The Samoyed languages

Salminen, Tapani

Vajda, Edward

2024-03-04

http://hdl.handle.net/10138/593421

Salminen, T 2024, The Samoyed languages. in E Vajda (ed.), The languages and linguistics of Northern Asia : Language families. The World of Linguistics [WOL], vol. 10.1, De Gruyter Mouton, Berlin, pp. 167–251. https://doi.org/10.1515/9783110556216-004

Tapani Salminen

4 The Samoyed languages

Abstract: Samoyed languages, centred in north-western Siberia, represent a branch of the Uralic language family that is syntactically and morphologically conservative but phonologically and lexically innovative. Their general structure exhibits all the typical features of a consistent SOV language, save perhaps their complex morphophonology. In short, they are entirely suffixal and can be shown to have derived from an agglutinating proto-language which also had vowel harmony, variously preserved in modern Samoyed languages. The Samoyed languages have much in common with each other, yet they appear quite different in many aspects, which is the main topic of what follows. The chapter moves from Proto-Uralic through Proto-Samoyed to modern Samoyed languages, with selected topics of grammar discussed in separate sections.

4.1 Introduction

4.1.1 Overall characteristics

This chapter examines the structure and development of Samoyed, the easternmost branch of the Uralic language family. The family's internal diversification is shown in Figure 1.

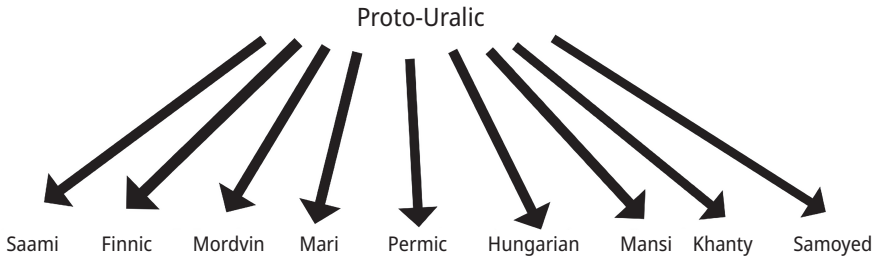

Fig. 1: The Uralic language family

The Samoyed languages constitute one of the nine primary branches of Uralic and represent a closely-knit but at the same time relatively diverse group of languages traditionally spoken by indigenous communities in north-western Siberia with extensions to south-western Siberia and north-eastern European Russia. They descend from an unattested yet quite straightforwardly reconstructable language, Proto-Samoyed, which existed roughly through the first millennium BC. In the light of language contacts in particular, the original homeland of Proto-Samoyed speakers was located in

https://doi.org/10.1515/9783110556216-004

the taiga region between the upper reaches of the Ob and Yenisei rivers, around the area where southern Selkup dialects were later recorded, while the expansion of Samoyed to both the tundra region in the north and the mountain region in the south must have involved both migration, i. e. movement of people, and diffusion, i. e. shifting to another language by aboriginal populations. Proto-Samoyed itself represents an early eastward migration of Proto-Uralic speakers whose language was heavily influenced by a local substrate. Proto-Samoyed was in contact with early forms of Indo-Iranian and Turkic, and possibly also with varieties of Mongolic, Tocharian, Tungusic, Yeniseian, and Yukagir. Most of these contacts took place around the beginning of the current era, when Proto-Samoyed had already started to diverge into its daughter languages.

Samoyed consists of seven sub-branches, viz. Nganasan, Enets, Yurats, Nenets, Selkup, Kamas, and Mator. The internal structure of the Samoyed branch is shown in Figure 2.

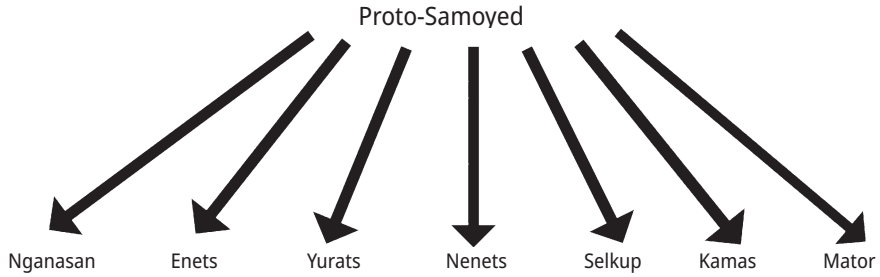

Fig. 2: The subdivisions of the Samoyed languages

Of these, Nganasan, Yurats, Kamas, and Mator represent single languages, while both the Enets and Nenets sub-branches are divided into two clearly separate languages, i. e. Tundra Enets vs. Forest Enets and Tundra Nenets vs. Forest Nenets, respectively. Selkup is a language complex with five distinct dialect groups, viz. Taz Selkup, Tym Selkup, Narym Selkup (also referred to as *h* Selkup), Upper Ob Selkup, and Ket Selkup. The fact that they have often been treated as dialects of one language or combined into a number of distinct languages will be discussed in more detail below. The various proposals to combine the uncontroversial sub-branches into intermediate subgroups are also referred to in what follows.

The demographic as well as geographic details of each language will be addressed in their specific sections, but in short, Tundra Nenets currently has by far the largest number of speakers, around 20,000, followed by Nganasan, Forest Nenets, and Taz Selkup, whose numbers of speakers range between 100 and 1,000. Both of the Enets languages have only a handful of speakers, and Narym Selkup is presumably spoken by two people, while Ket Selkup may still have one remaining though not recently contacted speaker. Tym Selkup, Upper Ob Selkup, and Kamas became extinct shortly before

Map 1: © Peter Palm, Berlin.

or after the beginning of the 21st century, whereas Yurats and Mator had already disappeared by the middle of the 19th century.

For Tundra Nenets, a written language has existed continuously since the 1930s, although with orthographic reforms, including a shift from Latin script to Cyrillic in the late 1930s. Tundra Nenets has since then been widely used in literature, and has a notable presence in media, to the extent that it is possible to study the language through

web searches. A Taz Selkup literary language briefly flourished in the 1930s in the Latin script, but few publications appeared after the subsequent shift to Cyrillic. For the remaining decades of the Soviet Union, only Tundra Nenets was used in print and, to a limited extent, in education. Shortly before the collapse of the Soviet Union, however, Taz Selkup was re-established as a written language in the form of a school dictionary (Irikov 1988), followed by similar publications in Forest Nenets (Barmič & Vėllo 1994), Forest Enets (Sorokina & Bolina 2001), and Nganasan (Kosterkina, Momde & Ždanova 2001). The other extant Samoyed idioms, Tundra Enets and Narym Selkup, although substantially different from Forest Enets and Taz Selkup, respectively, were never deemed worthy of literary expression by the Russian authorities.

The only accurate name for this group of languages in English is indeed Samoyed. It represents the totality of entities known in older literature as Tavgy Samoyed (Nganasan), Yenisei Samoyed (Enets), Yurak Samoyed (Nenets), Ostyak Samoyed (Selkup), and Sayan Samoyed (Kamas and Mator). Samoyed also enjoys primacy in English-language publications via Péter Hajdú's seminal *The Samoyed peoples and languages* (Hajdú 1963). The frequently encountered malformation "Samoyedic" derives from a calque of German *Samojedisch* or Russian *самодийский* but in English *Samoyedic* would productively refer to a group of languages wider than but named after Samoyed, say, a hypothetical but not an entirely implausible unit consisting of Khanty and Samoyed, for which, needless to say, "Khantyic" would work equally well. Similar cases within Uralic are Saami and Mordvin, not "Saamic" or "Mordvinic", respectively, while adjacent language families such as Turkic, Mongolic, and Tungusic are correct as they are formed on the basis of language names Turk-ish, Mongol-ian, and Tungus (Evenki), as is obviously Uralic as a derivative of a place name. The most extraordinary reason for preferring "Samoyedic" that I have met was that using Samoyed instead of the unnecessarily suffixed form would lead to confusion between Samoyed peoples and Samoyed dogs, but, since the canines in question are named after the humans, the claim should rather be that the dogs are renamed "Samoyedic" than vice versa.

