



**Universiteit
Leiden**
The Netherlands

Tocharian and Samoyed: on the question of Uralic substrate influence in Tocharian

Warries, A.R.

Citation

Warries, A. R. (2025, June 18). *Tocharian and Samoyed: on the question of Uralic substrate influence in Tocharian*. Retrieved from <https://hdl.handle.net/1887/4250485>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/4250485>

Note: To cite this publication please use the final published version (if applicable).

9 Potential loanwords

The most obvious sign of contact between languages often comes in the form of lexical borrowings. While most of the potential traces of contact between Tocharian and Uralic are structural, there are also a few possible loanwords, which will be discussed in this chapter.¹⁰⁸

9.1 Introduction

Janhunen was the first to make lexical comparisons between Tocharian on the one hand and Samoyed on the other (Janhunen 1983). The limits of this line of research were immediately acknowledged by him, as the inherent difficulties are manifold. For instance, there may have been early Indo-European languages spoken in South Siberia that were closely related to the ancestor of Tocharian, or there may have been branches that disappeared without leaving any descendants. To make matters more difficult, the Samoyed languages themselves have undergone further changes since the relevant period in time, so that “there is no reason to assume that very many traces of Indo-European influence would be left in the modern Samoyedic idioms” (Janhunen 1983: 118). The time depth is of course significantly large.

Nevertheless, Janhunen discussed two potential loanwords from specifically pre-Proto-Tocharian into pre-Proto-Samoyed, namely pre-PS **säjt³wə* ‘seven’ and **wäsa* ‘metal, iron’ (Janhunen 1983: 5–7). A few more suggestions have been made over the years, with especially Napol’skikh (2001) and Kallio (2004) dedicating articles to a discussion of lexical comparisons between Tocharian and Samoyed specifically or Uralic generally. Napol’skikh makes use of the concept of “para-Tocharian”, a hypothetical sister language of attested Tocharian, to explain certain unexpected sound correspondences present in his corpus. Kallio, meanwhile, criticizes the earlier loan etymologies, but does provide the new comparison in the form of PS **wɛn* ‘dog’ and pre-PT **kwɛnə* ‘dog (obl.sg.)’. The pool of suggestions has not greatly increased in the last 40 years, and in his treatment of the hypothesis that early Tocharian was in contact with early Samoyed, Peyrot lists only four that he deems “relatively good”, namely PS **säjt³wə* ‘7’, PS **wɛn* ‘dog’, PS ± **mänäwjə* ‘full moon’, and PS **wäsa* ‘metal, iron’ (Peyrot 2019a: 100–101).

I will give an evaluation of the etymologies that have been adduced in the publications mentioned, and added a few additional suggestions on the basis of not only the reconstructed Proto-Samoyed lexicon (Janhunen 1977), but also the individual

¹⁰⁸ I presented this topic at the Finno-Ugric Congress of 2022 in Vienna. This chapter is an adaptation of the resulting article (Warries fthc.), which will appear in the proceedings. I thank Tapani Salminen and an anonymous reviewer for their comments on the article version.

northern Samoyed languages. I searched through several dictionaries to consider per word if a Proto-Samoyed reconstruction matched a Tocharian word. I used Tereščenko (1965) for Tundra Nenets, Sorokina & Bolina (2001) for Forest Enets, and Kosterkina, Momde & Ždanova (2004) for Nganasan; for Tocharian I mainly consulted Adams' (2013) dictionary of Tocharian B, which means that I had a bias in favour of that language, possibly to the detriment of lexical material only found in Tocharian A. Despite my best efforts, I may still have missed some possible connections, and a study of more extensive (future) publications on the lexicon of especially Enets and Nganasan may yet yield a few additional results.

