in the |Gui-||Gana cluster of Kalahari Khoe (*gure* ‘bark’ in Tanaka 1978: 10), but in this case, the word seems to be exclusive for that particular cluster rather than reconstructible for PKK, implying possible borrowing from Taa rather than vice versa.
- Tuu-: No proper isoglosses between the three clusters, and the word itself is formally not reconstructible. Its frequent re-borrowing from Kalahari Khoe indicates that the concept itself is not very stable in Tuu languages.

4. BELLY [-]

- !Ui: (a) |Xam *!áu-tu;* (b) ||Ng!ke *||x’ã;* N|uu *||x’ã*. ◇ In most languages, available data do not allow to perfectly distinguish between the meanings ‘belly’ and ‘stomach’, although

at least W. Bleek's notes on |Xam suggest that *!áú-tu* 'belly' may have been opposed to */oaʔ* 'stomach'.

- Nossob: |ʔAuni *!ʔai*.
- Taa: (a) *!Xóõ !ʰūma*; (b) N|u|en *!a:ban*.
- Tuu-: Not properly reconstructible. \diamond 'Belly' / 'abdomen' as a concept referring to the external part of the body seems to be fairly unstable in Tuu, with each individual language essentially having its own equivalent (this assuming that the semantic interpretation in older, uncheckable sources actually holds water). 'Internal belly' = 'stomach, bowels' is actually more stable: Proto-Tuu *|oa- can be reliably reconstructed based on the correlation between |Xam */oaʔ* 'stomach' and *!Xóõ !ōãʰ* 'innards, bowels, stomach' (cf. also, perhaps, ||Xegwi-Z *!ʰu-ga*: 'stomach', ||Xegwi-B */u-bwa* id., although the second syllable in each of these forms remains unexplained).

5. BIG [!Ui + Taa] [-Nossob]

- !Ui: (a) |Xam *!ui-ya*; (b) N|uu *!xo*; (c) ||Xegwi-Z *!xeya* ~ *!x'eya*, ||Xegwi-LH *!x'e*. \diamond Not properly reconstructible. The adjectival meaning 'big' in general is unstable and its equivalents seem to be easily reinvented from various verbal stems, e.g. |Xam *!ui-ya* is most likely derived from *!ui* 'to grow'. N|uu *!xo* is clearly the same as |Xam *!xo*: 'upright, tall', but this does not guarantee that 'big' was the original semantics.
- Nossob: *o-si ~ *u-si (|ʔAuni *ús* ~ *ú:si* ~ *ú:ši*, |Haasi *ǎ-si*). \diamond For |Haasi, Story also records usage of *!xwa*: 'big' as a free synonym; this may actually be a borrowing from N|uu.
- Taa: *!xa-(i): *!Xóõ !xa-*, Kafia *!xai*, N|u|en *!xai*.
- Tuu-: Since !Ui *o : Taa *a is a recurrent correspondence (possibly indicative of a special Proto-Tuu phoneme such as *ɔ), it is formally admissible to postulate Proto-Tuu *!xɔ- 'big' on the basis of the isogloss between Proto-Taa *!xa- and N|uu *!xo*: (+ |Xam *!xo*: 'tall?'). This is, however, not a perfect onomasiological match in light of the overall instability of the concept and uncertainties about specific semantic glossing in separate languages.

6. BIRD [!Ui + [Nossob + Taa]]

- !Ui: *|q^hui (||Ng!ke */wí* ~ */wi:*, †Kho-M */wi-si*, N|uu */q^hui-si*, ||Xegwi-Z *!ʰwi*). \diamond Reconstruction based upon the presumably accurate efflux transcription in N|uu. The original root seems to have been narrowed down to the meaning 'vulture' in |Xam (*/wi:*), whereas two innovative forms are attested for 'bird': (a) *x'arri* ~ *x'ānni*, usually glossed as 'little bird'; this is clearly related to Proto-Khoekhoe *k'ani 'bird' (ironically, itself ← Proto-Khoe *k'ani 'vulture', see Vossen 1997: 441) and likely reflects a recent borrowing; (b) *!errri-tān*- 'large bird', transparently derived from *!errri(ya)* 'feather' and thus also clearly innovative.
- Nossob: *si=|u (|ʔAuni *si=|u*, |Haasi *si=|ǎ*). \diamond Reconstruction of voiced efflux and *-u is provisional (largely based on external data). Initial *si=* is a nominal prefix commonly observed in other words as well (nothing to do with the copula *si*; possibly the same singulative marker as in N|uu */q^hui-si*, etc., only prefixed rather than suffixed?). Note that Bleek also lists |ʔAuni */o*: as a free synonym; this may be a prefix-less variant of the same stem, perhaps from a different dialect since in this shape, the word is actually phonetically closer to the variant in |Haasi.
- Taa: *|u(ʔ)- (*!Xóõ !ūʔu*, pl. *!ūʔā-tê*, Kafia *ši=|u*, N|u|en *si=|óu*). \diamond Kafia *ši=|u* is probably a typo for **ši=|u*. It is notable that Kafia and N|u|en both share the prefix *si-* with Nossob forms; *!Xóõ*, however, shows no traces of it.

- Tuu+: Nossob and Taa forms are pretty much identical. The question of how they tie together with !Ui *|q^hui is more problematic, but etymological identity is possible assuming that (a) *-i is a fossilized class marker and (b) the aspirated uvular efflux in N|uu is somehow correlated with strident vowel articulation in Taa (in any case, there are additional examples where uvular efflux articulation in !Ui correlates to a lack thereof in Taa, cf. ‘horn’, etc.). We may tentatively reconstruct Proto-Taa *|u^ʕ- or even *|q^(h)u^ʕ- to account for this alignment.

7. BITE [!Ui + Nossob] + Taa]

- !Ui: *c'i (|Xam c'i: ~ c'i:, ||Ng!ke ci ~ c'i, †Kho-M, †Kho-D c'i, N|uu c'i, ||Xegwi-Z ci, ||Xegwi-LH c'i:).
- Nossob: |Haasi c'i:. ◊ In !'Auni, the form c'i: is only attested by Bleek in the meaning ‘to ache’, but given that the polysemy ‘bite / ache’ is also attested in †Kho, it is possible that this was the proper equivalent in !'Auni as well.
- Taa: *siʔ- (!Xóõ siʔi, N|u|len ce-ya). ◊ The N|u|len form is the same as the !Xóõ variable stem siʔ-JV. Not attested in Kafia.
- Tuu+: *siʔi may be reliably reconstructed as the original root. !Ui and Nossob forms seem to share the phonetic shift *siʔi → *siʔi → *c'i, in which the intervocalic glottal stop fused with the word-initial sibilant and turned it into a glottalized affricate. In some of the varieties of Taa and Nossob, the same root also serves as the derivational basis for ‘snake’ (see below).

8. BLACK [-]

- !Ui: *!(^h)oe (|Xam !we:n ~ !wèŋ, ||Ng!ke !we ~ !oe, N|uu !^hoe, ||Xegwi-Z čwa ~ nčwa, ||Xegwi-LH žwa: ~ žwã:). ◊ The forms in |Xam and the N|uu cluster are clearly related (nasal coda in |Xam is likely of suffixal origin). Relation of these forms to ||Xegwi (n)čwa or žwa: is less certain, but a probable scenario is [1] regular deletion of initial *!- (*!oe → *koe) with [2] subsequent palatalization before a front vowel (→ *čoe) and [3] lowering of the diphthong (→ *čwa). For [1], see TAIL, TWO, WATER below; for [3], cf. ‘one’; no clear examples of [2], but no contradictory cases either. For now, we may count all these forms as cognates.
- Nossob: |Haasi //e. ◊ Not attested in !'Auni.
- Taa: *ʔaʔ- (!Xóõ ʔáʔ-ŋa, N|u|len ʔa-na). ◊ Same root in !Xóõ ʔā-be ‘black person’; -ŋa is a common adjectival suffix also encountered in other color terms (see RED, WHITE below).
- Tuu-: Not reconstructible. All three branches have their own equivalents.

9. BLOOD [!Ui + Nossob] [-Taa]

- !Ui: *|xau (|Xam ||xáú-ka ~ ||xáú-kən ~ ||xau-ki, ||Ng!ke ||xau, N|uu ||xau-ke). ◊ Solid reconstruction. Probably not related to ||Xegwi-LH λ'ēũ, which should reflect something like *ʔ'āũ, without any clear external parallels (phonetic similarity to Proto-Khoe *ʔao ‘heart’ is likely accidental, since click effluxes do not match and semantic connections between ‘blood’ and ‘heart’ are not particularly common in the Khoisan area).
- Nossob: *|xau (!'Auni ||xau(?)u, |Haasi ||xau).
- Taa: *|aʔ (!Xóõ ʔáʔ ~ ʔáʔm, Kafia ʔáʔa, N|u|len ʔa:ʔa). ◊ Provenance of -m in ʔáʔm is unclear; perhaps the result of morphological reanalysis of the plural form ʔáʔ-ma-tê.
- Tuu-: Not reconstructible. ◊ An obvious isogloss between !Ui and Nossob, on one hand, and Taa, on the other. It is curious that in both of D. Bleek’s dictionaries, she re-

cords a !Ui-like form for ‘blood’ for Kakia: $\|x\tilde{a}\tilde{u}$ (1929), $\|x\tilde{a}\tilde{u}^{\text{f}}$ (1956). However, it is not confirmed by textual examples, not distinguished semantically from $\tilde{!}a^{\text{f}}a$, and is clearly not the principal word for ‘blood’ in Taa as a whole. It may be a borrowing from some variety of !Ui (which is hard to confirm without a systemic analysis of the entire corpus) or, if it is some sort of archaic retention in limited (bound?) contexts, it could be a valuable indication that Common Taa $*\tilde{!}a^{\text{f}}$ is innovative.

10. BONE [!Ui + Taa]

- !Ui: $*\mathcal{C}^{(\circ)}a$ (|Xam *!wá*, ||Xegwi-LH *!a*). ◊ Correspondences between |Xam and ||Xegwi are non-trivial, but regular, reflecting the “sixth click” and extra labialisation in |Xam. A different equivalent for ‘bone’ is seen in the N|uu cluster: ||Ng!ke *||abba*, N|uu *||aba*, evidently cognate with |Xam *||abba* ‘a piece of eland’s bone that forms part of the completed arrow’. It must be noted, however, that Bleek transcribes the plural form for ||Ng!ke *||abba* as *||a||a* (reduplication is typically indicative of plurality), and that similar forms are also found in N|uu compound plurals, e.g. $\dagger q^ha$: *||ai-ke* ‘chest bones’. This may indicate suppletivism and preservation of the original equivalent for ‘bone’ in the collective / plural forms. Since the expected reflex of $*\mathcal{C}^{(\circ)}a$ in N|uu would indeed be *||a*, such a solution is quite likely; it must be noted, however, that *||a* and *||abba*, despite phonetic similarity, can hardly represent the same root due to having different clicks in |Xam.
- Nossob: Not attested in either |’Auni or |Haasi.
- Taa: $*\dagger\tilde{a}$ (!Xóõ $\dagger\tilde{a}$: /poss./, $\dagger\tilde{a}$: /alien./, Kakia *||a*, N|u||en $\dagger\tilde{a}$). ◊ Lateral click in Kakia is probably mistranscribed, as in many other similar cases.
- Tuu+: !Ui $*\mathcal{C}^{(\circ)}a$ and Taa $*\dagger\tilde{a}$ represent a solid etymological and lexicostatistical match; the only discrepancy is nasalization in Taa, which may ultimately go back to a suffixal extension ($*\dagger a\text{-}\eta$).