Syntactically, Proto-Samoyed was very similar if not identical to Proto-Uralic, and in this respect it may well be regarded as the most conservative branch of Uralic. The modern Samoyed languages largely continue to exhibit features of an archetypal SOV language, insofar as they are strictly verb-final, with all typologically expected consequences, as detailed in what follows. The few exceptions in individual languages and dialects (other than Nganasan, see below) are clearly due to recent foreign, typically Russian, influence. The basic word order of a transitive sentence is, obviously, SOV, although when the focus is shifted to the subject, OSV would be used, in accordance with a more general rule that any focused constituent, representing new information, takes the preverbal position. As expected of a consistent SOV language, modifiers invariably precede their heads. Subordination is expressed by embedded non-finite complements, while coordination employs either parataxis, i. e. no marking at all, or connective adverbs, which, however, also have other functions, so that there are no word-classes such as relative pronouns or conjunctions. Question words appear in situ instead of

being moved to initial position. There are only postpositions, which consistently govern the genitive case. With some interesting exceptions discussed below, attributive modifiers show no agreement, and in basic predicative clauses, involving unmarked modal and temporal forms as well as excluding negation, no copula is required.

Morphologically, the Samoyed languages are also relatively conservative, as they are not only exclusively suffixal, but the bulk of their suffixes can be traced to Proto-Uralic. Proto-Samoyed was still presumably a more or less agglutinative language, while in the modern Samoyed languages, especially in the northern areal, a lot of language-specific morphophonology has emerged. In nominal inflection, presumably all of the Proto-Uralic number and case markers have been preserved in Samoyed, and the possessive suffixes appear to be almost equally archaic. Proto-Samoyed did innovate a prosecutive case, i. e. a case conveying a meaning such as 'via' or 'through', recently explained as agglutination of a Proto-Uralic postposition, and the Proto-Uralic translative case suffix turned into the (northern) Samoyed destinative marker, as discussed below. When it comes to derivation, there are few widespread Uralic derivational suffixes that lack a Samoyed cognate, and while the picture is more varied regarding verbal inflection, most elements are nevertheless of Uralic origin.

Phonologically, Samoyed is much more innovative than in terms of morphology or syntax, and the vowel system in particular has become notably richer. I am quoting Proto-Uralic vowels according to Aikio (2022), which, however, is identical to Janhunen (1981) except for one non-contrastive detail (Janhunen's *å = Aikio's *a), with eight Proto-Uralic vowel phonemes, i. e. *i *e *ä *a *o *u *į *ü. Proto-Samoyed underwent several phonemic splits producing five new vowels *ə̂ *ə *å *ę *ö, creating a system of 13 distinct units. The Proto-Samoyed consonant system, by contrast, lost several members of the presumed Proto-Uralic inventory. Notably, Proto-Samoyed had a single sibilant instead of the three sibilants *s *š *ś as well as the palatal affricate *ć customarily reconstructed for Proto-Uralic, but its status is given special attention below. The Proto-Uralic spirants *δ *δ' simply merged with *r and *j, respectively, while *γ (frequently, but inaccurately, referred to as a "laryngeal" and symbolized with *x) was vocalized or lost altogether. Prosodically, Proto-Samoyed continued to have predictable stress placement, with stress on the first syllable of any two-syllable foot. Phonotactically, Proto-Samoyed largely preserved earlier structures, e. g. PU *muna > PS *mə̂nå 'egg', *sarka > *tårkå 'branch', and *tumtį- > *tumtə̂- 'recognize', but it did acquire new canonical word structures *(C)VC and *(C)V, restricted to minor word classes such as pronouns and auxiliary verbs in Proto-Uralic, through apocope, e. g. PU *nimi > PS *nim 'name', and the loss of intervocalic *γ or *k, e. g. PU *kokį- > PS *ko- 'find', respectively.

Lexically, Samoyed is also on the innovative side, but not quite to the extent previously assumed. While several words otherwise widespread in Uralic, for example, PU *käti 'hand', *weri 'blood', or *säppä 'bile' as well as certain numerals, dealt with in section 6 below, lack a cognate in Samoyed, cf. PS *utå 'hand', *kęm 'blood', and *pətä 'bile', similar lacunae can also be found in other branches of Uralic (cf. Aikio 2002). Furthermore, thanks to the publications of Aikio in the first place, we now know that

Samoyed does have cognates for words such as PU *jį̈ɣį̈- 'drink' that were previously deemed to be "Finno-Ugric", referring to a long-held but now discredited hypothesis of a subgroup that consisted of all other branches of Uralic except Samoyed. At the same time, there are a number of Samoyed words that are often quoted as having a cognate in the other Uralic languages, notably PS *kâtå 'fingernail' or *paj 'stone' that are actually unrelated to, for instance, Finnish *kynsi* and *pii*, respectively. Nevertheless, a major part of the Samoyed core lexicon derives from Proto-Uralic, and exceptions can be straightforwardly explained as substrate influence from an earlier aboriginal language spoken in the Samoyed homeland.

Among the features supposedly unique to Samoyed, the destinative declension (see section 8.1), now explained as a direct development of the Uralic translative, has already been mentioned above. The contrast between subjective vs. objective conjugations in the inflection of transitive verbs (the number of the object being morphologically marked in the objective conjugation), based on the focused vs. non-focused status of the object (see section 8.2), and the consequently strict subcategorization between transitive and intransitive verbs (cf. section 9.2) have exact parallels in Khanty and Mansi, despite major differences in object marking within Siberian Uralic, in that the personal suffixes of the objective conjugation are or have been identical with the possessive suffixes of nouns. Another feature that creates subcategories among verbs is the category of aspect, which in Samoyed indeed appears much stricter than anything found in the other Uralic languages. Samoyed aspect bears a close resemblance to the Slavonic aspect system, but with a different time reference for the basic verb forms (cf. section 4.1).

4.1.2 General sources

The Samoyed languages have been documented very unevenly. Tundra Nenets has been recorded in a vast number of publications, including large dictionaries and major grammatical studies, with Taz Selkup taking a distant second place in this respect. Things have been improving in recent years thanks to a concerted effort of specialists in the other Samoyed languages, yet a notable imbalance remains. This is in marked contrast with the beginnings of the scholarly study of Samoyed, started by the pioneering Finnish field linguist M. Alexander Castrén, whose Samoyed grammar (Castrén 1854) and comparative lexicon (Castrén 1855) aimed at a fairly equal treatment of the Samoyed languages and sub-branches extant at the time, i. e. Nganasan, Enets, Nenets, Selkup, and Kamas. Brief word lists had been recorded for most of the Samoyed languages, with the notable exception of Forest Nenets, for which Castrén provided the very first documentation, and it is these early attestations that remain everything we have about Yurats and Mator, as evident below.

Following Castrén, there have been relatively few book-length publications discussing the entirety of Samoyed. Hajdú (1963), already mentioned above, offers a broad

introduction, while his Samoyed textbook (Hajdú 1968) focuses on Tundra Nenets and Taz Selkup. Tereščenko (1973) notably presents an early overview of the syntax of the Samoyed languages. The etymological dictionary *Samojedischer Wortschatz* (Janhunen 1977a) continues to be the most widely quoted source in the field, and this essay is no exception. Further milestones in Samoyed historical linguistics include the handbook of Mator (Helimski 1997) and the dictionary of Selkup (Donner, Sirelius & Alatalo 2004), both with extensive comparative data, as well as the posthumous work by Mikola (2004) on Samoyed historical phonology and morphology. Important chapter-length treatises include Tereščenko (1966), Hajdú (1988), Mikola (1988), Janhunen (1998), and Helimski (2001a). Major publications on individual languages and subgroups are mentioned below, although without aim at full coverage. Already at this point, it must be mentioned that no less than three substantial grammars, written in English, appeared in the last decade, i. e. Siegl (2013) on Forest Enets, Nikolaeva (2014) on Tundra Nenets, and Wagner-Nagy (2019) on Nganasan. As seen below, there are numerous monographs, mostly in Russian, on Selkup as well, and while Kamas lacks a reference grammar, its study has also progressed rapidly in recent years.

In spite of the large number of publications devoted to the Samoyed languages, a notable, and from an outside perspective unfortunate, feature of Samoyed linguistics is that several crucial sources only exist in the form of manuscripts. To start with Castrén, the extent of the material which Schiefner, the editor of his posthumous works, had left unpublished came as a kind of a surprise a few years ago. A project called *Manuscripta Castreniana* aims at creating critical editions of everything in Castrén's legacy, and while much has already been achieved, a lot remains to be done. Three lexical databases of vital importance collected by Helimski, on Nganasan, Tundra Enets, and Taz Selkup, respectively, currently circulate among specialists but are not available to the wider public. My own fieldnotes on Forest Nenets unfortunately remain unpublished as well, although this I hope will change soon. By far the most important manuscript within the scope of this article, however, is the detailed study of Selkup historical morphology by Alatalo, reflected throughout this essay.