The methodology I used to identify possible comparisons was quite strict. Since there is always the possibility of accidental resemblances in both form and meaning, a relatively high degree of similarity is to be desired at the current stage of research. Perhaps if it is established with greater certainty that the two languages were in contact, and a larger number of comparisons is already established, one could venture to see different patterns, but at the present moment I looked for a number of specific criteria. Of a root with the shape CVC(CV), the reconstructed consonants should match, or there should be a reasonable explanation as to why they do not fully match (e.g., if one side of the comparison is known to have reduced or disallowed a specific type of cluster that is found on the other side). The vowels should at least be similar, e.g., be similarly front or back, so far as we can understand their earlier values. Precise phonetic information unavoidably remains irretrievable, so that at best we can arrive at an approximation based on the pronunciation of the relevant phonemes in the attested languages and their more distant cousins. Full semantic identity is of course desirable, but a plausible shift can be argued for and may be permitted. Additionally, a verb in the donor language should ideally correspond to a verb in the recipient language, and a noun or adjective to the like as well.

9.2 Earlier comparisons

9.2.1 pre-PT **septə(m)* 'seven' ⇒ pre-PS **säjt³wə* 'id.'

The first word adduced as a loanword from Tocharian into Samoyed by Janhunen (1983) is the word for 'seven': PS **säjt³wə* in connection to PT **s³əptə* from PIE **septm̥*. This is later repeated by Janhunen as the most notable representative of Samoyed-Tocharian contact (Janhunen 1998: 477). The etymology requires a number of complicated steps to work, however, such as metathesis of **-pt-* to **-tp-* and a weakening of **-p-* to **-w-*. The vowel correspondence of simple **e* with a sequence **äj* (written by Janhunen still as **ej*, as per the old reconstruction of Proto-Samoyed; cf. Helimski 2005, Aikio 2006a) is also rather strange.

It appears furthermore that PS **säjt³wə* ‘seven’ can be derived from the same PU **ćjć(ć)imä* as Fi. *seitsemän* ‘seven’, according to Aikio (2020: 110–111). The original value of PS **t³* in **säjt³wə* cannot be determined from Samoyed, but it may indeed represent PS **s* from PU **ć*; the other possibilities are PS **c* (from PU **ć*), PS **t* (PU **s* or *t*) and PS **k* (PU **k*). The **-m-* was apparently irregularly lenited to **-w-* in the Samoyed reflex (Aikio 2020: 110–111). Despite this irregular development, I would agree with Aikio that the Uralic etymology is to be preferred over a loan etymology that shows several unexplained modifications of the source word.

9.2.2 pre-PT **wesa* ‘gold’ and PS **wäsa* ‘metal, iron’

Janhunen also discusses the similarity of the Tocharian words for ‘gold’ (TA *wäs*, TB *yasa* from PT **w^əasa*) to Samoyed words for ‘metal, iron’ (Ng. *basa*, NeT *yesya*, PSk. **kuäsə*, Km. *baza*, Mt. *be/äsä*). The story is quite complicated, however, as Janhunen admits (Janhunen 1983: 120–121; cf. also Kallio 2004: 132–133). The Samoyed words are connected with metal names in other Uralic languages, such as SaaN *veaiki* ‘copper’, Fi. *vaski* ‘id.’, Hu. *vas* ‘iron’, pointing to a reconstruction as approximately PU **wäckä*, although Hungarian rather indicates back vocalism; some other Uralic branches have similar words (Grünthal et al. 2022; supplemental material). The loss of **k* in the cluster PU **ćk* in Samoyed is regular (cf. 3.5.2), but the vowels are troublesome. Not only do the Saami, Finnic and Hungarian vowels not match nicely, also Ng. *basa* seems to imply a reconstruction PS **wäsa* with second-syllable **-a* of unexplained origin, as according to Aikio (2015: 42–43). Additionally, Aikio argues that the Selkup and Kamas forms obtained their first-syllable vocalism due to some effect caused by this same **-a* as seen in Nganasan.

However, according to Grünthal et al. the second-syllable *-a* in Nganasan may also be of analogical origin due to paradigmatic alternations between final *-i* (usually PS **-ä*) and *-a* (usually PS **-a*) in that language. Grünthal et al. also consider the vowel development of the initial sequence **wä-* in the Nganasan, Selkup and Kamas words to be regular without influence from a second-syllable following **-a* (Grünthal et al. 2022; supplemental material). This would make the Samoyed reconstruction PS **wäsä* as expected from PU **wäckä*. The complicated situation involving this word has inspired its analysis as a Wanderwort that entered different Uralic languages independently at an early stage of differentiation, and/or was borrowed from branch to branch (Aikio 2015: 43).