11. BREAST (CHEST) [!Ui + Taa] [-Nossob]

- !Ui: (?) $*\tilde{!}u\eta$ (|Xam *||wain-tu*; ||Ng!ke *||woen ~ ||woin-tu*; N|uu *||ũĩ-ɕu*, ||Ku||e *||ɕin-tu*). ◊ This word is clearly distinct from the word for ‘female breast / milk’ (|Xam *!hwai*, N|uu *||hũĩ*, etc.), and its semantic properties in individual languages are not always clearly distinguishable from those of close synonyms, e.g. |Xam *||axu* ‘chest’ (possibly the same as *||axu ~ ||ãxu* ‘side’) or N|uu $\dagger q^ha$: ‘sternum’, ‘breastbone’. The latter directly corresponds to ||Xegwi-Z/LH *ša-gu* ‘chest’ (see the exact same phonetic correspondence in ‘wind’ below), which makes the ||Xegwi form a less probable candidate for Proto-!Ui status (i.e. we reconstruct an original semantic opposition between $*\tilde{!}u\eta$ ‘chest’ and $*\dagger q^ha$ ‘breastbone’, with both meanings probably merged in one in ||Xegwi).
- Nossob: |’Auni *†an ~ †an*. ◊ Distinct from *||ēi-si* ‘female breast’.
- Taa: Kakia *||am*. ◊ All three varieties of Taa have different equivalents, of which !Xóõ $\tilde{!}u$: ‘chest’ (distinct from $\dagger q^h\tilde{e}$: ‘female breast’) phonetically coincides with Proto-Khoe $*\tilde{!}u$ ‘chest’ (Vossen 1997: 426) and quite likely represents a borrowing from the |Gwi-||Gana cluster; and N|u||en *||u* ‘chest’ is either a typo for $\tilde{!}u$ (see a similar case for BIRD above) or, alternately, could be compared with !Xóõ *||úí* ‘breastbone’. This leaves Kakia *||am* (distinct from *||xa:n-sa* ‘female breasts’) as the only form for which it is difficult to suggest a secondary origin.
- Tuu+: The correspondence between !Ui $*\tilde{!}u\eta$ and Kakia (= Proto-Taa?) *||am* is almost exactly the same as in the word for ‘liver’ (see below), likely reflecting Proto-Tuu $*\tilde{!}u\text{-}$

with different suffixal extensions (!Ui *-iŋ, Taa *-a/m). !'Auni *fan* is incompatible with this stem.

12. BURN (tr.) [?]

- !Ui: *||a (|Xam ||a(:) ~ ||e(:), ||Ng!ke ||a ~ ||e(:)). ◇ Although the forms by themselves are glossed as intransitive in Bleek's sources, textual examples clearly confirm transitive use as well, e. g. ||ōĩ *ŋe sa, ha ||a ŋ* "the sun comes, it burns me" (Bleek 1956: 545), etc. For modern N|uu, Sands et al. 2006 give *!xao* as the main equivalent; this root is listed for ||Ng!ke as *!xau* 'to kindle, make or light a fire', and its cognates in the Nossob languages have the same semantics (!'Auni *!xau* 'to light (fire)', |Haasi *!xau* 'to kindle'). Overall, the data are insufficient to reach a certain conclusion, but it is quite possible that this word, originally only taking 'fire' as its object, has widened its scope in N|uu. Examples of transitive usage also attested for ||Ku||e ||a: 'burn'. In ||Xegwi, ||a is only attested in the meaning 'to cook' by Bleek; no other equivalents are known for the meaning 'to burn' in available sources. On the whole, there are sufficient reasons to think that both 'burn (tr.)' and 'burn (intr.)' in Proto-!Ui were expressed by the same root *||a (|Xam and ||Ng!ke ||e represent a secondary morphophonological variant, probably fused with an agreement marker).
- Nossob: Not reconstructible. ◇ For !'Auni, the only attested form is *!á* 'to burn, light a fire, roast', with one accompanying example: *!á n !i* 'light the fire', meaning that the semantics could have actually been 'light, kindle'. For |Haasi, Story lists the form ||ɔ: 'to burn (tr.)', but it is not confirmed by textual examples — actually, the only textual example for this form is *!i ||a ||ɔ:* 'the sun is hot', which may, of course, be interpreted as 'the sun burns', but there is no explicit justification for this. Intransitive 'burn' = 'to be cooked' is actually attested as ||a (*θwi: k'i ||a ||a k'a* 'the meat is burning', *θwi: k'i ||a* 'the meat is cooked').
- Taa: Not reconstructible. ◇ The only solidly attested equivalent for 'to burn (both tr. and intr.)' is !Xóō *θ'á:*, clearly the same as Kafia *θwa ~ θwā* 'to make a fire' (e.g. *ši a θwa !a* 'we will light a fire'). Intransitive ||a 'burn' is also found in Kafia (*!a: wa ||a a* 'the fire is burning'); in !Xóō, however, the meaning of the cognate *||āha* is listed as 'set alight, set on fire, torch (e.g. tobacco, the veld, a hut), singe', indicating transitive use.
- Tuu-: Not reconstructible. ◇ Onomasiological reconstruction in this particular case is seriously hampered by what looks like incomplete and inaccurate semantic glossing in both older and newer sources of data, and by the difficulties in distinguishing between transitive and intransitive usages of verbal stems, as well as subtle semantic distinctions between 'to light, kindle' (= 'to make to begin to burn') and 'to burn down' (= 'to reduce to ashes by burning'). Clearly, the verbal root *||a is in itself well preserved in all three branches of Tuu, but whether it was indeed the basic equivalent of the meaning 'burn' in contexts like 'I burned [down] the house' remains unclear. For now, we should probably exclude this item from any calculations.

13. CLAW (= FINGERNAIL) [!Ui + Nossob + Taa]

- !Ui: *||qo-rV (|Xam ||ur(r)u, ||Ng!ke ||uri-si, †Kho-M ||oro(-si), N|uu ||qoro-si, ||Xegwi-B ||ɔla). ◇ The detachable origins of the second syllable are hinted at by plural forms in |Xam: *||u-||u-(t)tan*, where the original expression of plurality is represented by root reduplication. In all other languages, the nominal suffix of the singular form has completely

fused with the root. It must be noted that the form in ||Xegwi-Z is completely different: sg. *!ʼelo-lorj*, pl. *!ʼelo-le*, of unknown origin.

- Nossob: *!ʼAuni* *||ora-sa*. ◇ The situation in |Haasi is not clear. Story lists the form *kʼa=ǃü*, correctly identifies it as plural and further connects it with *ǃü* ‘finger’, which seems to be a phonetic variant of the same root. Further etymological connection of this word is clearly with *!Xóõ* *ǃü* ‘foot, spoor, track, hoof of an ungulate’ and its cognates (see FOOT below). Given that there are no textual examples confirming the semantics of ‘fingernail’, the glossing may very well be erroneous.
- Taa: (a) ***||qu-** (*!Xóõ* *||qû-le*, pl. *||qû-n-sâ*); (b) ***||aʔm** (*!Xóõ* *||aʔm*, Kafia pl. *!ʼam-te*). ◇ For *!Xóõ*, Traill lists two synonymous equivalents with the meaning ‘fingernail’ without specifying any semantic differentiations. The former is a perfect etymological match for Proto-!Ui ***||qo-rV** (right down to the detachable suffix of the sg. form), but is not supported by older data on Taa. The latter has no parallels in !Ui, but could be equated with Kafia *!ʼam-te* assuming that the dental click / here is a mistranscription for lateral // (these two symbols seem to be frequently mixed up in Bleek’s materials on Kafia). It must be noted that if *||aʔm* is analyzed as *||aʔ-m* (where *-m* is a fossilized plural marker, as in *ǃá*: ‘stick’, pl. *ǃá-m* id. and many other examples), the forms are comparable with Proto-Kalahari Khoe **||a* ‘fingernail’ (Vossen 1997: 436) and could be interpreted as old borrowings from a Khoe source, leaving only ***||qu-** as a viable etymon.
- Tuu+: The form ***||qu-rV** (where **-rV* is likely to have been a detachable segment, appearing only in sg. forms) may safely be reconstructed for Proto-Tuu based on equidistant evidence from all three branches.

14. CLOUD [-]

- !Ui: Not reconstructible. ◇ Each language or dialect cluster has its own equivalent: (a) |Xam *!waʔ:-gən*; (b) ||Ng!ke *tix-ke* (pl. form) = N|uu *ʒo:-si* (reflecting **to-* or **do-*); (c) ||Xegwi-B *||xe:ŋ* (dubious form).
- Nossob: Not reconstructible. ◇ *!ʼAuni* *!ʼhum-sa* cannot be compared with |Haasi *!al=xwai*; the second form is clearly of composite origin, but the two halves are not easily decipherable.
- Taa: Not reconstructible. ◇ *!Xóõ* and N|u||en employ different periphrastic expressions for the concept: *!Xóõ* *!qʰà:=[qʰūā*, lit. ‘water-hair’ vs. N|u||en *!xwe: ǃarri*, lit. ‘rain-sky’. Kafia *!wé* ‘cloud’ is unclear and without further connections.
- Tuu-: Not reconstructible. ◇ The generic concept of ‘cloud’ is clearly unstable in Tuu, although specific narrow terms denoting various types of clouds are encountered in *!Xóõ* (e.g. *qô*: ‘fairweather cumulus’) and other languages. This word should probably be excluded from comparison.

15. COLD [-]

- !Ui: Not reconstructible. ◇ Another unstable concept. In |Xam, no fewer than three equivalents are attested: (a) *xʼaoʔ* ~ *xʼáo* ‘cold’, (b) *sérri* ‘cool, cold’, (c) *||xwe*: ‘to be cold, become cold’. Attested examples are insufficient to draw clear semantic distinctions between these forms. In the N|uu cluster, most sources are in agreement on a single root, cf. ||Ng!ke *!ʰu*: = †Kho-M *!ʰu* = N|uu *!ʰū*: (but cf. additionally ||Ng!ke *si:-ya* ‘to be cold’, †Kho-D *kāriʔi* ‘cold’). ||Xegwi-Z *||keʔe* ‘cold’ contrasts with ||Xegwi-B *!xoa* ‘cold’. For most of these forms, it is hard to find etymological connections, but neither do they look like recent borrowings from Khoe or other sources.

- Nossob: |’Auni *ʃoʔa*. ◊ In the early source Bleek 1929: 29, *ʃxau* ‘cold’ is listed instead. Not attested in |Haasi at all.
- Taa: **ʃãʔũ* (!Xóõ *ʃãʔũ*, Kákia *ʃxʷe*, N|u|en *ʃkʷũ*). ◊ The phonetic discrepancy between Traill’s !Xóõ and Bleek’s earlier data is suspicious, but it may be argued that her transcriptions of ejective click effluxes actually reflect the same intervocalic glottal stop as in !Xóõ. In the case of Kákia, she glosses the word as ‘wind, cold’, but it is likely that she confuses here the reflexes of two separate roots, e.g. in *ʃi ia ti ʃxʷe*: “we are cold” the form *ʃxʷe*: = !Xóõ *ʃãʔũ* ‘cold’, but in *ʃxʷe*: *ʃxai* “a big wind” the form *ʃxʷe*: = !Xóõ *ʃqʰùe* ‘wind’ (see WIND).
- Tuu-: Not reconstructible. ◊ It is tempting to connect Taa **ʃãʔũ* with at least |Xam *ʃxwe*: (and possibly |’Auni *ʃxau* if this is indeed a real form), but the discrepancy in effluxes is disconcerting, with additional examples for such a correspondence being hard to find. In any case, since there are problems with confirming the archaic origins of *ʃxwe*: even on the !Ui level, this can hardly be counted as a lexicostatistical match on grounds of poor distribution.