4.1.3 Technical notes

The superscripts ^{1} ^{2} ^{3} found in a number of Proto-Samoyed reconstructions below represent "archigraphemes", introduced by Janhunen in SW (14) to circumvent gaps in the comparative material which preclude reconstructing exact proto-forms. For instance, the contrast between PS *č and *t is only preserved in (most of) Selkup, so when Selkup data is lacking, *t^{1} is employed to denote either *č and *t. Similarly, Selkup alone sheds light on a number of PS consonant clusters, the initial element of which is neutralized into a glottal stop elsewhere, in which case *t^{3} is used to cover no less than four PS consonants, *č *k *s *t. One additional archigrapheme used here is *i^{2} to denote uncertainty in reconstructing PS *i and *e (both SW *i), cf. Helimski (2005).

For the sake of simplicity, I only quote the back-vocalic variants of elements subject to vowel harmony, cf., for example, the Proto-Samoyed 3sg possessive suffix *-tå, which implicitly covers the respective front-vocalic variant *-tä. Two other options are found in literature: either consistently listing both variants or employing archiphonemes such as capital letters, e. g. *-tA.

Proto-Samoyed reconstructions presented below are often accompanied with a reference to *Samojedischer Wortschatz* in square brackets, e. g. PS *wenä [SW 175 *winä] 'strap'. If the SW reconstruction requires no update, only a page number is mentioned, e. g. PS *jåårå [SW 37] 'sand'.

While phonological transcriptions in general appear in italics as usual, Common Selkup (SkWb) is marked in boldface and Mator (MS) in italics and boldface, following the original practice of the sources. The lexical entries in these publications may be quoted with their running number instead of a page reference.

Most data in this chapter are presented in Finno-Ugric Transcription, close to the related systems employed for Turkic and other Siberian languages, so that perhaps only the symbols *ę į* as the back counterparts of *e i* require a clarification. Finno-Ugric Transcription has occasionally been called the "Uralic Phonetic Alphabet", which is a misnomer for every word in the term, as "Finno-Ugric" has been included in the name of the system from the very beginning, the system is decidedly linguistic rather than phonetic, and it by no means constitutes an alphabet. This incorrect term obviously arose in imitation of the International Phonetic Alphabet, or IPA, the use of which, as must be noted in rather harsh terms, has increasingly plagued Samoyed and Uralic studies in recent years. In short, IPA is a system designed by phoneticians for the study of phonetics, with no applicability to exclusively linguistic purposes, especially if the languages in question have a long history of using a superior system, which is clearly the case here.

Ideally, and only in apparent contradiction to the previous point, all languages and dialects should have a phonological transcription of their own, reflecting their particular phonemic contrasts in the most explicit manner. In this spirit, I continue to quote Nenets materials as I have done before, meaning that palatalization is expressed by the letter *y*, i. e. *CyV* represents the palatalized counterpart of *CV* for both Nenets languages, while the palatal glide, otherwise symbolized by *j*, is marked by *y* in Tundra Nenets, which, I should hope, does not cause major confusion. The symbol ° in the phonological transcription of both Nenets languages denotes the "schwa", a contrastive and syllable-carrying feature deriving from the Proto-Nenets reduced vowel *ə in unstressed positions as explained in more detail below. For example, the phonetic exponent of ° in forms such as TN *təm°na* (< PN *təməna < PS *tə̑ŋ-mə̑nå) 'about the summer' can be either an ultra-short vowel or extra length on the preceding consonant, occurring in free variation but always contrasting with zero, cf. *təmna* (< PN *təmna < PS *tə̑-mnå) 'still' (adverb).

4.2 From Proto-Uralic to Proto-Samoyed

That the attested Uralic languages are divided into exactly nine groups of closely related languages, viz. Saami, Finnic, Mordvin, Mari, Permic, Hungarian, Mansi, Khanty, and Samoyed, is uncontroversial as such, irrespective of the numerous efforts, traditional and more recent, to posit intermediate subgroups between these basic branches and Proto-Uralic. This picture is occasionally blurred in presentations where, for instance, the relationship between Mansi and Khanty, which are patently not closely related, is supposed to be on the same level as that between Komi and Udmurt, which actually are closely related. It is similarly untrue that Samoyed constitutes a major division of the Uralic language family as opposed to "Finno-Ugric", consisting of all the other branches, or that Samoyed would be in any major way different from the rest of Uralic (cf. Comrie 1981: 93–94), which shall become evident in what follows. Besides the long-held yet untenable peculiarity of Samoyed within Uralic, there are two other myths concerning the classification of the Samoyed languages. The first is the incorrect bifurcation of Samoyed into northern and southern sub-branches, while the second involves inaccurate language divisions with regard to Enets and Nenets on the one hand and Kamas and Mator on the other. These issues are dealt with in the next section.

With a couple of exceptions discussed below, there is a high degree of consensus about the consonant system of Proto-Uralic, which included four nasals *m *n *ń *ŋ, five stops and affricates *p *t *č *ć *k, three sibilants *s *š *ś, two spirants *δ *δ', three liquids *l *l' *r, and three approximants or glides *w *j *γ (or in terms of the place of articulation, labial *m *p *w, dental *n *t *s *δ *l *r, cacuminal or retroflex *č *š, palatal *ń *ć *ś *δ' *l' *j, and velar *ŋ *k *γ). As for the velar nasal *ŋ, I will later use the symbol *ŋ, which is much more commonly found in transcribing Samoyed languages and these symbols have exactly the same reference. I will not replace *γ with *x since the latter, after starting off as a rather unnecessary innovation by Janhunen (1981), does not reflect the evident position of this consonant in the system, i. e. a voiced velar approximant, which is still preserved unchanged in Khanty, e. g. PU *mi̯γi̯ > Eastern Khanty *męγ* 'earth', and which also makes the most sense in light of its development in the rest of Uralic, including Samoyed. Similarly, the spirants *δ *δ', reflected as such in Saami, e. g. PU *nüδi > North Saami *nađđa* 'shaft', should not be described as voiced stops. Transcribing them as **d and **d' should be explicitly explained as a technical simplification, albeit a confusing one in my view. On a more substantial level, the challenged status of PU *ć as opposed to *ś is discussed below in some detail, as is the PU palatal lateral *l' assumed here contrary to most recent treatments dealing with the Proto-Uralic consonant system.

As detailed below, the Proto-Samoyed consonant system simply represents a subgroup of the Proto-Uralic one, as it did not innovate any new phonemic contrasts, but eliminated six of the twenty PU units, i. e. *ć *š *ś *δ *δ' *γ, leading to a system of fourteen consonants *m *n *ń *ŋ *p *t *č *k *s *l *l' *r *w *j in PS. The first explicit description of the Proto-Samoyed sound system in SW (9) also contained fourteen consonants, yet with two differences addressed later, but in short, PS **ś, already attached with a question

mark in SW, was an allophone of *s conditioned by a following non-initial vowel *i, not yet recognized at the time, while the assumption of PS *l', unmentioned in SW, seems the only workable solution for a rare but significant contrast found in Proto-Samoyed and prior to it in Proto-Uralic.

In any case, Proto-Samoyed lost many of the phonological contrasts found among Proto-Uralic consonants. As for spirants, they merged straightforwardly with other consonant, viz. PU *δ > PS *r, e. g. *nüδi > *nir 'shaft', *śaδa- > *såråˑ 'rain', and PU *δ' > PS *j, e. g. *δ'ümä > *jimä 'glue', *kaδ'a- > *kåjä- 'leave', *śüδ'i > *sijə 'coal'.

The reduction of the PU sibilant system is slightly more complex. The original PU *s merges fully with *t in Samoyed, e. g. *sala- > *tålä- 'steal', *ipsi > *əptə 'smell', *mi̯ksa > *mi̯tə̂ 'liver', *pesä > *petä 'nest'. This is also the fate of the rare PU *š > PS *t, e. g. *kajšaw > *kåjtåw (> TN *xœdo*) 'epidemic' (cf. Aikio 2014a). It was long assumed that no PU etymologies pointing to *š were found in Samoyed, and indeed *š must have been a rather marginal phoneme in PU to start with, so the lack of relevant cognates would not have constituted a major surprise, but the above-mentioned comparison by Aikio seems too exact to be abandoned on the basis of tacit assumptions.