For the purposes of a comparison to Tocharian, an important aspect is that a **k* in the Finno-Ugric languages is not easily explained as a reflex of a pre-PT **wesa*, so that a loan etymology of PS **wäsa* from Tocharian would seem to necessitate a thorough separation of the Samoyed word from its Finno-Ugric counterparts. This seems undesirable. An additional complication is that PT **w^əasa* does not have a fully ironclad Indo-European etymology (on both points, cf. Thorsø & Wigman et al. 2023). A borrowing in the opposite

direction, from early Samoyed into Tocharian can also be considered, although it would be so far the only example of its kind. If this route is taken, it would really be early Samoyed, rather than another Uralic branch, that was the source of the borrowing, on account of the lost **k* (again, cf. Thorsø & Wigman et al. 2023). An important issue with this option, however, aside from the lack of comparable loans in the same direction, is that the loss of **k* in the cluster PU **ćk* cannot be established to have occurred particularly early in pre-Proto-Samoyed (see 3.9). This may have been a late change, potentially taking place after any possible contact with early Tocharian.

Even after all this time, this etymology is still rife with uncertainty. The Tocharian word may be innovative, or could be derived from Proto-Indo-European, while the Samoyed word cannot easily be derived from Tocharian and a borrowing from Samoyed into Tocharian seems unlikely. This means that, while a connection looks very attractive at first, due caution is advised.

9.2.3 pre-PT **mēnē*- ‘moon’ ⇒ PS ±**mānā*/*iwjə* ‘full moon’

A Proto-Samoyed word for ‘full moon’ can be tentatively reconstructed based on NeT *myenyuy*^o and EnT *menio*, both of the same meaning. No cognates are known from the other Samoyed languages, but it bears a strong semantic and phonological resemblance to the Indo-European and Tocharian word for moon, cf. TB *meñe*, TA *mañ* (Blažek *apud* Napol’skikh 2001). This derives via a pre-PT **mēnē*- from PIE **meh₂nē(n)*- ‘moon’, with a probably early change of PIE **eh₂*, to pre-PT **ē*. The Samoyed formation looks like it requires two extra suffixes, however, seemingly first a **-w* and then the suffix **-jə*, which is used to form adjectives, but also secondary nouns. Especially the latter suffix is very common in the Samoyed languages, but the need for extra derivational material to arrive at the attested word certainly does not add to the plausibility of this loan etymology.

There is a parallel of another Uralic language borrowing the Indo-European word for ‘moon’, namely in Saami: SaaN *mánnu* < PSaa. **mānō*, from Old Norse (Sammallahti 1998: 253–254). It looks striking that the rounded vowel in the second syllable is matched in both cases, but in Saami it derives from an **-ō* in Germanic; such a source does not seem to be available for Samoyed, if it is based on pre-PT **mēnē* (the ancestor to TB *meñe*). The semantic and formal match of these Samoyed and Tocharian words is quite close, so that a borrowing etymology of the base for NeT *myenyuy*^o, EnT *menio* is overall possible, but not entirely compelling.

9.2.4 pre-PT **monko*- ‘lack, deficit’ ⇒ PS **māŋkâ* ‘poverty, need’

PS **māŋkâ* ‘poverty, need’ is another one of the words listed by Napol’skikh (2001), this time originally from Helimski. Representative Samoyed reflexes are Ng. *məŋkə*, NeT

maŋk^o, all of the same basic meaning (Ru. ‘нужда, бедность’).¹⁰⁹ The semantics are virtually identical to TA *maŋk* ‘lack, deficit’. Napoľskikh somewhat confusingly compares the noun PS **māŋkâ* ‘poverty, need’ to the verb root TAB *mānk-* ‘to lack’, with supposed adoption of “para-Tocharian” **ə* as Samoyed **ă* after **m-*. This sound substitution rule with reference to “para-Tocharian” does not lend to the loan etymology’s credibility, and it is probably the reason why this example is not listed by Peyrot as one of the “relatively good” examples of loan etymologies (Peyrot 2019a: 100–101).