16. COME [!Ui + Nossob + Taa]

- !Ui: **ʃi* ~ **ʃa* (|Xam *s:e*, ||Ng!ke *si* ~ *se* ~ *se-ya* ~ *sa*, †Kho-M *si* ~ *si-ya* ~ *sa*, †Kho-D *sī-yā*, N|uu *sa*: ~ *ca*:, ||Kxau *sa*: ~ *se*:, ||Ku||e *sa* ~ *si*, ||Xegwi-Z, ||Xegwi-LH *sa*). ◊ In addition to seemingly random variation of the root vocalism, some sources also register a glottalic articulation of the initial sibilant, e.g. |Xam *s:e* ~ *ssʰe* ~ *s:a*: ~ *ssʰa*: (W. Bleek). The reason for these variations is unclear; some of them may represent fusions of the root with agreement markers, but since few other verbal roots with codas in either *-a* or *-i* display so much variation, this is clearly not the only reason.
- Nossob: **ʃʰi* ~ **ʃa* (|’Auni *sa* ~ *sé* ~ *sí*, |Haasi *cʰi*). ◊ Nossob language data shows more or less the same variation as !Ui.
- Taa: **ʃi* ~ **ʃa* (!Xóõ *sî*, Kákia *si* ~ *ša*, N|u|en *sa* ~ *se* ~ *si* ~ *ša*). ◊ The precise !Xóõ forms are glossed as follows: *sî*: ‘come arrive’, *sî*: (var. form *sa-V*) ‘come to, come up to’. Cf. also *sî*: ‘go’.
- Tuu+: **ʃi* ~ **ʃa*. ◊ Vocalic variation in this root clearly goes all the way back to Proto-Tuu. If the (presumably accurately defined) situation in !Xóõ is deemed indicative, **ʃi* may be thought of as the original unbound form (infinitive, etc.) while **ʃa* would be the stem variant used in conjunction with agreement markers. Still, the general issue remains open.

17. DIE [!Ui + Nossob + Taa]

- !Ui: **ʃa* (|Xam *ʃa*, ||Ng!ke *ʃa*, †Kho-M, †Kho-D *ʃa*, N|uu *ʃa*, ||Kxau *ʃa*, ||Ku||e *ʃa*, ||Xegwi-Z *ʃa*, ||Xegwi-B *ʃa*). ◊ Apart from a strange lack of glottalic articulation in some of the attested varieties of ||Xegwi (cf. also ||Xegwi-LH *ʃa*: ‘dead’), all languages clearly reflect a single form **ʃa*.
- Nossob: |’Auni *ʃã* ‘dead’. ◊ |Haasi seems to have lost the old root, since Story only records *ʃʰo* ‘to die’, *ʃʰwa*: ‘dead’ — an innovation without a definitive etymology (cf., perhaps, !Xóõ *ʃʰũ* ‘be old?’).
- Taa: **ʃa* (!Xóõ *ʃã*, Kákia *ʃa* ~ *ʃa*, N|u|en *ʃa*).
- Tuu+: **ʃa*. ◊ This is one of the most stable and widely distributed verbal roots in Tuu.

18. DOG [!Ui + Nossob] + Taa]

- !Ui: ***ʃ^hu-** (|Xam *!wínj* ~ *!úinj* ~ *!^hwinj*, ||Ng!ke *!winj*, †Kho-D *ʃān*, N|uu (W) *ʃ^hun* ~ (E) *ʃ^hun*, ||Kxau *ʃ^huni*, ||Ku||e *!winj*, !Gā!ne *!inyi*, Seroa *kuenia*, ||Xegwi-Z *λwa* ~ *λweŋ* ~ *ɓwe*, pl. *λume*, ||Xegwi-LH *λ^hwinj*, pl. *λ^hu-minj*). ◇ Despite the seemingly chaotic array of reflexes, all forms are related. Original palatal articulation of the click is preserved in N|uu and confirmed by the regular development into a lateral affricate in ||Xegwi. Aspiration of the click is strongly confirmed by the same two languages (for ||Xegwi, only in the LH doculect, which seems to be more phonetically reliable than ||Xegwi-Z). As for the coda, most of the languages reflect the stem ***ʃ^hu-ni** (preserved as such in ||Kxau, losing the final vowel in N|uu, weakened to **ʃ^hu-ĩ* ~ **ʃ^hu-inj* in |Xam and ||Xegwi), but it seems that at least some dialects of ||Xegwi had different stem extensions (*λwa* ← **ʃ^hu-a?*).
- Nossob: ***ʃ^hɔŋ** (|ʔAuni *ʃɔ̃*, |Haasi *ʃ^hāŋ*). ◇ Reconstruction of the coda is highly approximate.
- Taa: ***ʃ^hq^ha-** (!Xóõ *ʃ^hq^hàì*, pl. *ʃ^hq^hà-ba-tê*, Kakia *ʃ^hxai* ~ *!xài* ~ *!ài*, N|u||en *ʃ^hi* ~ *ʃ^hĩ* ~ *ʃ^hxi*). ◇ The paradigm in !Xóõ shows that **-i* is a detachable class marker.
- Tuu+: ***ʃ^hɔ-**. ◇ All attested forms are related through recurrent correspondences (for N|uu *-^h* vs. !Xóõ *-q^h-*, see HAIR, TOOTH; rounded vocalism in !Ui vs. unrounded vocalism in Taa is very frequent, see BREAST, etc.). The Nossob forms with their nasal coda are notably closer to !Ui than to Taa; it seems that the full stem ***ʃ^hɔ-ni** separates !Ui and Nossob from Taa ***ʃ^hɔ-i**.

19. DRINK [!Ui + Nossob + Taa]

- !Ui: ***k^ha-** (|Xam *k^hwā* ~ *k^hwī* ~ *k^hwū*, ||Ng!ke *k^ha* ~ *k^hā* ~ *k^hē* ~ *||x^hā*, †Kho-M *k^hā* ~ *k^hēĩ*, N|uu *k^hāĩ*, ||Ku||e *k^hwā* ~ *||x^hwāĩ*, Seroa *k^hā*, !Gā!ne *k^ha*, ||Xegwi-Z pres. *k^hi*, past *k^ha*, ||Xegwi-LH *k^hēĩ*). ◇ Secondary labialisation in |Xam under unclear conditions, as in many other examples.
- Nossob: ***k^ha-** (|ʔAuni *k^hā* ~ *k^hē*, |Haasi *k^ha*).
- Taa: ***k^ha^h-** (!Xóõ *k^hā^h*, var. *k^ha^h-V*, Kakia *k^hā* ~ *k^hā* ~ *k^he* ~ *||x^hā*, N|u||en *k^ha-a* ~ *k^ha-u*).
- Tuu+: ***k^ha^(h)-**. ◇ An extremely stable basic verbal root, well preserved in every language. At least some of the attested variants, most notably **k^ha-* and nasalized **k^hā*, must go all the way back to Proto-Tuu where they may have been, as in !Xóõ, indicative of free and bound (“variable”) usage. Other variants (*k^hĩ*, *k^hē*, *k^hāĩ*, etc.) probably represent fusion with various auxiliary particles. It is worth noting that this root represents one of the best known isoglosses between Tuu and Khoe, cf. Proto-Kxoe **k^ha* ‘to drink’ (Vossen 1997: 497), but since in both cases the word is clearly reconstructible to the topmost level, no a priori judgment can be made on the direction of borrowing, or even on whether this is indeed a borrowing or a super-archaic retention from a common linguistic ancestor of both Tuu and Khoe.

20. DRY [!Ui + Taa] [- Nossob] (?)

- !Ui: (?) ***||o** (|Xam *||o* ~ *||ɔ* ~ *||ò*, N|uu *||o*). ◇ This concept is not too well attested for !Ui languages; additionally, it is not easy to distinguish between the required semantics of ‘dry = not wet (e.g. of clothes)’ and ‘dry = dessicated, dried up’. Still, such examples as *||ainj ||ai:e se ||ɔ*: ‘...so that the inside of the house may dry’ (Bleek 1956: 581) confirm reliability of the item in |Xam, and the overall reconstruction is based on the correlation between |Xam, N|uu, and ||Xegwi-B *||o*: ‘thirsty’ (‘dry’ is not attested for ||Xegwi, but the semantic shift or extension from ‘dry’ → ‘thirsty’ is trivial).

- Nossob: !'Auni $\|x'om$. \diamond Not attested in |Haasi.
- Taa: (a) !Xóõ $\|d\grave{o}$; (b) !Xóõ $\|úá^f$. \diamond Not attested in either Kafia or N|u|en. For !Xóõ, Traill lists two synonyms without specifying the distinctions. It may, however, be reasonably conjectured that $\|d\grave{o}$ is an areal root, since it is well attested across Khoe (Vossen 1997: 497), whereas for $\|úá^f$ no immediate source of borrowing can be detected.
- Tuu+: (?) $\|o^f$. \diamond Proto-!Ui $\|o$ and !Xóõ $\|úá^f$ are formally traceable back to a single source, although pharyngealized articulation in !Xóõ vs. lack thereof in N|uu is somewhat puzzling. The only attested Nossob form, !'Auni $\|x'om$, has no known parallels; if it is phonetically and semantically accurate (no guarantee), it can only be treated as an innovation of unknown origin.

21. EAR [!Ui + Nossob + Taa]

- !Ui: $\|fu-$ (|Xam $\|fu-ntu$, ||Ng!ke $\|we:(-ntu) \sim \|fu-ntu$, †Kho-M $\|fui(-si)$, N|uu $\|fui-si$, ||Kxau $\|we-ntu$, ||Ku||e de , ||Xegwi-Z $\|we$, pl. $\|fu-me$, ||Xegwi-LH $\|wĩ$). \diamond All languages show traces of the original root $\|fu-$ (click correspondences are regular; the development $\|f \rightarrow \|b$ in ||Xegwi is unique, but not contradicted by any other evidence, and ties in well with the general tendency of loss of palatal click articulation; ||Ku||e $d-$ is also a regular reflex of both the alveolar and the palatal clicks); codas are different across most of major dialect clusters, reflecting such morphological variants as $\|fu-ntu$ and $\|fu-i$.
- Nossob: $\|fu-$ (!'Auni $\|fui$, (?) |Haasi $\|k'u=\|fa-am$). \diamond The attested form in |Haasi contains the 1st p. possessive prefix $\|k'u=$ and the plural prefix $\|fa-$. The discrepancy in vocalism between |Haasi and !'Auni is more serious, but if the original suffixed stem was $\|fu-a$ (cf. Taa), elision of the labial element in such a complex form could be a possibility (hard to confirm or disprove). In theory, it would be possible to think of $\|fa-$ as the original root shape in Proto-Nossob assuming that !'Auni $\|fui$ is a form influenced by or directly borrowed from N|uu, but there is no conclusive evidence for such an assumption.
- Taa: $\|fu-$ (!Xóõ $\|fũá^h$, Kafia $\|wa$, N|u|en $\|fu-ša$, pl. $\|fu-i-te$). \diamond As in !Ui, all forms reflect the base root $\|fu-$ with different suffixal extensions ($\|fu-ā$, $\|fu-sa$).
- Tuu+: $\|fu-$. \diamond Although it is hardly possible to unambiguously reconstruct the original paradigm for this root, given the massive amount of variation across different lineages, all languages (with the possible exception of |Haasi) clearly show that $\|fu-$ was the original root. No specific morphological isoglosses across the three branches.