Another PU consonant that for a long while did not seem to be reflected in Samoyed material was *ć, for a similar reason of already being rare in PU, and possibly not contrasting with PU *ś at all, which is the latest view of Aikio (2022). Aikio (2012), however, still reconstructs a PU verb *kaća- 'give (as a gift)' with a Samoyed reflex found in the deverbal noun PS *kåså-w 'gift, offering' (> TN *xaso*), hence PU *ć > PS *s. It has also been generally assumed that PU *ś > PS *s, following the merger of PU *s with *t, and also because the attested reflexes of PU *ś overwhelmingly point to PS *s, e. g. *śilmä > *səjmä 'eye', *kośki̯ 'rapids' > *kåsə̂ 'dry riverbed'. As reviewed below, this view has been challenged recently, yet notably the issue does not involve any phonemic contrast within Samoyed.

Often excluded from the consonant inventories of both Proto-Uralic and Proto-Samoyed, *l' still seems the best candidate for the initial consonant of PU *l'i̯pśi̯ (or *l'i̯pći̯) > PS *l'ępsə̂ [SW 41 *jępsə̂ (? *jęə̂psə̂) ~ *lęə̂psə̂ (ng, en)] 'cradle' > Ng *labsə* ~ TE *l'it'o* ~ TN *yebc*° ~ Sk (№ 1452) **ćopsə** ~ Km *t'epsə* ~ Mt (№ 205) ***ćębsę*** ~ ***sębsę*** [? = *ǯębsə*], both on the basis of the Enets *l'* and the otherwise inexplicable variation between a lateral found in Enets and Nganasan and potential reflexes of *j elsewhere. My conclusion is supported by older sources such as UEW (№ 510) and MS (№ 205) which do reconstruct initial *l' for Proto-Uralic and Proto-Samoyed, respectively, while Aikio (2022) simply has *li̯pśi instead. The only other presumably valid example of an initial *l' in UEW (№ 514) with a reflex in Samoyed would be PU *l'upša (Aikio 2022 *lupsa) > PS *jə̂ptå [SW 35] 'dew', but since both *l' and *l would give *j before *u in PS, the point is moot from a Samoyed point of view. Curiously, initial PU laterals would only seem to be preserved before the vowel *i̯, of which another example would be the word for 'bone', reconstructed as PU *li̯wi by Aikio (2022), which in my view might be better reconstructed as PU *li̯ya, whose phonotactic structure, however, is not allowed in Aikio's framework (see section 7). Before other vowels, initial PU *l > PS *j, cf. PU *lumi̯ > PS *jom [SW 46

*jom-] ‘snow’ or *lewi- > *jew- > Sk (№ 1434) **ćü-** ‘shoot’. PU *l has also merged with *j in (primary or secondary) syllable-final positions, e. g. PU *śilmä > PS *səjmä ‘eye’, *tulį > *tuj ‘fire’, but has been preserved in intervocalic positions, e. g. *kala > *kålä ‘fish’.

The case of PU *ć is more tenuous, and Živlov in particular argues that there was no contrast between PU *ć and *ś, so that *ć is his preferred reconstruction not only for Proto-Uralic but also for Proto-Samoyed. Aikio (2022) agrees that there was no contrast in PU, but continues to reconstruct *ś. I still think that words for ‘cradle’ and ‘dowry’ discussed by Aikio (2022: 7) as well as some other etymologies such as the one involving Hungarian *csomó* and Finnish *solmu* ‘knot’, not mentioned by Aikio, would rather represent cases with a distinct *ć.

When it comes to Samoyed, Živlov’s arguments include the following points, where Proto-Samoyed reconstructions with *s have been traditionally assumed. (i) The Selkup reflex of *ns is **ś** in two words, PS *kunsâ [SW 77] > Ng *kunsə* ~ Sk (№ 2212) **küśə** ~ Km *künzə* ~ Mt (№ 596) ***kunǯɜ*** ‘urine’, and PS *kənsä- [SW 52 *kânsâ- (? ~ *kânså[1]-)] > TN *xəncye-* ~ Sk (№ 2167) **qaśə-** ~ Km *kănzə-* ‘cool down’. (ii) The Tundra Enets reflex of clusters such as *ns, *ms, and before front vowels *rk is *d′* (~ FE *s*), e. g. PS *âmså [SW 15] > TE *ud′a* ~ FE *osa* ~ TN *ŋəmca* ‘meat’, PS *märkä [SW 93 *merkä] > TE *med′e* ~ FE *mese* ~ TN *myercya* ‘wind’. (iii) The weak grade of *s* in Nganasan is *d′* (= *j* in a strictly phonological transcription, see below). (iv) PS *ę > Ng (> *į*) > *i* after *ń and *į* after *s, e. g. PS *ńęr > Ng ńir ‘cartilage’, PS *ńęrkå > Ng *ńirku* ‘aspen’, PS *sęjâ- > Ng *sįə-* ‘sink’, PS *sęr > Ng *sįr* ‘ice, salt, white’, PS *sęrå > Ng *sįru* ‘widow’ (in other positions PS *ę > Ng *a*). (v) PS *s > Mt ***k*** before front vowels, e. g. PS *sejə > Mt (№ 455) ***keje*** ‘heart’, PS *sünsə > Mt (№ 595) ***künǯü*** ‘chest’, cf. PS *sįrå > Mt (№ 891) ***sirä*** ‘snow’. (vi) Common Selkup **ć** > southern (= Upper Ob and Ket) Selkup *k* before front vowels, e. g. Sk (№ 1433) **ćü** > Ket *kū* ‘belt’, Sk (№ 1513) *ćinti > Ket *kindi* ‘bowstring’, cf. PS *jâptå > Sk (№ 1445) *ćaptu > Ket *t′aptu* ‘dew’, Sk (№ 1470) *ćęmə > Ket *t′ivvə* ‘bird cherry’. All of the above-mentioned points are noteworthy in the context, hence their detailed enumeration here, yet it would be typologically quite unusual to reconstruct Proto-Samoyed without sibilants at all. A provisional, even if a bit trivial, solution would be that Proto-Samoyed possessed both [s] and [ć] as allophones of a single phoneme.

Besides segmental consonant phonemes, Proto-Samoyed also inherited the Proto-Uralic phonotactics. Notably, there were, and largely still are in modern Samoyed languages, no initial or final consonant clusters, and internal consonant clusters consist of only two consonants. As for Proto-Samoyed, however, the glides *w *j were not counted as consonants in this respect but *wC *jC in final positions and *wCC *jCC as clusters were permitted freely. Since, as detailed below, preconsonatal glides only have segmental reflexes in Nganasan and Mator, and marginally in Kamas, this feature can only be seen through comparative analysis in Enets, Nenets, and Selkup, for instance, TN *yík°* ~ FN *wyeqk°* < PN *wyæqkə < PS *wajkkâ [SW 173 *wäjkkâ] ‘neck’, the only source of the Proto-Nenets *æ being a Proto-Samoyed *Vj sequence.

Another phonotactic restriction in Proto-Uralic concerned the ban on initial *ŋ *r *γ, which, again, applied to Proto-Samoyed as well insofar as the process of nasal pro-

thesis in northern Samoyed is regarded as secondary, which, while nearly universally accepted, was actually questioned a long while ago by Jaakko Sarvela, who also held other unorthodox views of Uralic historical phonology, including the status of sibilants along similar lines to the recent comments by Živlov discussed above. The issue is, however, moot in the sense that just like in the case of the Proto-Samoyed sibilant, no phonological contrasts are involved.

Words beginning with *r are occasionally reconstructed for Proto-Uralic but they rather represent loanwords spreading across parts of the Uralic language area, so that some branches gradually allowed initial *r. This is not the case of Proto-Samoyed or several of the attested Samoyed languages, where initial *r is still often substituted with *l* in loanwords, notably in the widespread word for 'Russian', reflecting the Nordic origins of the medieval *Rus'*, i. e. TN *lúca* etc. < late Finnic *ruocca 'Swedish'. More recent Russian loanwords include Ng *urubaakə* 'shirt' (possibly via Dolgan mediation) < *rubáxa*, or TN *lyes°ka* '(unleavened) flat cake, pie; (in Siberian dialects also) dough' < *rézka*, although in those Tundra Nenets dialects that have been subject to more extensive language contacts *ryes°ka* is allowed along with several other borrowings with initial *r* not only from Russian but also Komi and Northern Khanty. Selkup has many words beginning with *r*, all of recent origin accordingly, while a general sound change *l > *r* in Forest Nenets obviously resulted in *r* being present word-initially as well.