However, Napoľskikh’s workaround is unnecessary. The Tocharian A noun *maŋk*, from PT **menke*, can derive regularly from an *o*-stem nominalization early pre-PT **monko-*, which provides a quite adequate source for the Proto-Samoyed noun **māŋkâ*; especially when considering that the common change from PU **o* to PS **ă*, if it applied after this borrowing, would allow for an even more faithful adoption of the pre-Proto-Tocharian vowel. The phonological and semantic match between the Tocharian and Samoyed words is thus very high, and I am not aware of any cognates for PS **māŋkâ* in other Uralic languages. Borrowing of this word from pre-Proto-Tocharian into pre-Proto-Samoyed is therefore probable.

9.2.5 Pre-PT **kwēnə(m)* acc. of ‘dog’ ⇒ pre-PS **wēn* ‘dog’

The Proto-Samoyed word **wēn* ‘dog’, with reflexes such as Ng. *bay*, EnF *bu?*, NeT *weh*, SkTym der. *kanak*, has no Uralic etymology. Kallio (2004) has proposed that it could be a borrowing from the Tocharian word for dog, not from the nominative TAB *ku*, but from the oblique TB *kweŋ*, TA *koŋ*. The oblique forms go back to PT **kwenə*, from pre-PT **kwēnə(m)*, PIE **k̑uon̑n̑*. The initial cluster **kw-* could not be borrowed as such in Samoyed, which lacks word-initial clusters altogether, so it may have been simplified to a simple **w-*. There are no known parallels for this, but the development is plausible. The Samoyed vowel **ē* may point to the de-rounded **ĕ* (\pm *[Λ]) of pre-Proto-Tocharian, from earlier IE **o* (Kallio 2004: 133–134). The semantic match is of course exact.

Interestingly the vowel in PS **wēn* does not correspond to the vocalism of the previous etymology PS **māŋkâ*, where PIE **o* would rather be reflected as PS **ă* (via a pre-PS **ă* or **o*). Perhaps the difference could be due to an earlier adoption of **māŋkâ* than of **wēn*, with in between the pre-Proto-Tocharian sound shift of **o* to **ĕ*, but at this point it is difficult to say.

¹⁰⁹ There is also Enets *maggô* ‘arm : poor’ in Mikola (1995: 141), and Sorokina & Bolina (2001: 66) give derived *magutuda* ‘бедный : poor’ and *magutut* ‘бедно : poorly’.

9.3 New comparisons

9.3.1 pre-PT **lonko-* ‘steep slope; cliff’ ⇒ PS **l̥ɲkâ* ‘steep bank; steep’

The underived Samoyed word is found in Nenets, NeT *l̥ɲk*^o, NeF *l̥ɲk*^o ‘обрыв, яр; крутой : precipice, cliff, steep bank; steep’, and derivations are found in EnT *logori*, EnF *logar* ‘гора : mountain’ and Ng. *l̥ɲkəri* ‘порор : rapids’. The Tundra Nenets acc.pl. is *l̥ɲko* (Salminen 1998). This means that it is of the type where nom.sg. -^o changes to acc.pl. -*o*, i.e., PS **-â* to **-âj*, which points to the word being back vocalic PS **l̥ɲkâ*. Front vocalic **l̥ɲkâ* should have yielded an acc.pl. **l̥ɲcye* instead in Tundra Nenets.