22. EARTH (= SAND) [-]

- !Ui: $\|q\tilde{a}\tilde{u}$ (|Xam $\|k'\tilde{a}\tilde{u}$, ||Ng!ke $\|a\tilde{u}$, N|uu $\|a\tilde{u}$). \diamond This concept is rather poorly attested in extinct languages and is usually not distinct from 'sand' (typically of the entire Khoisan area). At least the isogloss between |Xam and N|uu is reliable, though reconstruction of the click efflux is ambiguous (Bleek and Lloyd's transcription of the form with $\|k'$ suggests something other than a standard glottalized efflux — possibly a uvular release — but this seemingly contradicts the N|uu transcription which shows no signs of uvular articulation).
- Nossob: (?) $\|a\tilde{a}$ (!'Auni $\|á:a \sim \|a$ 'ground', $\|a:a$ 'dust'; |Haasi $\|aa$ 'ground'). \diamond Assuming that there is no lexical distinction between 'earth' (as substance) and 'ground' (as surface), which is a rather typical situation for San languages, we can tentatively set up $\|a\tilde{a}$ as the Proto-Nossob equivalent; reconstruction of the coda as $\|a\tilde{a}$ is confirmed by the transcription both in !'Auni (where it is reflected as variation between $\|a(:)a$ and $\|a$)

and in |Haasi (doubled *-aa*). Additionally, cf. |’Auni *!ãũ* ‘dust’, which could either be an archaic retention with a slight semantic shift, or a borrowing from N|uu.

- Taa: **ɬx’um* (!Xóõ *ɬx’úm*, Kakia *!um ~ ||um ~ ||k’om*, N|u|en *!om-sa* ‘ground’, *ɬum* ‘ground, sand’). ◊ We rely on the accurately transcribed !Xóõ form for the phonological reconstruction; Bleek’s transcriptions of Kakia and N|u|en probably reflect the usual inaccuracies characteristic of items with original palatal clicks.
- Tuu-: Not reconstructible. ◊ Each of the three main branches has its own equivalent. Of these, only Proto-Nossob **!aʔa* has a transparent external etymology in !Xóõ *!ãʔa* ‘below; to lower’, indicating that ‘ground’ (surface) rather than ‘earth’ (substance) was, after all, the original meaning in the Nossob languages.

23. EAT [!Ui + Nossob + Taa]

- !Ui: **ɬã* (!Xam *hã: ~ ha:*, ||Ng!ke *ã ~ ë ~ ëĩ*, †Kho-M *ã ~ ãĩ*, †Kho-D *ʔãĩ*, N|uu *ʔã*, ||Kxau *ʔa*, ||Ku|e *ẽ*, ||Xegwi-Z pres. *ʔĩ*, past *ʔã:*, ||Xegwi-LH *ʔĩ: ~ ʔiŋ*). ◊ Vocalic variation here is similar to the situation with COME q.v.; original root vocalism *a* is strongly suggested by external data.
- Nossob: **ɬa ~ *ɬã* (|’Auni *ã ~ hà ~ hàa*, |Haasi *à:*).
- Taa: **ɬã* (!Xóõ *ʔã:*, var. form *ʔa-V*, Kakia *ã ~ a: ~ e: ~ é:*, N|u|en *ã ~ ë*).
- Tuu+: **ɬã*. ◊ All languages preserve the original root. Nasalization must be reconstructed as an intrinsic property of the original root vowel: it is extremely frequent across all three branches, and emerges clearly in such diagnostic forms as, e.g., the !Xóõ nominalization *ʔã:-sà* ‘eating, food’.

24. EGG [-]

- !Ui: (?) **ɬaʔwi* (!Xam *!áúi ~ !àúwi ~ !k’áú:wi*, ||Ng!ke *!ʰãũ*, pl. *!wi-tən*, †Kho-D *ɬwi* ‘ostrich egg’, N|uu *ɬui*, ||Xegwi-Z *λwiŋ*, ||Xegwi-LH *λwĩ*). ◊ All of these forms are most likely related, since they all contain regular reflexes of the palatal click **ɬ-* (including the shift to a lateral affricate in ||Xegwi) and of the diphthong *-ui* (*-wi*). The overall shape of the root, however, is less clear. Perhaps the solution is hinted at by the quasi-suppletive paradigm recorded by D. Bleek for ||Ng!ke, which can be historically interpreted as going back to sg. **ɬaʔwi*, pl. **ɬui-ten* with contraction of the singular stem in a long plural form; in this case, *ɬui* in modern N|uu would be a back-formation from the original plural form. The form **ɬaʔwi* (← **ɬaʔbi?*) would also agree perfectly with |Xam data, as well as explain the variation between presence and lack of glottalic articulation in the different doculects of ||Xegwi. Still, the reconstructed shape remains speculative in the absence of similar corroborating examples.
- Nossob: (a) |’Auni *!ũĩ* ‘ostrich egg’; (b) |Haasi *k’ii*. ◊ The |’Auni form is likely related to !Ui **ɬaʔwi*, but it is unclear in which capacity — given the glossed semantics, and the glaring discrepancy with |Haasi, it could actually be a borrowing from N|uu (with misspelled click articulation). As for the |Haasi form, it has no external etymology at all.
- Taa: **ɬu-* (!Xóõ *ɬúú*, dimin. *kâ=ɬú:-bê*, Kakia *||wa:*, N|u|en *ɬwõĩ*). ◊ All listed forms are compatible, given how frequently the palatal click is transcribed as lateral or alveolar in Bleek’s Kakia and N|u|en records (see multiple other examples on this list); the basic root shape without suffixal extensions is **ɬu-* as seen in the !Xóõ diminutive form. It is not quite clear if !Xóõ *ɬú:*, pl. *ɬúã-tê* ‘empty ostrich egg’ is a phonetic variant of the same root (with a voiceless click) or a completely different etymon.

- Tuu-: Not reconstructible. ◊ Despite the obvious phonetic resemblance between !Ui (especially N|uu) and Taa forms, there is no easy scenario that would allow to reconcile them with each other (an original bisyllabic root like **ʃaʔbi* would not be expected to contract to **ʃu-* in Taa, since there are plenty of stems with the shape CV(?)bi in !Xóǝ).

25. EYE [!Ui + Nossob] [- Taa]

- !Ui: **c'a-xu* (|Xam *caxáú*, ||Ng!ke *cáxu* ~ *ca:xem*, †Kho-M *c'axau* ~ *c'axu* ~ *c'axəm*, †Kho-D *c'āxám*, N|uu *c'axam*, ||Kxau *c'axɔʔ*, ||Ku||e *caxu*, ||Xegwi-Z *sagu*, ||Xegwi-LH *c'agu*, pl. *c'a-ŋ*). ◊ Unique example of a bisyllabic stem in which the intervocalic consonant is not a resonant; this implies that the stem is historically a compound formation. The first root is unequivocally reconstructible as **c'a-* (most of the phonetically reliable sources mark glottalic articulation of the affricate); the second alternates between several variants (*-xau* ~ *-xu* ~ *-xam*), of which *-xu* is the most frequent one and is also often encountered as a nominal suffix in various words denoting surfaces (cf. in ||Ng!ke: *ʃa:xu* 'foot', *!a:xu* 'sky', *ʃa:xu* 'side'). It is very tempting to equate it with Proto-!Ui **xu* 'face' (|Xam *xú*, N|uu *xu* etc.), although this still leaves variants like *-xam* without a satisfactory explanation.
- Nossob: (?) **cxo* (|'Auni *cóo* 'eye/s', *c'a:xu(-ke)* 'eyes', |Haasi *cxɔ*, pl. *cxɔɔ*). ◊ The cluster *cx-* is extremely rare in |Haasi, making it all the more probable that the form *cxɔ* is contracted from an earlier bisyllabic form, clearly equatable with !Ui **c'axu*. If so, |'Auni *cóo* may further be regarded as its true cognate (with further simplification: **cxo* → *co*), while the doublet form *c'a:xu*, also attested by D. Bleek, could be interpreted as a re-borrowing from one of the dialects of N|uu (alternately, it could be an archaic preservation of the uncontracted form in some peripheral dialects or a higher register of the language, but this is unverifiable).
- Taa: (?) **ʃũ-* ~ **!ũ-* (!Xóǝ *!ũĩ*, pl. *!ũũ-tê* ~ *!ũũ-nî*, Kakia *ʃx'wĩ*, N|u||en *ʃũ*). ◊ Correspondences are unique: !Xóǝ clearly shows an alveolar click, whereas all other varieties of Taa speak in favor of palatal articulation (N|u||en in particular, but Kakia *ʃ* in Bleek's records very often transcribes an etymological palatal click, and almost never an alveolar one). This could be a serious argument for rejecting cognacy between !Xóǝ and Kakia-N|u||en; however, since there is no evidence in any of these languages for two different roots, and since 'eye' is typically one of the most stable items on the Swadesh list, it seems more prudent to admit the possibility of an irregular development in one of the two clusters (perhaps contamination with some other root).
- Tuu-: Not reconstructible. ◊ For this case, much rides on whether it is possible to demonstrate that !Ui **c'a-xu* (as well as Nossob **cxo*, which looks like a contracted variant of the former) is a compound form of secondary origin. While its composite nature is evident from its structure, the first component is not immediately identifiable, but it is phonetically and semantically possible to equate it with !Xóǝ *sàʔã* 'face, surface'. The optimal, though not the only possible, scenario here would be: (a) Proto-Tuu **saʔ-* with typologically common polysemy 'eye / face'; (b) Proto-Taa: **saʔ-* is retained in the meaning 'face, surface', but replaced by an innovation in the meaning 'eye'; (c) Proto-!Ui: **saʔ-* is replaced in the meaning 'face, surface' by the innovation **xu* (which has no cognates in Taa); the meaning 'eye' is eventually transferred to the new compound **saʔ-xu* → **c'a-xu*. Notably, such a scenario would hardly be compatible with the idea of a common ancestor for Taa and Nossob, but quite compatible with the idea of a common ancestor for Nossob and !Ui.

26. FAT (n.) [!Ui + [Nossob + Taa]]

- !Ui: ***so-** (|Xam *s:wéŋ*, ||Ng!ke *soa* ~ *süŋ*, †Kho-M *sōē*, N|uu *sun* (W) ~ *suŋ* (E), ||Xegwi-LH *swĩ*). ◊ Coda correspondences for this stem are extremely similar to the ones for DOG q.v., speaking in favor of reconstructing ***so-ni** for Proto-!Ui (original root vowel is **o* rather than **u*, which accounts for two different paths of assimilation: → **su-ni*, leading to ||Ng!ke *sü-ŋ*, N|uu *su-n*, or → **so-ne*, leading to |Xam *s:we-ŋ*, †Kho-M *sō-ē*). It is possible that ||Ng!ke *so-a* actually reflects the same root with a different suffixal extension.
- Nossob: ***so-** (|Haasi *cwa*). ◊ Not attested in !'Auni, but cf. *sā:a* 'fat' (adj.; polysemy 'fat /n./' : 'fat /adj./' is quite common for this word in Tuu languages). The |Haasi form regularly reflects **so-a* (with expected affricativization); !'Auni *sā:a* is somewhat strange due to lack of labial articulation, but there are no solid counterexamples for the potential change **-oa-* → **-a-*.
- Taa: ***sā^f** (!Xóō *sā^f*, Kakia *šā*). ◊ Note pharyngealized articulation in !Xóō.
- Tuu+: ***sɔ^f**. ◊ A clear isogloss between all three branches; vocalic correspondences between !Ui / Nossob, on one hand, and Taa, on the other, are recurrent, possibly reflecting **ɔ*. Pharyngeal articulation of the vowel in !Xóō may be archaic (it finds no correlation in N|uu, the only !Ui language where pharyngealization is marked accurately, but pharyngeal articulation seems to be prohibited in this language in structures like **CVn* ~ **CVɲ* anyway). Morphologically, the stem in Nossob seems to be closer to Taa than to !Ui (**so-a* or **so-ã* vs. **so-ni*).