Needless to say, there were and are further restrictions on which consonants may combine as clusters. A notable innovation in Proto-Samoyed was the introduction of clusters consisting of an obstruent and a sonorant, such as *tw, *tm, or *kl, which did not occur in Proto-Uralic, cf. PS *wåt^{3}wâ [SW 173] 'bed' (> TN *waqw°*) or *pət^{3}mä [SW 115 *pât^{3}mä] 'sharp' (> TN *pəqmya*). Words containing clusters of this kind, obviously lacking cognates elsewhere in Uralic and effectively representing substrate loanwords, were not numerous in Proto-Samoyed although their number can be augmented by taking into account less widespread but nevertheless rather old words such as TN *səqla* ~ TE *sola* < PS *sât^{3}lå 'idiot' not found in SW because that would require a cognate in southern Samoyed.

A major development concerning consonant clusters involves PU *k, which, whether the first or the second segment, was lost in all combinations with another obstruent, e. g. PU *i̯kta- > PS *i̯tå- 'hang', *koska > *kåtå 'grandmother', *mi̯ksa > *mi̯tâ 'liver', *pučki̯ > *pučâ 'tube', as well as between a sonorant and a high vowel, for instance in the Samoyed words for (PU *kurki >) 'crane' (TN derivative *xəryo*). Otherwise, Uralic clusters largely survive in Proto-Samoyed, a further exception being PU *ŋs > PS *nt, cf. *ji̯ŋsi̯ > *(j)i̯ntâ 'bow'.

The vowel system of Proto-Uralic is widely described as containing eight phonemes *i *e *ä *a *o *u *i̯ *ü, and I follow here Janhunen (1981) and Aikio (2022) when it comes to their distribution and marking, only replacing Janhunen's *å with Aikio's *a, but at the same time retaining Janhunen's *i̯ in non-initial syllables as the harmonic counterpart of *i. As for the latter issue, Aikio and many other specialists only reconstruct *i in non-initial syllables for PU irrespective of vowel harmony, i. e. also in stems with initial

back vowels, yet assuming non-initial *į makes sense from a Samoyed point of view, as seen below. The high back illabial vowel *į in initial syllables has more recently been characterized, on varying bases, as *ę, i. e. a mid back illabial vowel, but, since no contrast is implied, the matter is rather moot, and reconstructing *į seems more accurate from a systemic point of view.

As noticed above, Proto-Samoyed has added five new vowels *ə *å *ę *ö to the Proto-Uralic vowel inventory to create a system of 13 vowel phonemes in comparison to the eight PU vowels. While simple phonemic splits, under complementary conditions, have occurred in most instances, the picture is complicated by the fact that the Proto-Uralic *ü apparently merged with *i in Proto-Samoyed, cf. PU *nüδi > PS *nir ‘shaft’, *δ'ümä > *jimä ‘glue’, and *śüδ'i > *sijə ‘coal’. Yet a new *ü has emerged in certain positions from PU *i, as in PU *wiγti ‘five’ > PS *wüət ‘ten’, but also from PU *ü, presumably via an intermediate stage of *i, as in PU *künti ‘fog’ > PS *küntə [SW 79 *küntâ] ‘steam’. What is more, Proto-Samoyed *ö, found only in a handful of cases and contested by Alatalo, derives from a secondary split from PS *ü, the only widely attested etymology being *könsä [SW 74] ‘insides’ > Ng *kunsį* ~ TE *sôd'e* ~ TN *syoncya* ~ Sk **śünćə**, and it may be possible to treat it as an allophone of *ü.

The other phonemic splits responsible for the enrichment of the Samoyed vowel system simply represent conditioned sound changes: (i) PU *u > PS *ə, e. g. *muna > *mənå ‘egg’; (ii) PU *i > PS *ə, as in *ipsi > *əptə ‘smell’; (iii) PU *į > PS *ę, as in *įptį > *ęptə ‘hair (on the head)’ or *pįnį- > *pęn- ‘put’; (iv) PU *a > PS *å, as in *sarka > *tårkå ‘branch, fork’, although this development is more common than the actual PU *a > PS *a, which is, however, found in several cases such as PU *a > PS *a, e. g. *taka- > *takə- ‘behind’.

The Proto-Uralic non-initial vowel system only had two contrasting entities, the harmonic pairs of a lower (*a / *ä) vs. a higher vowel (*į / *i, often characterized and symbolized as a reduced vowel *ə). Their primary reflexes in Proto-Samoyed are *a / *ä vs. *ə / *ə, respectively, yet two kinds of mismatches occur. First, apocope has eliminated the higher vowel after certain consonants, e. g. PU *nimi > PS *nim ‘name’ vs. *śüδ'i > *sijə ‘coal’, cf. also PU *suŋį > PS *təŋ (in Enets and Nenets) ~ *təŋə [SW 148] ‘summer’ possibly with an ancient variation. Second, in a number of cases, notably after an original consonant cluster, the PU low vowel is reduced in PS, as in *mįksa > *mįtə ‘liver’. Janhunen (SW) followed the PU pattern of minimal contrasts among non-initial vowel presuming only *å *ä *ə beyond the first syllable. It has, however, become increasingly clear that high vowels need to be included in the Proto-Samoyed non-initial system, for instance, PS *jänti [SW 43 *jentə] ‘bowstring’ (> Ng *d'enti* ~ TE *d'edi* ~ TN *yen*° ~ FN *jeny*° [< PN *jennyə] ~ Sk **ćinti**) does not derive directly from PU *jänti but must reflect an original derivative, possibly cognate with Finnic *jäntek id., but as yet details remain unclear. Helimski (MS) also postulated mid vowels in non-initial syllables, but they may be better analyzed as sequences such as *aw, cf., for instance, Helimski’s **sarö ‘rain’ for what is reconstructed as *såraw below (see section 5).

A phonotactic peculiarity of Proto-Samoyed, with direct retentions in Nganasan and Enets as well as in Yurats and Mator, are vowel sequences, e. g. PS *jåårå [SW 37] ‘sand’

> Ng *d'uəru* ‘dye’ ~ TE *d'aara* ~ TN *yara* ‘sandbank’ ~ Sk **ćūrə** ‘sandbank’, or *ńuâ [SW 111 *ńuâ (? *ńuâj)] ‘(own) child’ > Ng *ńuə* ~ TE *nio* ~ TN *nyú* ~ Sk **ńū** ‘small’ ~ Mt (№ 779) ***ńuh*** [? = *ńuγ*]. Aikio (2012) has explicitly disputed Janhunen’s idea that PS *Vâ < PU *Vγ (= *Vx in Janhunen 1981) sequences, yet I should think that at least in a minority of the relevant cases reconstructing such a PU sequence is well-founded. Notable examples include PU *wiγti ‘five’ > PS *wüət [SW 177 *wüt] ‘ten’ (> Ng *biiq* ~ TE *biuq*) already mentioned above, which Aikio (2022) reconstructs as PU *wij(i)t(t)i ‘five / ten’, but *j does not have vocalic reflexes in Samoyed (and ‘five’ is sufficient as the PU gloss since PS has undoubtedly experienced a semantic shift here), and PU *käγli > PS *käəj [SW 66 *keâj] ‘tongue’ (> Ng *śiəd'ə*, generalized from PS vowel stem *käəj-ə-), for which Aikio simply reconstructs *käli ‘tongue’; cf. also TE *sioro* id., recorded as (?) *siqoro* with an apparently unetymological glottal stop by Helimski, which could presumably reflect an earlier vowel stem *käəl-ə- if the phonemization is correct.

For Proto-Samoyed, Janhunen (SW) only reconstructed vowel sequences consisting of a full vowel plus the reduced vowel *ə, while Helimski (MS) further assumed vowel sequences where the latter component is a low vowel, e. g. PS *tua [SW *tuâj] ‘feather, wing’ > Ng *t'üə* ~ TE *tua* ~ FE *tue* ~ Yr ‹túa› ~ TN *to* ~ FN *tŏ* ~ Sk **tū** ~ Mt (№ 1070) ***tua***, ***tuga*** (? ***toga***) [? = *tuγa*], where the reflexes of Tundra Enets, Yurats, and Mator (for *γ*, see below) in particular would have simply retained the PS form according to Helimski. Janhunen obviously noticed the contrast but applied internal reconstruction to the latter group involving a final *j, apparently on the basis of accusative plural forms such as TN (*ya* ‘earth, place' :) *yo* < PS *jåə-j [SW 36]. This analysis, however, does not work for either Nganasan or Mator, where final *j is regularly preserved, cf. Ng *ŋüaj* ‘island’ < PS(?) *wuaj, Ng *juəj* ‘fish dam’ < PS *juâj > Mt (№ 234) ***čuj***, or Ng *ŋoj* ‘foot, leg’ < PS *åj > Mt (№ 9) ***aj***, so that, while questions concerning particular reflexes remain, Helimski’s solution is clearly superior.