This fits with back vocalism in pre-PT **lonko-*, the source for TB *leñke* ‘slope’, which is very similar in both meaning and form to the Nenets words. It is an *o*-stem nominalization from the verb TAB *l̥ɲk-* ‘hang, dangle’, from PIE **h₁leng^{wh}-* ‘move lightly’ (Peyrot 2013: 476), thus quasi-PIE **h₁long^{wh}-*. The initial **l-* of PS **l̥ɲkâ* before a vowel other than PS **e* or **i* points to this word being borrowed into Samoyed after Proto-Uralic initial **l-* had changed into **j-*, as in, e.g., PU **lupsa* ‘dew > PS **jâpta* ‘id.’ (Janhunen 1981: 223). Since the semantic and formal match with Tocharian is so close, this could provide a suitable source of borrowing. Especially the consonant skeleton is the same, only the vowels do not match up exactly. If we project PS **l̥ɲkâ* back to pre-Proto-Samoyed, we get **luɲkâ*. The vowels **u* and **o* are close neighbours, so that the discrepancy may be allowed, but it is notable that the vowel reflex is not identical to that seen in PS **m̥ɲkâ* from a pre-PT **moɲko-* (cf. above).

Anikin and Helimski connect the Nganasan (*l̥ɲkəri*) and Tundra Enets words (*l̥ɲkəri*, *logori*) with an Evenki word *лэңирү* ‘колода : (chopping) block (of wood)’, and consider it to be a borrowing from Tungusic into northern Samoyed (Anikin & Helimski 2007: 125).¹¹⁰ Evk. *лэңирү*, in turn, they derive from *лэңэ-/лаңа-* ‘сломать (зуб) : break (a tooth)’. I think that the semantic connection for this loan etymology is more distant, and it separates the Nganasan and Tundra Enets words from the semantically more closely related Nenets word. The Evenki/Tungusic loan etymology can therefore be rejected, with a possible Tocharian etymology to replace it.

9.3.2 pre-PT **lenk-* ‘dangle, sway, hang’ ⇒ pre-PS **laɲkâ-* ‘hang, dangle’

Tocharian has a verb *l̥ɲk-* ‘dangle, sway, hang’, which is very similar to a word of the same meaning found in Samoyed. In Tundra Nenets, there are only derivatives attested: NeT *lyɲkabta-* ‘повесить, развесить; понести что-л. свешивающееся : hang (tr.); carry sth. that is hanging down’, *lyɲkada-* ‘развесить что-л. : hang up sth.’, *lyɲkasa-* ‘висеть (в большом количестве); бродить без дела : hang (in great number); wander around’,

¹¹⁰ I am grateful to Kaisla Kaheinen for pointing this out to me.

etc. The underived root would be *lyajka-*¹¹¹, which can be mechanically reconstructed as PS **lajkā-*.¹¹¹ Outside of Tundra Nenets, a direct cognate to *lyajkabta-* is found in NeF *lyajkaptā-* with the same meaning, and a borrowing from Tundra Nenets is found in EnF *lagata-* ‘id.’.

Due to its absence in the other Samoyed languages, the distribution of this word leaves something to be desired. Nevertheless, the correspondences in form and meaning between Tundra Nenets and Tocharian are quite close, so that we may suppose that the word was lost elsewhere in Samoyed (or remains unattested), similar to the word for ‘full moon’, and the underived word for ‘steep slope’. Like with the previous word, the initial **l-* in Samoyed points to this word being borrowed at some point in the specific prehistory of the Samoyed branch, as it cannot be inherited from Proto-Uralic; pre-Proto-Tocharian seems to be a possible source.

9.3.3 pre-PT **der-* ‘split, separate’ ⇒ pre-PS **tar-* ‘divide, separate’

Reflexes of PS *tar-* ‘divide, separate’ is attested in all branches of Samoyed, with pretty much the same meaning everywhere: Ng. *tar-*, NeT *tyar-*, EnF *tār-*, Sk. *tār-*, Km. *tarə-*, Mt. *tār-*. In SW (154-155), the reconstruction is given as **tār- ~ *tārâ-*, which in the new vowel system notation is **tar(â)-*. The form with second-syllable **-â* is for the Kamas reflex, and may well be secondary. The word also has a suggested Uralic etymology in UEW: **šur3-* ‘schneiden, teilen’ (UEW: 503–504). However, PU **u* does not become PS **a*. An alternative etymology can thus be proposed, namely a borrowing from the ancestor of PT **tšər-* ‘be separated; caus. separate’ from PIE **der-* ‘split, flay, tear’. Semantically, the Tocharian and Samoyed verbs are very similar. In Tocharian texts the verb is often used to refer to people being separated from one another, but given the original more physical meaning of ‘split, tear’ attested in other Indo-European languages, Samoyed ‘divide, separate’ may represent a very fitting intermediary stage also for the Tocharian semantic development. The Samoyed meaning is more distant from, for instance, Gr. δέρω ‘flay’, Skt. *ḍṛṇāti* (root *dar-*) ‘rend, split, burst (tr.)’, and also English *tear*. The early Indo-Iranian **dar-*, from which the Sanskrit derives, could otherwise theoretically constitute an alternative source for the Samoyed word PS **tar-*, if Samoyed is regarded to independently have undergone a similar semantic shift as seen in Tocharian **tšər-*.