27. FEATHER (= HAIR)

- It is preferable to exclude this word from comparison due to scant and dubious attestation. In both languages which have relatively modern descriptions (N|uu and !Xóō) the equivalent for FEATHER is the same as for HAIR q.v. In many others the word is not explicitly attested (||Xegwi; both Nossob languages; N|u||en), and those few equivalents which are distinct from HAIR are dubious (e.g. Kakia *dohé* 'feather' = !Xóō *dū^he* 'white ostrich plume' and may in reality be a more specialized term; |Xam *||erre* ~ *||árre* 'feather' is concurrent with FEATHER = HAIR and may actually mean 'wing' or a special type of feathers, etc.).

28. FIRE [!Ui + Nossob [+ Taa]]

- !Ui: ***'i** (|Xam *'i*, ||Ng!ke *'i*, †Kho-M *'i*, N|uu *'i*, ||Kxau *'i*, ||Ku||e *'e*, ||Xegwi-Z, ||Xegwi-LH *'i*). ◊ Lack of glottalized efflux in ||Xegwi is surprising — it is hardly a transcriptional error, being recorded independently in two doculects — but still probably secondary, given the overwhelming testimony of other languages.
- Nossob: ***'i** (!'Auni *'i*, |Haasi *'i*). ◊ |Haasi shows the same lack of glottalized efflux as ||Xegwi, but in this case it is not so surprising, since Story very rarely marks ejective articulation in clicks anyway (see PERSON, for example).
- Taa: ***'a-** (!Xóō *'ā*, Kakia *'ā* ~ *'a*, N|u||en *'ā*). ◊ Nasalization in the coda is of morphological origin (the word belongs to Class 2 in !Xóō, typically marked by nasal suffixes).
- Tuu+: ***'i**. ◊ Although all three forms are quite likely related, reconstruction of the original root vocalism poses problems due to discrepancy between !Ui-Nossob **-i* (quite unambiguous) and Taa **-a* ~ **-ã*. Purely phonetic reasons are out of the question here, since the correspondence is non-recurrent. From a morphological perspective, a scenario deriving **'i* from **'a-i* is not too likely, since there are numerous examples of

-ai ~ -ae diphthongs in !Ui languages, and it is not clear what might have caused such a tight fusion in Proto-!Ui. The most probable hypothesis, therefore, is that there is an underlying contraction in the Taa form: */i- (root) + -ã (class suffix) → */ã with elision of the original root vowel. Of note is the near-total identity of this item with Proto-Khoe */(a)e ‘fire’ (Vossen 1997: 435), but since in both cases the items are clearly traceable all the way back to the proto-language, no assumptions may be made at this point about the reasons for this similarity (ancient borrowing, common ancestry or even chance resemblance).

29. FISH [-]

- This word is excluded from comparison due to the near-total lack of the corresponding reality in the Tuu-speaking area and, subsequently, in Tuu languages as well. (Curiously, Doke records the form *ʃebē* ‘fish’ for †Kho-D, but it has no parallels anywhere and its origins are obscure).

30. FLY (v.) [!Ui + Nossob + Taa]

- !Ui: (a) */^h/au (|Xam *||au* ~ *||^hau* ~ *||^hóu* ~ *||xáu*, ||Ng!ke *||óu* ~ *||^hou*); (b) *ze^ʃ (N|uu *ze:ʃ*). ◊ Unfortunately, this word is not attested in many languages, which makes the situation difficult to resolve. On one hand, the isogloss between |Xam and Bleek’s records of ||Ng!ke is fairly strong, despite some phonetic problems (e.g. confusion about the click efflux), and speak in favor of an original !Ui root such as */^h/au. On the other hand, N|uu *ze:ʃ*, attested in a more modern variety of the N|uu cluster, is strongly confirmed as the original word for ‘fly’ by its external parallels in both Nossob and Taa. Unclear if the former is really some sort of secondary synonym (perhaps ‘to fly up, to rise’ as opposed to simply ‘to fly?’), or if, vice versa, the latter was somehow reintroduced into modern N|uu from an outside source (|’Auni?); better to take both as technical synonyms.
- Nossob: |’Auni *zé*. ◊ Not attested in |Haasi.
- Taa: *zōē^ʃ (!Xóō *zāi^{ʃh}*, Kafia *žōī^ʃ* ~ *žwe^ʃ*). ◊ The parallel between !Xóō and Kafia is straightforward enough, right down to pharyngealization, but the labial vs. non-labial articulation of the vowel is surprising.
- Tuu+: *z(o)e^ʃ. ◊ The parallel between N|uu *ze:ʃ*, |’Auni *zé*, and Taa **zōē^ʃ* hints at a common Tuu origin for all these forms, despite some issues with vocalism (particularly on the Taa side) and distribution (see notes on !Ui). There are further areal connections to Khoe, cf. the clearly related Naro *cāē^ʃ* ‘to fly’ (Visser 2001: 98); however, this Naro word has no further Khoe etymology, meaning that it might itself be of Taa origin (the Taa word seems to have also made it into †Hoan, cf. †Hoan *zòe^ʃ* ‘to fly straight’ in Honken 1988: 65).

31. FOOT [-]

- !Ui: *ča (|Xam *čwa*, ||Ng!ke *ča(-xu)*, ||Kxau *na-xu-ŋ* ~ *na-xu-si* ‘leg’). ◊ Forms in |Xam, “old N|uu” (Bleek’s ||Ng!ke), and ||Kxau (for which Meinhof lists the meaning ‘Bein’, but there is no separate ‘Fuß’) agree well with each other and point at a protoform with the “sixth click”. Other forms are less clear. For modern N|uu, Sands et al. 2006 list the form *!x’u:-ke* but warn that the recording may be inaccurate; this may, in fact, be the same form as *!u:-ke* ‘shoe’ (Collins, Namaseb 2011: 35), which also puts the semantics

in doubt. ||Xegwi-Z /^hiʔi = ||Xegwi-B /x'e 'foot' has no etymological parallels in !Ui (but see below).

- Nossob: *!Xai (|'Auni /x'ai, |Haasi n=!'ai). ◇ |Haasi n= is a pronominal prefix ('my'). Both forms are clearly related, but the click efflux is ambiguous, probably due to mistranscription in one out of the two cases, or in both.
- Taa: *!fu- (!Xóõ fū̃, Kákia fo ~ ʃo ~ ʃð, N|u|len fu). ◇ Also frequently encountered in the bisyllabic variant *!fu-ma, most likely a former diminutive: !Xóõ fū̃-ma-tê pl. 'feet', Kákia ʃu-ma ~ ʃu-mma 'foot', N|u|len fu-ma 'claws, little feet'.
- Tuu: Not reconstructible. ◇ Each of the three subgroups has its own root to denote the required meaning, with no obvious etymologies in the other ones. It is quite tempting to relate Taa *!fu- with !Ui *!ca, especially considering the labialization in |Xam ʃwa; however, its secondary nature is strongly hinted at by ||Ng!ke ʃa- and ||Kxau ɲa-xu-, and the correspondence !Ui *a : Taa *u finds no reliable confirming examples. The meaning 'foot' does seem to be generally unstable in Tuu; for a possible example of semantic shift, cf. ||Xegwi /^hiʔi 'foot' = !Xóõ /q^hi: 'to walk (pl.)' (very likely the same root, implying a nominalization in ||Xegwi).

32. FULL [-]

- !Ui: *!qauŋ (|Xam /áúin ~ /áúin ~ /áú:enyā, ||Ng!ke /xan, N|uu /qāi-ya). ◇ Not attested anywhere other than |Xam and the N|uu cluster. The root is verbal in origin ('to fill / be filled'). Protoform is approximate (*!q- reconstructed based on N|uu as well as occasional |Xam transcriptions with velar ejectives, e.g. /k'āú 'to fill'; coda *-auŋ could also be *-aiŋ, since labialization frequently turns out to be of secondary origin in |Xam).
- Nossob: |'Auni /x'an-si. ◇ Cf. also /āú 'to fill', which can only be related if the click efflux in one of the forms is mistranscribed. Not attested in |Haasi.
- Taa: *!um (!Xóõ /ú^hm, Kákia /úm, N|u|len /um). ◇ Another !Xóõ equivalent is the verb /ola, but the latter has no parallels in other varieties of Taa.
- Tuu: Not reconstructible. Each subgroup has its own equivalent for this concept.

33. GIVE [-]

- !Ui: *a (|Xam /á: ~ /à: ~ /ã: ~ /a-a ~ /a-ã, ||Ng!ke /a ~ /a:, N|uu /ã:). ◇ Although the verb is not attested beyond |Xam and the N|uu cluster, it is clearly the most basic equivalent for 'to give' in both of these nodes and is easily reconstructible for Proto-!Ui. Nasalization is infrequent and likely secondary; quality of root vocalism is notably stable. The only other language where the main equivalent for 'to give' is perfectly clear is ||Xegwi, cf. ||Xegwi-Z /sa, ||Xegwi-LH /sa ~ /s- (as in /in za s-e 'I will give'). It correlates with ||Ng!ke /sa: 'to bring, fetch' and with ||Kxau /ŋ-sa 'to give' (where /ŋ- may be 'me'); ultimately, all these forms can be explained away as originally causative formations from *sa ~ *si COME q.v.
- Nossob: Not reconstructible. ◇ The situation in Nossob languages is complicated. For |Haasi, the only recorded equivalent is the monovocalic verb /i; its cognacy with !Ui *a is not excluded, but given the total lack of vocalic alternations in this root in any of the !Ui languages, there is nothing to confirm it. For |'Auni, Bleek records (a) rather marginal /a 'to give', only attested in one or two dubious examples; may be a result of mistaken analysis or a N|uu form; (b) /ã ~ /ɔ ~ /o, most often used in an imperative function and consequently comparable with |Xam /ã 'let, give' (also typically an imperative). Because of this variety and the relative unreliability of Nossob data, it is better to exclude the word from comparison.

- Taa: ***!q^ha-** (!Xóõ *!q^hǎ:*, Kákia *!xa: ~ !xe:*). ◇ Apart from this autonomous verb, the meaning ‘give’ is also expressed in !Xóõ by the auxiliary “verb-postposition” *ǎ:*, e.g. *ǎa ǎ:* ‘to pass to, give to’, where *ǎa* by itself = ‘hold, grab, grasp’. This is probably the same word as N|u|en *ǎi* ‘give’, but its usage in !Xóõ makes it somewhat less eligible for inclusion (and there is no way to verify if it actually displaced ***!q^ha-** in N|u|en or just accidentally happened to be the only recorded variant for GIVE).
- Tuu: Not reconstructible. ◇ The only secure isogloss between more than one branch of Tuu is the auxiliary verb ***ǎ-**, commonly used in the imperative function (‘give!’) in some of the !Ui and Nossob languages, and in a postpositional function in !Xóõ. The principal indicative forms are, however, clearly different between !Ui and Taa, and somewhat obscure in Nossob. Available data do not allow to reconstruct a precise historical scenario.