In both Nganasan and Enets a large number of secondary vowel sequences have developed through a loss of an intervocalic consonant, e. g. PS *kəńəli [SW 53 *kâńâlâ (? *kâjnâlâ ~ *kâjâlâ)] > Ng *kəəlį* (~ TN *xəyəl°*) ‘tear’, *kajwå [SW 63 *käjwå] > TE *sea* ~ FE *see* (~ Ng *kajbu* ~ TN *syíwa*) ‘spade, shovel’. Nenets also has vowel sequences, but they invariably result from contractions after a loss of an intervocalic consonant, while the original PS vowel sequences are reflected as single vowels. In Tundra Nenets, the most frequent contraction concerns the so-called aorist, marked by PS *-ŋå (except in Nganasan as explained below), e. g. PS *ko-ŋå-tå > TE *kôaδa* ~ TN *xo°da* ~ FN *koŋata* ~ Sk **qoŋətə** ‘(s)he found it’. It has been suggested that there might have been another aorist marker consisting of a single vowel (cf. Janhunen 1998: 472), yet the only possible predecessor of TN *xo°da* is a form that contained a consonant (< PN *koŋata, as preserved in FN), and the idea that a similar element could be identified inside the Nenets plural objective marker *-jə (as inTN *xoy°da* < PN kojəta ‘(s)he found them’) or the Nganasan perfective aorist marker *-qə* are manifestly outdated.

4.3 Problems and prospects in Samoyed internal taxonomy

As evident from above, Samoyed is one of nine primary branches of Uralic, with its own peculiarities, yet not notably more distant from Proto-Uralic than the other branches of the family. This section deals with the other remaining myths, or dogmatic positions if you wish, concerning the taxonomy of Samoyed.

First, the traditional division between northern (Nganasan, Enets, Yurats, and Nenets) and southern (Selkup, Kamas, and Mator) Samoyed languages is patently obsolete, as the few innovations that supported it to start with can be shown to be retentions or secondary areal influences, cf., for instance, the discussion on nasal prothesis typical of northern Samoyed in the phonology section below. Alatalo has suggested a simple remedy to the problem by moving Mator, the southernmost of the Samoyed languages, to a redefined "northern Samoyed", which would leave Selkup-Kamas, further discussed below, as a primary subgroup within Samoyed, yet the enlarged "northern Samoyed", even disregarding the terminological issue, still appears to lack a substantial number of indisputably shared innovations.

As an example of a sound change that could formally meet the criteria of a northern Samoyed innovation one might mention the merger of the PS word-final *n > *ŋ, as evident from PS *wẹn [SW 173] 'dog' > Ng *baŋ* ~ TE *buq* ~ TN *weh* ~ FN *wĕŋ* ~ Sk (derivative) **kẹnaŋ**, insofar the Enets and Tundra Nenets mergers with the glottal stop can be assumed to represent *ŋ. A more serious setback for the assumption, however, would be the fact that the PS final *ń appears to be treated differently, viz. > *j* in Nganasan and > *ŋ* in Forest Nenets.

The Sayan Samoyed languages, or rather the Samoyed languages spoken in the Sayan region, Kamas and Mator, undoubtedly share several innovations, but they invariably derive from parallel influences from the neighbouring Turkic languages. These obviously include many lexical changes caused by an influx of loanwords, but also phonological features, notably involving glottalization as well as nasalization in the case of PS initial glides when a word-internal nasal follows, as seen below.

Janhunen (1998), realizing the problems with the old northern-southern division as well as the numerous unique features of Nganasan, created an alternative binary classification, which, although in my view not strictly necessary, corresponds much better to the actual relationships among the languages. In his taxonomy, Nganasan would be the first to branch off, followed by Mator, which leaves a large Enets-Nenets-Selkup-Kamas entity as the core group as opposed to the geographically most peripheral languages. This "core Samoyed" would then be split into Enets-Nenets (obviously including Yurats) and Selkup-Kamas as the lowest-level intermediate sub-branches within Samoyed. All of the nodes implied by this novel scheme may have some basis in actual data, Enets-Nenets and Selkup-Kamas in particularly sharing a number of morphological and lexical features. At the same time, it is possible to find counterexamples of apparent innovations

that unite, say, Nganasan and Enets, or Nenets and Selkup, and therefore highlight the chain-like structure of Samoyed, obscured by any binary classification. More crucially, even the supposedly closest entities, Enets-Nenets and Selkup-Kamas, share so few actual sound changes that the potential Proto-Enets-Nenets and Proto-Selkup-Kamas reconstruction levels would not differ significantly from Proto-Samoyed.

One sound change uniting Enets and Nenets that became clearly recognizable only now that the SW vowel system has been updated is the fronting of PS *a [SW *ä]. This is most conspicuous in connection with the assibilation process mentioned below, cf. PS *kat [SW 65 *kät] 'face' > TE *seq* ~ TN *syaq* ~ Sk (№ 1874) **qāt** 'forehead', or *kajwå [SW 63 *käjwå] 'spade, shovel' > Ng *kajbu* ~ TE *sea* ~ TN *syíwa* ~ Km *kō* 'spade, shovel, oar'. A similar development may have taken place in Mator, but as seen above, is definitively absent in Nganasan. In morphology, the most striking Enets-Nenets innovation is the emergence of a preterite marker attached to the personal suffix, i. e. appearing as the last element of a word-form, e. g. TE *oδi-ma-do-śi* 'you showed up' (in the distant past as opposed to the aorist form *oδi-ma-do* 'you showed up' in the immediate past) ~ TN *ŋədyi-mya-nə-sy°* (: aorist *ŋədyi-mya-n°*) < Pre-Enets-Nenets (?) *ətiwm-mä-ntə-si. At the same time, the original PS preterite suffix *-så has developed into an interrogative mood marker in Forest Enets and the Nenets languages, cf. late FE *uu mosra-sa-d* 'Did you work?' (Siegl 2012: 402) ~ TN (*pidər°*) *mənc°ra-sa-n°* id. For further information see Siegl (2012) and the literature therein. Another major topic in this context, the personal pronouns, is discussed in section 5.

As in the case of Enets-Nenets, there are morphological and lexical peculiarities that join Selkup and Kamas. Within morphology, the most notable commonality is perhaps a secondary comitative case, with **-sē** in Selkup and *-se'* in Kamas. Among the many lexical items shared by Selkup and Kamas alone is the very basic verb Sk (№ 14) **ī-** 'take' (cf. TN *me-* 'take' with cognates in Enets and Mator). Selkup and Kamas also show similar conditional reflexes of PS *å, although actual shared innovations are few.

The second dogmatic position still held by many specialists, albeit in a decreasing manner, sees Enets and Nenets as single languages. As mentioned in the introduction and evident from what follows, both Enets and Nenets are not languages but subgroups consisting of two languages each, neither of them coming even close to a borderline case between language and dialect, which, on the other hand, could be said about Selkup. Curiously, while older literature was content with an "Enets language" and a "Nenets language", it was until recently customary to refer to the dialects of Kamas (Kamas proper and Koibal) and Mator (Mator proper, Taigi, and Karagas) as separate languages, despite the fact that they are both actually quite unified, with minor dialectal differences in phonology and lexicon, i. e. nothing comparable to the Enets and Nenets subgroups. Furthermore, two additional "languages" have occasionally been listed within (the strictly areal grouping of) Sayan Samoyed, viz. "Abakan Samoyed", which, however, is a phantom idiom based on a corpus with a mixture of Kamas and Mator materials (Klumpp 2019), and "Soyot", which simply represents a synonym of the Karagas dialect of Mator.

In what follows, each of the Samoyed languages will be discussed briefly, with concise information about dialects, demography, and geographic distribution. The focus, however, is on historical phonology, with some comments on lexical differences as well as references to morphological and syntactic issues. To start with, I mention six post-Proto-Samoyed innovations that have such a wide distribution that will be relevant to most subsections below.

(i) The odd Proto-Samoyed affricate *č has merged with *t everywhere except in Selkup, and even there the same merger has occurred, although apparently very recently, in part of the language area.

(ii) Syllable-final *s, *t (as well as *č), and *k have been replaced with a glottal stop in what has been traditionally called northern Samoyed, but they also have glottalized reflexes in Sayan Samoyed. Glottalization is absent in Selkup.