The development of PIE **d* to Tocharian *tš*, as in this word, is not well understood as a diachronic change, especially as regards intermediate stages and phonetic interpretation (see 2.3.4.3). It is thus uncertain how PIE **der-* would have been pronounced at the time that this word could be borrowed into early Samoyed, although

¹¹¹ The Tundra Nenets morphological dictionary (Salminen 1998: s.v.) gives an uncertain analysis of the root-final vowel as either *-ə* or *-a*, but Tapani Salminen (p.c.) has informed me that the correct interpretation of the root is *lyajka-* with an *-a*.

some kind of dental is likely. In theory, even starting from pre-PT $\pm *t^{\text{ser-}}$, one could also posit a borrowing in the shape of pre-PS $*sar-$, whence PS $*tar-$ by the well-known Samoyed change of PU $*s$ (and $*\xi$) to $*t$, depending on the time depth of the relevant changes (not easily established, see 3.4.1 and 3.9).

Formally, PS $*tar-$ ‘divide, separate’ can be derived from the ancestor of PT $*t^{\text{ser-}}$ ‘separate’, or from a Proto-Indo-Iranian source $*dar-$ ‘split, tear’, both of which descend from PIE $*der-$ ‘split, tear’. Semantically, the Tocharian verb provides the better match with the Samoyed, so that this is the more likely source.

9.3.4 pre-PT $*newo-$ ‘new’ \Rightarrow Pre-PS $*\acute{n}uwV$ ‘child, young boy’

The Samoyed word for child is reconstructed in SW as $*\acute{n}u\hat{a}$ (p. 111). However, a vowel sequence like $*u\hat{a}$ often derives from sequences $*VwV$, $*Vji$ and $*V\gamma i$ (Aikio 2012), which allows for an internal reconstruction of $*\acute{n}u\hat{a}$ as pre-PS $*\acute{n}uwV$. This is very close to pre-PT $*newo-$ (perhaps already $*[n\acute{e}wo]$, but the status of phonetic palatalization at the possible time of contact is unclear, see 4.4), whence TB $\acute{n}uwe$, TA $\acute{n}u$. The semantics are not identical: the Tocharian word really means ‘new’ rather than ‘young’, but the latter meaning is attested for cognates in other Indo-European languages, e.g., Gr. $\nu\acute{\epsilon}\omicron\varsigma$ ‘young, youthful’, Skt. $n\acute{a}va-$ ‘new, young’, and the semantic connection between the concepts ‘new’ and ‘young’ is close in any case.

Terms for children seem to be borrowed relatively easily: compare for instance the North Saami words $b\acute{a}rdni$ ‘boy, son’, and $g\acute{a}nda$ ‘boy’ with Swedish $barn$ ‘child’ and Old Norse $ganti$ ‘boy’ respectively (Sammallahti 1998: 230, 243) and Dutch colloquial $kids$ ‘children’ from English. On account of this, borrowing of a word for ‘child’ would not be out of the ordinary.

Nevertheless, the etymology remains tentative due to the lack of semantic exactness, as well as the theoretically possible alternative internal reconstructions for pre-Proto-Samoyed, namely $*\acute{n}uji$ or $*\acute{n}uyi$. The Samoyed word might be derived from early Tocharian, but it might also be of some yet unknown different origin.