34. GOOD [-]

- !Ui: Not reconstructible. ◇ Most languages have their own equivalents, including some roots of clearly non-!Ui origin (||Xegwi-Z *luga-ge*, probably a Bantuism; †Kho-D *kǎm-ǎé* ← Khoekhoe **kǎm* ‘right; true’) and some with very weak distribution (|Xam *a:-kǎn* ‘good’; |Xam *twá:i-ǎ* ‘good’; ||Ng!ke *kǎi ~ kǎiʔ*, N|uu *ǎǎ-kǎ* ← **tǎi ~ *tǎiʔ*).
- Nossob: Not reconstructible. ◇ Not attested in |Haasi. For !’Auni, Bleek records the variants *xwe ~ xwoi*, without any etymology.
- Taa: Not reconstructible. ◇ !Xóõ *ǎǎ* has no parallels in Kákia or N|u|en. Not attested in Kákia; N|u|en *ǎǎ* also has no etymology.
- Tuu: Not reconstructible. ◇ The meaning GOOD is clearly very unstable in Tuu; the concept itself seems rather diffuse, and the relations between all these forms and similar forms in other Khoisan groups may reflect a complex network of areal interaction (cf. Proto-Khoe **ǎǎ* ‘good’; †Hoan *qǎǎ* ‘good’, etc.).

35. GREEN [-]

- !Ui: Not reconstructible. ◇ Highly unstable and poorly documented meaning. In modern N|uu, expressed with *ǎǎo-a*, a borrowing from Khoe (cf. Nama *ǎǎo* ‘to turn green; to grow’).
- Nossob: Not reconstructible. ◇ Not attested in !’Auni. |Haasi *ǎǎu* is probably borrowed from the same Khoe source as the N|uu item.
- Taa: ***ǎǎiʔ-** (!Xóõ *ǎǎiʔ^h*, Kákia *ǎǎiʔ*). ◇ Cf. !Xóõ *ǎǎiʔ^h-sǎ*, pl. *ǎǎiʔ^hm-sǎ* ‘dung beetle’, most likely containing the same root; morphological structure of the noun suggests that *-i-* is an original class suffix and that the semantics of ‘beetle’ might be primary. One might also speculate about further links with **ǎǎana* (← **ǎǎna* ?) LEAF q.v.
- Tuu: Not reconstructible. ◇ The concept is generally unstable, not very well documented, and most of the languages have their own ways of expressing it.

36. HAIR [!Ui + Nossob + Taa]

- !Ui: ***ǎǎu** (|Xam *ǎǎ(-kǎn) ~ ǎǎú(-kǎn)*, ||Ng!ke *ǎǎ ~ ǎǎú*, †Kho-M *ǎǎ ~ ǎǎu*, N|uu *ǎǎu:-ke*, ||Xegwi-Z *ǎǎu:-zi*, ||Xegwi-LH *ǎǎú*). ◇ A super-stable word with fairly transparent phonology, though the aspirated articulation of the efflux tends to go unnoticed in older transcriptions.
- Nossob: ***ǎǎo** (!’Auni *ǎǎo*, |Haasi *ǎǎ*). ◇ Aspirated articulation explicitly marked by Bleek for !’Auni, but noticeably absent in |Haasi.

- Taa: ***q^hu-** ($q^h\ddot{u}\ddot{a} \sim |q^h\ddot{u}\ddot{a}$, Kakia $|wa:-ni$, N|u|en $|uun-te$). \diamond Kakia and N|u|en forms are obviously plurals. It may be tentatively assumed that the complex voiced aspirated uvular click in !Xóõ is primary, although there is variation between voiced and voiceless articulation even within !Xóõ itself. Nasalized coda vowel in !Xóõ is detachable as a class 2 suffix.
- Tuu: ***|^hu-**. \diamond All forms are clearly related. The correspondence N|uu ^{-h} : !Xóõ ^{-q^h} is recurrent (see DOG). Nossob forms are slightly closer to !Ui due to lack of nasalization, but since nasal codas are occasionally attested in !Ui as well (cf. ||Xegwi-LH $|^h\ddot{u}$), this cannot be a classificatory argument.

37. HAND [!Ui + Nossob + Taa]

- !Ui: ***|x'a** (|Xam $|x'a$, ||Ng!ke $|x'a$, †Kho-M $|x'a$, N|uu $|x'a:$, ||Ku|e $|x'a:$, ||Kxau $|x'a$, ||Xegwi-Z pl. $|x'a-\eta$). \diamond Curious replacement (no external parallels) in ||Xegwi: ||Xegwi-Z *kyi* (with assumed suppletivism between singular and plural forms) = ||Xegwi-LH q^hi . Otherwise, a very stable item.
- Nossob: ***|x'a/N/** (|'Auni $|x'a/n/$, |Haasi $n=|xan$). \diamond |Haasi $n=$ is probably the 1st person possessive prefix. Nasality in the coda is either the same as the ||Xegwi plural $-\eta$ or the nasal class suffix in Taa languages.
- Taa: ***|x'a** (!Xóõ $|x'\ddot{a}$, Kakia $|x'a$, N|u|en $|x'a$).
- Tuu: ***|x'a**. \diamond One of the most stable and securely reconstructed items on the list.

38. HEAD [!Ui + Taa] [? + Nossob]

- !Ui: ***ǃa** (|Xam $\tilde{a} \sim \tilde{a}:$ $\sim \tilde{a}:$, ||Ng!ke $\tilde{a} \sim \tilde{a}:$, †Kho-M \tilde{a} , N|uu $\tilde{a}:$, ||Ku|e $\tilde{a}:$, ||Kxau $\tilde{a}:$, ||Xegwi $\tilde{a}:$). \diamond Stable and securely reconstructed. Plural form may have been * $\tilde{a}-\eta$ (as in ||Xegwi, etc.).
- Nossob: (a) ***ǃa** (|'Auni $\tilde{a}:$); (b) ***xu** (|'Auni $x:uu$, |Haasi $\eta=x\ddot{a}$). \diamond Difficult situation. On one hand, Common Nossob **xu* is clearly the same as Proto-!Ui **xu* 'face' (|Xam *xu*, etc.) and reflects a very likely semantic shift 'face' \rightarrow 'head' (especially in light of external evidence from Taa which clearly shows * \tilde{a} as the original equivalent for 'head'). Since Bleek records both the older form $\tilde{a}:$ and the innovative form $x:uu$ in the meaning 'head' for |'Auni, it might be assumed that the Proto-Nossob form was still * \tilde{a} . On the other hand, it is also possible that |'Auni had simply reinstated the original word (at least in some contexts) under the very common influence of N|uu (in fact, this scenario is explicitly advocated by D. Bleek herself, see Bleek 1937: 211). Available evidence does not allow to make a definitive decision, meaning that both items have to be counted as technical synonyms on the Proto-Nossob level.
- Taa: ***ǃa-** (!Xóõ $\tilde{a}n$, Kakia $\tilde{a} \sim \tilde{a}\eta$, N|u|en $\tilde{a}\eta$). \diamond Status of the nasal coda is unclear, but probably suffixal in light of external data.
- Tuu: ***ǃa**. \diamond Another highly stable Tuu root, albeit largely replaced by 'face' in the Nossob group (see discussion on the |'Auni situation above).

39. HEAR [!Ui + Nossob] [- Taa]

- !Ui: ***tu** (|Xam $t:u \sim t:úi$, ||Ng!ke $tu \sim tu:i$, †Kho-M $\phi^hu \sim \phi^hu:-wa$, N|uu $\phi u:$, ||Kxau $tú$, ||Xegwi-Z tu). \diamond This stem is attested in many different morphological variants (cf. ||Xegwi-Z past stem $tu-wa$, present stem $tu-bi$, etc.; ||Ng!ke tu : 'hear' vs. $tu-\ddot{a}$: 'heard' vs. $tu-i$ 'listen' in Bleek 2000: 24), but root vocalism is almost always *u* regardless of the morphological environment.

- Nossob: |ʼAuni *tu:* ~ *tu:i*. ◇ Not attested in |Haasi. Curiously, the earliest source on |ʼAuni (Bleek 1929: 46) lists the forms *ta:ã ~ ta:a* in the meaning ‘hear’, but in Bleek 1937: 203–206, probably the same word is transcribed as *tiãñ ~ kiãñ* (reflecting palatalized articulation of *t-*) with the meaning ‘feel’. These look like two different etyma rather than morphemic variants of each other.
- Taa: ***tã**^f (!Xóõ *tã*^f, Kákia *tãa ~ ta:*^fη, N|u|len *tãñ*). ◇ Variable form of the !Xóõ stem is *ta*^f-, but nasalization is so pervasive in all doculects that we should probably project it onto the proto-level (or even reconstruct **ta*^fη with an actual velar nasal coda).
- Tuu: (?) ***tu**. ◇ In the !Ui branch (and possibly also in |ʼAuni, though the data here are limited and may also reflect !Ui influence), there is a rather clear distinction between the verbs **tu* ‘to hear’ and **ta ~ tã* ‘to feel’, cf. |Xam *tã ~ ta: ~ tã:* ‘to feel, try, seem, beware of’ with multiple text examples in Bleek 1956: 184. In Taa (or at least !Xóõ), both meanings seem to have been merged in the same root **tã*^f. Whether or not **tu* and **tã*^f are etymologically connected cannot be determined at this point, but since there is no strong evidence for grammatical Ablaut of any kind in !Ui or Taa, we should certainly treat them as two different roots, and postulate a probable lexical replacement in Taa. It should be noted that |ʼAuni is closer in this respect to !Ui than to Taa.

40. HEART [!Ui + Nossob + Taa]

- !Ui: *|**ʼai** (|Xam *ʼi:*, ||Ng|ke *ʼai ~ ʼe*, †Kho-M *ʼe:-ç^hi*, N|uu *ʼe:*, ||Ku|e *ʼẽ:*, ||Kxau *ʼae ~ ʼai-si*). ◇ A stable item, lost only in ||Xegwi where it is replaced by a Bantuism (||Xegwi-Z *kele*, ||Xegwi-LH *keleñ*). However, click efflux correspondences are unique, with a lot of variation between simple velar and glottalized articulation which cannot be fully ascribed to mistranscriptions; we are either dealing with an original root structure like *|*aʼi* (with metathesis of glottalization) or with the unique reflexes of a rare click type (see below).
- Nossob: *|**ʼe** (|ʼAuni *ʼe:* ~ *ʼe:*, |Haasi *n=|a-ʼe*). ◇ Structure of the form in |Haasi, except for the usual 1st p. possessive prefix *n=*, is unclear (reduplication?). Note the same discrepancy in click efflux articulation (glottalic in |ʼAuni vs. velar in |Haasi) as in !Ui.
- Taa: *|**qʼa-** (!Xóõ *qʼãñ*, pl. *qʼã:*, Kákia *ʼi:*, N|u|len *qʼañ*). ◇ Kákia *ʼi:* is unusual here because of the vocalism, but in light of external cognates in !Ui and Nossob it may actually be seen as more archaic in that respect. Perhaps what we see are the results of morphological variation in Proto-Taa, e.g. *|*qʼa-i* (→ Kákia *ʼi*) vs. *|*qʼa-n* (!Xóõ *qʼãñ*, N|u|len *qʼañ*).
- Tuu: *|**qʼa(-i)**. ◇ All the forms seem related, but reconstruction of click efflux and vocalism runs into problems. There may be a correlation between the glottalized / non-glottalized effluxes seen in !Ui and Nossob, on one hand, and the glottalic uvular efflux *-qʼ-* in Taa, in which case the latter should be set up for the protoform; clear additional evidence for this correlation is, however, lacking at present. As for the root vowel, variation in Taa suggests *a*, but this is also inconclusive, given the prevalence of front vocalism elsewhere. Only the !Xóõ-N|u|len node suggests *|*qʼa-n* as a full stem, so this may be an innovation; on the other hand, grammatical variants *|*qʼa-i* ~ *|*qʼa-n* may also reflect some meaningful opposition in Proto-!Ui, with only the former variant fused and preserved in !Ui, Nossob, and Kákia.