(iii) Proto-Samoyed *k has been palatalized and fricativized into *ś* before front vowels in most languages, although with different scopes, notably PS *ü (or *ö) does not trigger this change in Nganasan, cf. PS *küntə [SW 79 *küntə̑] 'steam' > Ng *kintə* (~ TE *śudo* ~ TN *syun*°) vs. PS *kitä 'two' > Ng *śiti* (~ TE *śiδe* ~ TN *syidya* ~ Sk **śitə** ~ Km *šide* ~ Mt ***kide***). As seen from the Mator reflex, PS *k has not undergone this change there, and what is more, as mentioned above, Mator shows a typologically unusual counter-change PS *s > Mt ***k*** before front vowels, e. g. PS *sejə > Mt (№ 455) ***keje*** 'heart'. All in all, the lack of uniformity within Samoyed points to an innovation that spread to most parts of Proto-Samoyed at a time when it was already disintegrating.

(iv) Proto-Samoyed intervocalic *ń has only been preserved in dialects of Mator and Selkup as well as, apparently, next to a vowel sequence in Nganasan (cf. *poə̑ń-ə̑-t > Ng *huońəq* 'borders, junctures'), but merged with *j everywhere else, cf. Sk (№ 2072) **qańi** 'tear' < PS *kəńi ~ *kəńəli > TN *xəyəl*° ~ Ng *kəəlị̈* id.

(v) The Proto-Samoyed reduced vowels *ə̑ and *ə have merged in the initial syllable everywhere, the only potential exception being the Mator reflex of PS *tər [SW 149 *tə̑r] > Mt (№ 991) ***ter*** 'body hair'. They do, however, contrast in the reflexes of Nenets non-initial sequences *ə̑j (> TN *o*) vs. *əj (> TN *ye*), for instance in the TN accusative plural forms *ŋebto* < PS *e̮ptə̑-j (: *ŋebt*° < PS *e̮ptə̑ 'hair (on the head)') vs. *ŋəbtye* < *əptə-j (: *ŋəbt*° < PS *əptə [SW 16 *ə̑ptə̑] 'smell'), cf. also TN acc.pl *tərye* < PS *tər-ə-j : 3sg *tərtya* < *tər-tä of the above-mentioned PS *tər [SW 149 *tə̑r] > TN *tər* 'body hair'. As suggested by SW reconstructions such as *ə̑ptə̑ and *tə̑r, the contrast between *ə̑ and *ə was established only later by Helimski (1993), where the symbol *ə suggested by him and included in the title can, and should, simply be replaced with *ə. Helimski based the distinction on Nganasan vowel harmony where *ə triggers the same harmonic variants as any other front vowel in the same manner as PS *tər-tä > TN *tərtya* above, cf. PS *mir-tä > TN *myirtya* 'its price'.

(vi) The initial *w in the sequence *wo- is lost (making the resulting *wo- subject to nasal prothesis where applicable) outside Selkup, where it is reflected as **ko-** as explained below, cf. *wota [SW 177 *wotɜ] 'berry' > Ng *ŋuta* ~ TE *ôδe* ~ Yr ‹ngóde› ≈ ***ŋoδä*** ~ TN *ŋodya* ~ Sk (№ 1917) **kotə** 'bog bilberry'.

The most notable feature that appears to unite members of the "core Samoyed" group, i. e. Enets-Nenets-Selkup-Kamas, would be the loss or merger of syllable-final glides. This change created highly complex morphophonological alternations in Enets and Nenets in particular, enriched the vowel systems of Nenets and Selkup, and is connected with secondary palatalization in (at least) late Kamas. Examples can be seen in PS *tuj [SW 166] 'fire' > Ng *tuj* ~ TE *tu*(*u*) ~ FE *tu* ~ Yr ‹tu› ~ TN *tu* ~ FN *tŭ* ~ Sk **tŭ** ~ Km *šü* ~ Mt ***tuj***, or (without a Selkup cognate) PS *åj [SW 17] 'foot, leg' > Ng *ŋoj* ~ TE *ŋa* ~ FE *ŋo* ~ TN *ŋæ* ~ FN *ŋă* ~ Km *üjü* / *ujə* ~ Mt ***aj***; in Mator, the 3sg possessive form was ***asta*** (= *ašta*), where *š* is the regular reflex of *j before obstruents. Another notable cognate set is Ng *ŋənduj* ~ TE *odu* ~ FE *odu* ~ Yr ‹ngáddu› [≈ *ŋədu*] ~ TN *ŋəno* ~ FN *ŋănu* ~ Sk **antu** ~ Km *eńi* / *äńi* ~ Mt ***ondoj*** 'boat', earlier reconstructed either as *ântâj [SW 15] or more accurately as *əntoj (MS № 809), which in the current framework should be rendered as *ântâwj (see section 5). In consonant clusters where *j is followed by a sonorant, both Nganasan and Mator have preserved the glide, e. g. PS *kåjmå [SW 58 *kåjmå (? *kåâjmå)] 'bone marrow' > Ng *kojmu* ~ TN *xœwa*, or PS *påjnå [SW 115] 'mattress' > Mt ***hajŋa*** ~ TN *pœŋa*, with a secondary vowel derived exclusively from *Vj sequences in Nenets. Curiously, there appear to be a few cases where first-syllable *Vj sequences have been preserved in Kamas, which further weakens, if not obviates, the "core Samoyed" hypothesis, e. g. Km *sejʼbə* 'seven' < PS *säjt³wə [SW 139 *sejt³wâ (? *sejkwâ)].

4.4 Samoyed languages and their subgroups

4.4.1 Nganasan

The Nganasan language (Ng *ńaa buəδu*) is spoken on the Taimyr Peninsula by perhaps 100 people. Nganasan is a highly unified language, with minimal phonetic and lexical differences between the western (Avam) and the eastern (Vadeyev) dialects. Wagner-Nagy has published not only a large grammar but also a monograph on derivational morphology (Wagner-Nagy 2001, 2019). The school dictionary edited by Gusev (Kosterkina, Momde & Ždanova 2001) contains not only lexical but also morphological information in a useful format. Katzschmann (2008) compiled a chrestomathy based on several publications, while a traditional handbook was already published by Tereščenko (1979).

Nganasan shows numerous unique features that separate it from the rest of Samoyed, of which three productive morphophonological alternations, consonant gradation, vowel harmony, and nunnation, discussed below, make the structure of Nganasan in many ways strikingly complex. Nganasan also possesses a substantial number of words without cognates in related languages or loan originals in known contact languages. A particularly interesting case is the culturally significant word for 'shaman', which in Nganasan would be *ŋəq*, which obviously looks like an archetypal ancient basic word from a phonotactic point of view, while the other Samoyed language share a word

which would be reconstructed as PS *tåtäpä (> TN *tadyebya*), which not only seems morphologically complex, perhaps a petrified participle in *-pä, but the stem is also disharmonic for an unknown reason, all of which makes it seem that it represents a lexical innovation shared by a subgroup ranging from Enets to Mator.

The major features uniting the other Samoyed languages as opposed to Nganasan are, however, grammatical. In phonology, the Proto-Samoyed contrast between *i and *e is only preserved in Nganasan (Helimski 2005) in the form of *i* vs. *į* as it happens, the reflex of PS *e being still represented by a likely *ę* in old Nganasan as recorded by Castrén however, cf. PS *wenä [SW 175 *winä] 'strap' > Ng *bįnį* (Castrén ‹béne› = *bęnę*) ~ TN *yinya* vs. *wiŋəw [SW 176 *wiŋâ (~ *wijâ)] 'son-in-law' > Ng *biŋi* (Castrén ‹biŋiŋ›, ‹biŋi›) ~ TN *yiy*° etc., the Tundra Nenets cognates representing the merger *e > *i that took place in all other Samoyed languages.

In morphology, two major isoglosses separate Nganasan from the other Samoyed languages. The locative (see section 6) marker in Nganasan reflects PS *-ntâ-nå (> Ng *-ntənu* with allomorphs) while the so-called co-affix is different elsewhere, cf. Ng *kəntə* 'sledge' : *kəndətənu* (< Pre-Nganasan *kânčâ-ntâ-nå), where nunnation does not apply, hence the locative allomorph *-tənu*, but the weak grade in *kəndə-* points to an original consonant cluster. The other languages reflect PS *-kâ-nå instead as in *kânčâ-kâ-nå > TE *kodoxone* ~ TN *xən°xəna* ~ Sk **qančōqən** 'in the sledge'.