9.3.5 pre-PT $*m\acute{e}ywya$ ‘power, strength’ \Rightarrow pre-PS $*mej\hat{a}$ ‘sturdy, solid’

We can reconstruct a word PS $*mej\hat{a}$ ‘sturdy, solid, strong’, with reflexes in all the northern Samoyed languages: Ng. der. $maag\acute{a}$, NeT der. caus. $mey^{\circ}ra-$ ‘strengthen’, der. $meyo$, NeF der. mju ‘id.’, EnF mu . It apparently mainly refers to ‘strong’ in the sense of ‘sturdy, solid’ (‘крепкий, прочный, твёрдый’ are the translations given, notably not ‘сильный’). This is slightly different from the Tocharian word TB $maiyya$ ‘power, might,

strength', which rather seems to be linked to some kind of general strength as potency.¹¹² However, the different shades of 'strong' are semantically close, and may have shifted somewhat over time. The word could in principle be borrowed from pre-Proto-Tocharian with loss of *-w-* in an original cluster *-ywy-* reconstructed by Adams (2013: 508), if it had not yet been simplified on the Tocharian side already, and a change from PIE **o* to **ɛ* reminiscent of the one found in pre-PT **kwonə-/kwënə-* > pre-PS **wən* 'dog'.

For both the Samoyed and Tocharian words there exist alternative loan etymologies. For Nganasan *maagəə* specifically, a connection with Evk. *maja* 'strong' has been proposed (Wagner-Nagy 2019: 503).¹¹³ However, this would separate the Nganasan word from its counterparts in Enets and Nenets, despite the regular phonological correspondence and identical semantics. A shared origin of Ng. der. *maagəə*, NeT der. *mey^ora-*, EnF *mui* 'sturdy, solid, strong' from PS **mejâ* is clearly to be preferred, and this cannot be Evk. *maja* 'strong'.

TB *maiyya*, meanwhile, has been connected to Iranian **māyā-* (Av. *māiīā-* 'craft, capacity, strength') with a development of **āy* to **ayy* in the Old Steppe Iranian language that was in contact with pre-Proto-Tocharian (Bernard 2023: 58–60). Adams (2013: 508) derives the Tocharian from TB *maiwe* 'young' with addition of a suffix **-ya*, i.e., **moywya-* with loss of the medial **-w-*, but this would separate TB *maiyya* from the very similar Indo-Iranian words (see also Bernard's criticism of this etymology). TB *maiwe* 'young' itself is not borrowed from Iranian, but it is formally not as easy to connect to Samoyed **mejâ* due to the cluster /jw/: such clusters occur in Proto-Uralic and Proto-Samoyed (e.g., PU **ojwa* > PS **âjwâ*), so that the adoption of pre-PT **moywo-* or **mëywë-* as **mejâ* rather than **mejwâ* seems strange.

Semantically, 'young' is also not the same as 'sturdy, solid, strong', although in principle a shift could be imagined on the Samoyed side from 'young' to 'new, newly made' to 'sturdy, solid', as the opposite of 'old and almost falling apart'. An advantage of *maiwe* as a source for the Samoyed would be that it is also an adjective, rather than *maiyya*, a noun, and that it is with greater certainty inherited in Tocharian rather than borrowed from Iranian. Perhaps a connection of the Samoyed word with Iranian **māyā-* can be considered instead, adding to a small group of words of this kind (e.g., PS **jaâ* 'flour' cf. Plr. **yawa-* 'grain'; Janhunen 1983: 123–124; Holopainen 2019: 105).

In conclusion, there exist similar words denoting some concept of 'strength' in Tocharian and Samoyed. The Tocharian noun TB *maiyya* 'power, might, strength' provides the closest phonological match to PS **mejâ* 'sturdy, solid, strong', but the

¹¹² Adams (2013: 508) cites passages where it refers to 'bodily strength' or abstract "force". Similarly, the derived adjective *maiyyätstse* is used to describe a powerful 'ignorance' and 'hatred', as well as a 'powerful man'.

¹¹³ Pointed out to me by Kaisla Kaheinen.