41. HORN [!Ui + Nossob + Taa]

- !Ui: *||**ãi** (|Xam *ʼhẽ:* ~ *ʼhẽi:*, ||Ng|ke *ʼãi*, †Kho-M *ʼẽi*, N|uu *qʼh^oe-si*, ||Xegwi-LH *ʼi:*). ◇ In most old sources, the word is hopelessly confused with TOOTH q.v., but N|uu and ||Xegwi

data clearly show that these are two different (albeit phonetically similar) etyma. However, there is additionally a serious incongruence between N|uu $\|q^hoe-$ and the rest of !Ui forms which rather go back to something like $\|āĩ$; the worst problem is the vocalism, since the other languages (as well as external cognates in Nossob and Taa) show no signs of labial vowels. It is possible that modern N|uu $\|q^hoe-$ is not related (although in that case, its provenance is a mystery); in any case, the reconstruction is primarily based on the correlation between |Xam and ||Xegwi as the most distant members of the !Ui branch.

- Nossob: |'Auni $\|ēĩ$. \diamond Not attested in |Haasi.
- Taa: $\|ā-$ (!Xóõ $\|āē$, Kakia $\|ān-ša$, N|u|en $\|ā$). \diamond Kakia $\|ān-ša$ = !Xóõ pl. $\|ān-sā$.
- Tuu: $\|ā-$. \diamond Nasalization of the vowel tentatively projected onto the proto-level due to its presence almost everywhere. Different codas most likely represent old morphology (e.g. sg. $\|ā-i$ vs. pl. $\|ā-n$, as in Taa, with generalization of the sg. form in !Ui and Nossob?).

42. I [!Ui + Nossob + Taa]

- !Ui: $\|ŋ$ (|Xam η , ||Ng!ke $\eta \sim n$, †Kho-M $\eta \sim na \sim n \sim na$, N|uu η , ||Ku|e η , ||Kxau $\eta \sim n$, ||Xegwi-Z $?η \sim ?n \sim ?in \sim ?iη \sim ?m \sim ?im \sim ?ĩ$). \diamond The primary and most common form of the root is that of a syllabic velar nasal; everything else is the result of contextual assimilations or combinations with various emphatic particles.
- Nossob: $\|ŋ$ (|'Auni $\eta \sim n \sim na \sim m$, |Haasi η). \diamond See notes on !Ui.
- Taa: $\|ŋ$ (!Xóõ n , Kakia $n \sim na \sim η \sim ηa \sim nya$, N|u|en $\eta \sim n \sim na$).
- Tuu: $\|ŋ$. \diamond The complete original paradigm (including stressed forms, clitical forms, emphatic forms, assimilated variants, etc.) is hard to reconstruct, but the monoconsonantal core of the Proto-Tuu 1st p. pronoun was undoubtedly a velar nasal, as still preserved in modern N|uu.

43. KILL [-]

- !Ui: $\|ʰa \sim \|ʰi$ (|Xam $|á(:) \sim |ʰá(:) \sim |i$, ||Ng!ke $|a(:) \sim |i$; $\sim |ʰi$, †Kho-M $|x'a$, N|uu $|ʰa$). \diamond Reconstructible on the |Xam-N|uu level. Vowel gradation is similar to what is observed in several other cases, but difficult to explain based on extant data (for modern N|uu, only the *a*-grade form of the root is attested, suggesting analogical leveling in recent times). In ||Xegwi, the equivalent is $\lambda iη \sim \lambda eη$ 'hit, strike; kill' (Z) = $\lambda'eun$ 'to hit' (LH); etymology is unclear, but the attested polysemy suggests semantic innovation (common semantic shift 'hit' \rightarrow 'kill').
- Nossob: |Haasi $!au$. \diamond Clearly the same word as |'Auni $!au$ 'to beat; to knock down', but not 'to kill', although the word 'kill' is not attested in |'Auni at all, so it cannot be excluded that the meaning 'kill' was expressed by the same form ($\|!au$) in Proto-Nossob.
- Taa: (?) $\|qa-$ (!Xóõ $qâi$, var. form $qa-JV$, Kakia $\|a$, N|u|en $\|wan$). \diamond It is unclear if the lateral click in Kakia and N|u|en is a real reflex of Proto-Taa $\|q-$ or if it represents a failed attempt to transcribe uvular articulation, but similar examples exist (e.g. !Xóõ $qâĩ$ 'pretty' = N|u|en $\|xai$ id.; !Xóõ $qâla$ 'to dig' = Kakia $|kâlaa$ id., where $|$ is clearly a typo for $\|$), so we tentatively list all these forms as cognate. Note that !Xóõ $qâi$ also conveys the meaning of 'forceful downward movement' (e.g. $qâi$ $\|úm$ 'stamp', etc.), which may be a case of homonymy but may also reflect an original meaning similar to 'hit', as in ||Xegwi or Nossob languages.

- Tuu: Not reconstructible. \diamond |Xam-N|uu $*|^{h}a \sim *|^{h}i$, with its seemingly archaic vowel gradation, looks like a good candidate, but has no recognizable cognates outside of that cluster. On the whole, the etymon ‘kill’ looks unstable and easily replaceable by words with the semantics of ‘hit’.

44. KNEE [-]

- !Ui: $*|^{h}u-$ (|Xam $\tilde{|\delta}a\eta \sim \tilde{|\acute{u}a\eta$, ||Ng!ke $\tilde{|\delta}i: \sim \tilde{|\delta}i:$, N|uu $\tilde{|\acute{u}}i:-si$, ||Xegwi-Z, ||Xegwi-LH $\tilde{|\delta}o-ma$). \diamond Same root in all three languages, but with different suffixation ($*\tilde{|\acute{u}}-/a/\eta$ in !Ui-N|uu, $*\tilde{|\acute{u}}-ma$ in ||Xegwi; we can still see the pure root form in |Xam pl. $\tilde{|\acute{u}}-\tilde{|\acute{u}a-d:e}$, with reduplication).
- Nossob: |’Auni $\tilde{|\omega\grave{e}}-\tilde{|\omega\grave{e}}$ (Bleek 1937); sg. $\tilde{|\omega\grave{e}}$, pl. $\tilde{|\omega\grave{e}}:-\tilde{|\omega\grave{e}}$ (Bleek 1956). \diamond Not attested in |Haasi.
- Taa: $*||x\tilde{u}$ (!Xóõ $\tilde{|\acute{x}\acute{u}}:-\tilde{|\grave{a}n}$, Kafia $\tilde{|\delta}-\tilde{|\grave{a}\eta}$, N|u|en $\tilde{|\acute{u}}-\tilde{|\acute{i}}$). \diamond All Taa doculets yield a compound in the meaning ‘knee’, where the second part is clearly HEAD q.v. (so it may be assumed that the original meaning of the compound was rather ‘knee-cap’). The root is tentatively reconstructed with a voiced velar fricative efflux based on !Xóõ data (Kafia and N|u|en transcriptions are unreliable).
- Tuu: Not reconstructible. \diamond Superficially, |’Auni $\tilde{|\omega\grave{e}}$ resembles Taa $*||x\tilde{u}$, and a common origin is not excluded (assuming inadequate click transcription and different suffixation). But it is even more similar to Proto-Kalahari Khoe $*||\omega\grave{e}$ ‘knee’ (Vossen 1997: 457), meaning that borrowed origin is more likely here than inherited. If the Nossob form is excluded, !Ui and Taa forms may reflect the original Proto-Tuu ‘knee’ with comparable probability.

45. KNOW [-]

- !Ui: (?) $*||x\tilde{a}e$ (||Ng!ke $\tilde{|\acute{a}i}$, N|uu $\tilde{|\acute{x}a}e$). \diamond This seems to be the main, if not only, equivalent of the required meaning in the N|uu cluster, cognate with |Xam $\tilde{|\acute{a}i} \sim \tilde{|\acute{a}i}^f$ ‘to take notice, be(come) aware of smth.’ (Bleek 1956: 550). In |Xam itself, the meaning ‘know’ is usually correlated with the root $\tilde{|\acute{f}en(n)} \sim \tilde{|\acute{f}e}$ ‘to know, to think’, further related to ||Ng!ke $\tilde{|\acute{f}e}$, N|uu $\tilde{|\acute{f}i}$: ‘to think’, still further to !Xóõ $\tilde{|\acute{f}an}$ ‘to think’ and ultimately to Khoe $*\tilde{|\acute{f}an}$ ‘to think (\rightarrow to know)’, from which this root may have been diffused into different branches of Tuu. ||Xegwi-LH $\tilde{|\acute{c}i}$, ||Xegwi-Z $\tilde{|\acute{c}i-ya}$ remains without any etymology. In light of all this, the N|uu root with its semantically similar |Xam cognate remains the optimal, if still weak, candidate for Proto-!Ui ‘to know’.
- Nossob: (a) |’Auni $\tilde{|\acute{x}a}i \sim \tilde{|\acute{x}e-ki}$; (b) |Haasi $\tilde{|\acute{u}ma}$. \diamond The |’Auni form is clearly connected with N|uu $\tilde{|\acute{x}a}e$, but it may be a borrowing rather than a genetic cognate. The |Haasi form is just as clearly related to !Xóõ (below), yet this could also be interpreted as an areal isogloss. Unclear.
- Taa: !Xóõ $\tilde{|\acute{u}m\grave{a}}$ (var. form $\tilde{|\acute{u}BV}$). \diamond Kafia $\tilde{|\acute{a}}$ ‘to know’, published only in the early source Bleek 1929: 51, is somewhat dubious.
- Tuu: Not reconstructible. \diamond Although there is a clear isogloss between |’Auni and !Ui, on one hand, and between |Haasi and !Xóõ, on the other, both may have areal rather than genetic interpretations, and given the overall unstable nature of this concept in Tuu (cf. the Khoe root $\tilde{|\acute{f}an}$ with its wide diffusion over Tuu territory), we should probably exclude this word from comparison for safety reasons.

46. LEAF [-]

- !Ui: Not reconstructible. \diamond ‘Leaf’ is not a basic concept for !Ui speakers, and the word is usually borrowed (N|uu *bla:r-si* ← Afrikaans, ||Xegwi-LH *li=k^hasi-zi* ← Swazi), not attested, semantically questionable (e.g. |Xam *juhm*: ‘leaf, stick’), or completely isolated (||Ng!ke *xerro*: ‘leaves, foliage’ without any external cognates).
- Nossob: Not attested.
- Taa: (?) *|**ana** (!Xóõ *āna*, Kafia *ana*). \diamond Formally reconstructible for Proto-Taa; however, the root is completely identical with Khoe *|*ana* ‘leaf, grass’ (Vossen 1997: 424), which strongly suggests a borrowed origin (which agrees well with the general instability of this concept in Tuu). Cf. also N|u|len *ābu* ‘leaf’ (Bleek 1929: 52; incorrectly listed as SIV = |’Auni in Bleek 1956), a different root without any external etymology.
- Tuu: Not reconstructible. All correlated items may be of secondary origin.

47. LIE [!Ui + [Nossob + Taa]]

- !Ui: ***ta** (|Xam *t:a* ~ *ta* ~ *t:ē* ~ *t:n* ~ *t:η*, ||Ng!ke *tia* ~ *kia*, †Kho-D *qâ*, N|uu *qa*, ||Kxau *ta* ~ *da* ~ *tn*). \diamond Despite some phonetic variation in the coda (most likely reflecting various morphologic variants), the most common and probably original shape of the root should be reconstructed with final *-a. The situation in ||Xegwi is unclear, with Bleek and LH contradicting themselves and neither of the variants (Bleek *la*: ‘lie’, LH *iη=θijne* ‘I lie down’) having external connections.
- Nossob: |’Auni *tàa*. \diamond Cf. also *tūa* ‘to lie curled up’ and transitive *tōā-a* ‘to lay down, to bury’. Not attested in |Haasi.
- Taa: ***tu** (!Xóõ *tū*., Kafia *tu(:)* ~ *tá*: ~ *tā*., N|u|len *tu*.). \diamond Note the variation in Kafia, unclear on its own but instructive in light of external comparison.
- Tuu: ***ta** ~ ***tu**. \diamond Although the dominant variant of this root in !Ui is clearly **ta* and in Taa clearly **tu*, scant evidence for the opposite also exists: Bleek 1929: 53 records the variant *tu* for ||Ng!ke, and Bleek 1956 has *tá*: for Kafia. The situation is slightly reminiscent of HEAR (except that the distribution of vocalizations is reversed), but in that case there was additional evidence to argue in favor of two original roots (‘hear’ vs. ‘feel’). Here, it is rather advisable to treat both forms as morphological variants of a single original root. The Nossob form is morphologically closer to Tuu, but could actually represent an “intermediate” variant (*to-a* ← **tu-a*, while !Ui languages go further and contract **tu-a* → **tá*?)

48. LIVER [!Ui + Taa]

- !Ui: *|**aN** (|Xam *āwan* ~ *āán*, ||Ng!ke *āi:n*, N|uu *āan* ~ *āan* ~ *āan*, ||Kxau *ηanηa*). \diamond Not attested in ||Xegwi. Variants fluctuate between *|*an/a/* and *|*a-ni*.
- Nossob: Not attested.
- Taa: *|**am** (!Xóõ *ām*, N|u|len *ām*). \diamond Not attested in Kafia.
- Tuu: *|**aN**. \diamond It can hardly be doubted that all listed forms belong together, but reconstruction of the coda is somewhat problematic, given all the variety between !Ui and Taa. It is possible that the original root was simply *|*a-*, particularly in light of !Xóõ plural forms (possessive *ā*., alienated *ā*) which seem to drop the labial nasal as a suffix. On the other hand, nasality is such a persistent feature for all reflexes that it is hard to believe it was not, in some way or other, an intrinsic part of the root. The provisional reconstruction *|*aN* reflects that uncertainty.

49. LONG [!Ui + Nossob?]

- !Ui: *|a (||Ng!ke /a:, N|uu /ã:, ||Kxau /a:, ||Xegwi-Z /ã, ||Xegwi-LH /ã:). ◇ The isogloss between N|uu, ||Kxau, and ||Xegwi clearly identifies */a as the optimal candidate for Proto-!Ui ‘long’, although there are some phonetic problems — in N|uu, the coda has a nasalized vowel, and in ||Xegwi we see unpredictable variation between /ã (no glottalisation) and /ã̃ (nasal efflux) depending on the dolect. This may be due to a more complex protoform (e.g. something like */ã̃ã̃ with different types of contractions) or to several original morphological variants (e.g. */a vs. */a-/a/N); the issue requires a better understanding of !Ui adjectival morphology. Curiously, the root is not at all attested in |Xam, where the semantic definition ‘tall / long / high’ is instead attributed to the form !xó:-wa, pl. !xó-!xó-ka — transparently derived from the verb !xo: ‘to grow up, climb up; make upright, make tall’.
- Nossob: |ʼAuni /ã̃-si. ◇ Glossed as ‘big, long, tall’, but ‘big’ is probably incorrect (the proper equivalent for this meaning in |ʼAuni is ús/i/, see BIG). Not attested in |Haasi.
- Taa: (a) !Xóõ /ám; (b) Kakia /úm. ◇ Not reconstructible (the two forms are clearly not related).
- Tuu: Not reconstructible. ◇ The |ʼAuni form is clearly the same as the !Ui form, although it is impossible to determine if it is inherited or historically borrowed from N|uu.

50. LOUSE [!Ui + Taa]

- !Ui: *|u- (|Xam /ũiŋ, ||Ng!ke /ũin-ya, N|uu /u-si, ||Xegwi-LH /ũe-zi). ◇ Root vowel reconstruction is provisional (labial variant is the most common, but it could be assimilated to the labial click).
- Nossob: Not attested.
- Taa: !Xóõ /ú. ◇ Plural: /ú-ú. Not attested in Kakia or N|u|en. Another synonym is !Xóõ /xóni ‘louse’; semantic differences between the two forms are unclear.
- Tuu: *|u-. ◇ The lexical match between !Ui and !Xóõ is transparent and allows to reliably project the root onto the Proto-Tuu level, even despite relatively scant attestation of the word in both branches (and a complete lack of attestation in Nossob). It should be noted that the alternate !Xóõ synonym /xóni bears an uncanny resemblance to the common Kalahari Khoe term for ‘louse’, *k’uni, which in term is somewhat irregularly connected with Khoekhoe *k’uri id. (Vossen 1997: 462); however, presence of a lateral click in !Xóõ is somewhat befuddling, since it does not allow to explain the word as a (quite common) relatively recent borrowing from Kalahari Khoe. Could this be another piece of evidence from a “pre-Tuu / pre-Khoe substrate”?

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

- Crawhall, Nigel. 2004. *!Ui-Taa language shift in Gordononia and Postmasburg Districts, South Africa*. PhD thesis: Faculty of Humanities, University of Cape Town.
- Doke, Clement M. 1936. An outline of ǀKhomani Bushman phonetics. *Bantu studies* 10: 433–461.
- Güldemann, Tom. 2005a. “Tuu”: A New Name for the Southern Khoisan Family. In: Tom Güldemann (ed.). *Studies in Tuu (Southern Khoisan)*. *University of Leipzig Papers on Africa, Languages and Literatures* 23: 2–9. Leipzig: Institut für Afrikanistik, University of Leipzig.
- Güldemann, Tom. 2005b. 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. 2017. Casting a Wider Net over Nǀng: The Older Archival Resources. *Anthropological Linguistics* 59(1): 71–104.
- Güldemann, Tom. 2018. Historical linguistics and genealogical language classification in Africa. In: Tom Güldemann (ed.). *African Languages and Linguistics*: 58–444. Berlin: DeGruyter Mouton.
- Hastings, Rachel. 2001. Evidence for the genetic unity of Southern Khoesan. *Cornell working papers in linguistics* 18: 225–246.
- Kassian, Alexei, George Starostin, Anna Dybo, Vasilij Chernov. 2010. The Swadesh wordlist. An attempt at semantic specification. *Journal of Language Relationship* 4: 46–89.
- Köhler, Oswin. 1981. Les langues khoisan. In: G. Manessy (ed.). *Les langues dans le monde ancien et moderne. I: Les langues de l’Afrique subsaharienne*: 455–615. Paris: Editions du CNRS.
- König, Christa, Bernd Heine. 2001. *The !Xun of Ekoka. A demographic and linguistic report*. *Khoisan Forum Working Paper No. 17*. Köln: University of Cologne.
- Lanham, Leonard, D. P. Hallows. 1956a. Linguistic relationships and contacts expressed in the vocabulary of Eastern Bushman. *African studies (Johannesburg)* 15(1): 45–48.
- Lanham, Leonard, D. P. Hallows. 1956b. An outline of the structure of eastern Bushman. *African studies (Johannesburg)* 15(3): 97–118.
- Maingard, Louis F. 1937. The ǀKhomani 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.
- 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 Nǀuu clicks. *Journal of the International Phonetic Association* 39(2): 129–161.
- 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. 2014. *Adoption, Maintenance and Loss of Click Contrasts*. Talk presented at: Sound Change in Interacting Human Systems, 3rd Biennial Workshop on Sound Change, University of California, Berkeley, May 28–31, 2014. Available online at: http://linguistics.berkeley.edu/SCIHS/abstracts/5_FridayAfternoon/Sands.pdf.
- Sands, Bonny, Amanda Miller, Johanna Brugman. 2007. The Lexicon in Language Attrition: The Case of Nǀuu. 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. 2008. From Modern Khoisan Languages to Proto-Khoisan: the Value of Intermediate Reconstructions. In: Ilya Smirnov (ed.). *Aspekty komparativistiki III [Aspects of comparative linguistics III]*. *Orientalia et Classica, Vol. XIX*: 337–470. Moscow: Russian State University for the Humanities.
- 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. 2016. From wordlists to proto-wordlists: reconstruction as “optimal selection”. *Faits de langues* 47/1: 177–200.
- 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 leksikostatisticheskoy klassifikacii. Tom I: Metodologija. Kojzanskije jazyki*. Moscow: Jazyki slav'anskoj kul'tury.
- Story, Robert. 1999. *K'u^ha:si manuscript (MS collections of the Ki^hazi dialect of Bushman, 1937)*. Ed. by Anthony Traill. *Khoisan Forum Working Paper No. 13*. Köln: University of Cologne.
- Traill, Anthony. 1994. *A !Xóǀ dictionary. Quellen zur Khoisan-Forschung/Research in Khoisan studies, Bd 9*. Köln: Rüdiger Köppe Verlag.
- Traill, Anthony. 1995. Interpreting |Xam phonology: the need for typological cleansing. In: Anthony Traill, Rainer Vossen, Megan Biesele (eds.). *The complete linguist: papers in memory of Patrick J. Dickens (Afrikanische Sprachen und Kulturen)*: 509–523. Köln: Rüdiger Köppe Verlag.
- Traill, Anthony. 2018. *A Trilingual !Xóǀ dictionary: !Xóǀ – English – Setswana. Quellen zur Khoisan-Forschung/Research in Khoisan studies, Bd 37*. Köln: Rüdiger Köppe Verlag.
- Vossen, Rainer. 1997. *Die Khoe-Sprachen: Ein Beitrag zur Erforschung der Sprachgeschichte Afrikas*. Köln: Rüdiger Köppe Verlag.
- Westphal, Ernst O. J. 1962. On classifying Bushman and Hottentot languages. *African Language Studies* 3: 30–48.
- Westphal, Ernst O. J. 1971. The Click Languages of Southern and Eastern Africa. In: Jack Berry, Thomas A. Sebeok (eds.). *Linguistics in Sub-Saharan Africa*: 367–420. The Hague / Paris: Mouton.
- 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. Лексикостатистические исследования по койсанским языкам II/1: к вопросу построения списка Сводеша для пра-ту языка

В статье, представляющей собой первую из двух частей исследования, представлены результаты общего лексикостатистического обзора языковой семьи ту (= южнокойсанской семьи), в ходе которого частично реконструируется список Сводеша для языка пра-ту и разъясняется ряд сложных моментов, касающихся внутренней классификации языков ту. В настоящей публикации представлен краткий обзор источников, перечислены основные методологические проблемы, связанные с диахроническим изучением языков ту, и приведены комментарии относительно исторической фонологии этих языков. Большую часть статьи занимает Приложение, в котором дается попытка реконструкции первых 50 элементов из списка Сводеша для трех промежуточных узлов семьи ту (пра-!ви, пра-носсоб и пра-та).

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