Tocharian etymology as a direct inheritance from Proto-Indo-European is highly uncertain. The adjective TB *maiwe* ‘young’ provides a worse phonological match, and a larger semantic difference between it and the Samoyed word requires a semantic change on the Samoyed side. A derivation of PS **mejâ* from Tocharian may thus be considered as a possibility, but it is definitely not the best example of this category. Some connection between the Samoyed and Indo-Iranian words could be considered instead.

9.4 Discussion

The number of promising loanwords from pre-Proto-Tocharian into pre-Proto-Samoyed remains very low. At the same time, some words do show quite tantalizing correspondences (like **mâŋkâ*, **lâŋkâ*, **laŋkâ-*, **mänäwjâ*). A summary of the evaluations is provided in Table 9.1.

Table 9.1: A summary of the loanword evaluations.

pre-Proto-Tocharian	Proto-Samoyed	Evaluation
<i>*septam</i> ‘seven (7)’	<i>*säjt³wä</i> ‘seven (7)’	rejected
<i>*wesa</i> ‘gold’	± <i>*wäsa</i> ‘metal, iron’	problematic
<i>*mēnē</i>	± <i>*mänäwjä</i> ‘full moon’	possible
<i>*monko</i> ‘lack, deficit’	<i>*mâŋkâ</i> ‘poverty, need’	plausible
<i>*kwēnäm</i> ‘dog obl.sg.’	<i>*wēn</i> ‘dog’	plausible
<i>*lonko</i> ‘steep slope, cliff’	<i>*lâŋkâ</i> ‘steep bank; steep’	plausible
<i>*lenk-</i> ‘dangle, sway, hang’	<i>*laŋkâ-</i> ‘dangle, hang’	possible
<i>*der-</i> ‘split → separate’	<i>*tar-</i> ‘divide, separate’	plausible
<i>*newo</i> ‘new’	<i>*ñuâ</i> ‘child, young boy’	possible
<i>*meywya</i> ‘power, strength’ or <i>*maywe</i> ‘young’	<i>*mejâ</i> ‘sturdy, solid’	problematic problematic

At first blush, the overall number of possible loans appears to contrast with those from early Indo-Iranian in the other Uralic languages, but the difference is not necessarily that significant at the level of individual branches. Holopainen lists 13 etymologies in the pre-Indo-Iranian category (Holopainen 2019: 336). However, if we check the distribution of these loanwords, it becomes clear that for instance Mansi and Khanty only have very few, one each in this category. This may be due to various factors that are difficult to test, including later lexical replacement. It is well-known that Samoyed has undergone some significant lexical replacement in even the basic vocabulary (e.g., Aikio 2002: 52), providing ample opportunity for secondary loss of any early loanwords. If there was only

one very old layer of borrowings from early Tocharian into early Samoyed, the number of reflexes is expected to be low. A lack of ample evidence for early loan contact is thus not a reliable indication that there was no early loan contact at all. We might imagine that retrievable early Tocharian loans from Samoyed behave according to a similar pattern as the earliest Indo-Iranian borrowings in Ob-Ugric.

That is not to say that the possible loans discussed in this chapter are (a) necessarily all accurately identified, as due caution is expressed where appropriate, or (b) incontrovertibly prove ancient contact between pre-Proto-Samoyed and pre-Proto-Tocharian; simply that the small number of etymologies is not an indictment of the proposed borrowings or of a contact hypothesis in and of itself. The limitations of this line of enquiry were already voiced by Janhunen (1983), as briefly addressed in the introduction to this paper as well. Furthermore, if the majority of phonological and grammatical contact phenomena are in line with influence on Tocharian through language shift from an early form of Samoyed, that particular variety Samoyed could logically not have been the direct ancestor of the modern Samoyed languages, so that limited borrowing from Tocharian might be expected. At least the new comparisons adduced here and the rehabilitation of the older suggestion of PS **māŋkâ* with a closer nominal origin in Tocharian should provide an illustration of the possibility for further results in this early contact situation, and brings us a step closer towards a better understanding of the prehistoric interactions of the early Samoyeds with Indo-Europeans.