An even more fundamental difference between Nganasan and the rest of Samoyed concerns the so-called aorist marking, i. e. the most basic forms in verbal inflection as regards the category of aspect. Outside Nganasan, aspect has no formal exponent but is only shown in the time reference, for example, TN *xada-* 'kill' (a perfective verb) vs. *nyoda-* 'follow' (an imperfective verb) share the same morphology to create forms such as *xadaəw*° 'I killed it' as opposed to *nyodaəw*° 'I follow it', i. e. otherwise unmarked finite forms of perfective verbs refer to immediate past while those of imperfective verbs express present time. Nganasan, by contrast, has two unrelated suffixes *-qə* for perfective verbs and *-ntu* (with allomorphs) for imperfective verbs, e. g. Ng *kotu-* 'kill' : *koδaqamə* 'I killed it' vs. *d'eru-* 'not know' : *d'erutumə* 'I do not know it'. The background of the suffix *-qa* remains unclear to me, while the origin of *-ntu* is clearly the PS imperfective participle suffix *-ntå (whereas the Nganasan imperfective participle in *-ntuə* reflects a combination of the original suffix and an adjectival marker *-jâ). The other Samoyed languages invariably employ the suffix *-ŋå in their stead, e. g. TN *xadaəw*° (< PS *kåtå-ŋå-mâ) means 'I killed it', while a cognate of *-ŋå appears to survive in Nganasan in the form of an (aorist) interrogative marker, so that the reflex of PS *kåtå-ŋå-mâ would be Ng *kotuŋumə* 'did I kill it'.

The Nganasan consonant system is quite conservative, and its development from Proto-Samoyed can be described in a straightforward manner, yet on a synchronic level there appear three instances of allophony that are usually included in the transcription of Nganasan, and I will follow this practice, as undesirable as that may be from a strictly structural point of view. First, PS *p has been fricativized in Nganasan, and recorded as ‹f› by Castrén, but further delabialized in modern Nganasan to a glottal fricative *h*

except before the vowel *a*, where it is still represented as a bilabial fricative *φ* which must have been the universal pronunciation before. This *φ* is often spelled as ‹xy›, and most phonological analyses of Nganasan assigned the labial feature to the following vowel instead. Second, PS *j in initial and intervocalic positions (where it has only been preserved after vowel sequences) has undergone a phonetic sound change yielding a voiced stop *d'*, which despite this obviously complementary distribution, is often found in specialist literature. Third, the weak grade reflex of PS *t appears as *δ* intervocalically and *d* after a nasal, again with no phonological contrast (expect perhaps in recent loanwords), yet both are kept in transcription.

In comparison with Proto-Samoyed, Nganasan has the same four nasals *m n ń ŋ*, four voiceless stops *t t' k q* (= glottal stop) as opposed to PS *p *t *č *k, three fricatives *h*/*φ* (< PS *p), two sibilants *s ś* where PS only had one sibilant *s, three liquids *l l' r* (similarly to PS but with a caveat discussed immediately below), and four voiced stops or approximants *b d*/*δ d'*/*j g* in contrast with only *w *j in PS. The palatal consonants *t' ś l'* have resulted from a relatively recent palatalization process, and, notably, the rare PS *l' merged with *l before such a process applied, i. e. *l'ẹpsə̂ ‘cradle’ > Ng *labsə* (~ TE *l'it'o*). Nganasan *b* has two origins, < PS *w in syllable-initial position and as the weak grade of *p via consonant gradation. The same goes with *d'* to the extent that it remains an allophone of *j* as noted above but also represents the weak grade of *s. Furthermore, gradation is the only source of *d*/*δ* as the weak grade of *t and *g* as the weak grade of *k.

Consonant gradation in Nganasan involves the weakening of the single obstruents *t k s*/*ś h* into *δ g d' b* (but compare the so-called nunnation below), and the homorganic nasal-obstruent clusters *nt ŋk ns*/*ńś ŋh* into *nd ŋg ńd' mb*, respectively. Two types of gradation can be distinguished as follows.

In syllabic, or radical, gradation both the single consonants and the clusters in question weaken in the beginning of an originally closed syllable, e. g. *kuhu* ‘fur’ : plural *kubuq* (< PS *kopå : *kopå-t) or *kəntə* ‘sledge’ : *kəndəq* (< *kânčə̂ : *kânčə̂-t). The word “original” is crucial in the context, as many of the conditioning syllable-final consonants have disappeared in the course of the history of Nganasan, cf. the respective accusative and genitive forms) *kubu* (< PS *kopå-m / *kopå-n) or *kəndə* (< *kânčə̂-m / *kânčə̂-n) that are the result of a recent loss of final nasals, still preserved in Castrén’s material. What is sometimes called reverse gradation refers to a special case where the stem itself ends in a consonant and adds an epenthetic vowel. For instance, the genitive of *nagür* ‘three’ (< PS *nakur [SW 99 *näkə̂r]) would be *nakürə* (< PS *nakurə̂n) on the same basis of the (original) syllable structure.

In rhythmic, or suffixal, gradation, which only concerns the single obstruents, weakening takes place after an unstressed syllable, for instance, *t > *δ* is found in the 3sg possessive form *kuhuδu* (< PS *kopå-tå) but not in the corresponding form *bakunutu* (< PS *wekånåtå), cf. *bakunu* ‘sturgeon’ < *wekånå [SW 174], because the third syllable in *bakunutu* carries secondary stress. That vowel sequences consists of two syllables is easily shown by this type of gradation, cf. the 3sg forms *ńuəδu* of *ńuə* ‘(own) child’ and *d'uərutu* of *d'uəru* ‘dye’.

As shown by Gusev (2008), a final *w needs to be reconstructed in Proto-Samoyed in many instances where it was unrecognized in SW and regarded as a secondary extension in MS. Interestingly, its reflexes are ***h*** [? = *γ*] in Mator and *ŋ* in older Nganasan, which, however, has been lost along with other final nasals in the modern language, while its labiality is reflected in the preceding, or in the case of modern Nganasan, final, vowel, e. g. PS *kåčâw [SW 57 *kåcâ] > Ng (Castrén ‹koduŋ› >) *koδu* ‘blizzard’, PS *såraw [SW 135–136 *sårâ-jâj] > Ng (Castrén ‹soruaŋ› >) *sorå* ‘rain’. The weak grade in *koδu* further points to a final consonant, while a contrastive *å* (see below) found in *sorå* only appears in this particular context.

Another morphophonological alternation related to nunnation and affecting suffixal consonants in particular is nunnation, whereby a suffix-initial nasal is present if and only if the last consonant of the preceding stem is also a nasal, cf. the locative forms *bi̯ni̯-ntəni̯* of *bi̯ni̯* ‘strap’, *kümaa-ntənu* of *kümaa* ‘knife’, or even *məu-ntənu* of *məu* ‘earth’ as opposed to *koli̯-təni̯* of *koli̯* ‘fish’ or *kəndə-tənu* of *kəntə* ‘sledge’. The final form reveals that the consonant cluster is original, i. e. *kəndə-tənu* < PS *kânčâ-ntâ-nå (with the specifically Nganasan co-affix *-ntâ).

In contrast with the relatively archaic nature of Nganasan consonants, the Nganasan vowel system has experienced a major upheaval. It is usually, in a tradition deriving from Helimski, described as containing four high vowels *i i̯ ü u*, three mid vowels *e ə o*, one low monophthong *a*, and two (low) diphthong[oid]s ^{i}a ^{u}a in the initial syllable. In that position, however, it turns out that there is never a three-way contrast among the mid vowels. Instead, *ə* can actually be interpreted either as an allophone of *e* or *o* depending on the preceding consonant. For instance, in above-mentioned words *məu* ‘earth’ and *kəntə* ‘sledge’, the initial CV sequences would phonologically be /mo/ and /ke/, respectively, yet I continue employing the over-differentiating practice. The initial element of the “diphthong” ^{u}a rather represents a feature of the allophone *φ* discussed above, hence Helimski’s *h*^{u}a- is replaced with *φa*- except in non-initial syllables where ^{u}a actually is contrastive in a few positions, where *å* is used below instead of the traditional marking. A similar notational simplification is carried out in the case of ^{i}a, contrastive both in initial and non-initial syllables and marked with *ä* in what follows. Besides the unusual distribution of *å*, the system of non-initial syllables is similar to that of the initial syllables, except that no contrast appears among mid vowels, but only *ə* is transcribed there.

Disregarding certain subsequent neutralizations for the moment, Nganasan has undergone chain-type vowel changes, i. e. PS *ü > Ng *i*, *u > *ü*, *o / *ö > *u*, and *å > *o*, as well as *i̯ > *i*, *e > (*ę* >) *i̯*, *ä > *e*, and *ę > *a*, while the PS reduced vowels *â *ə, while evident in the functioning of vowel harmony, have merged into what is transcribed here as *ə* despite its actual contrast with *o* or *e* having been neutralized depending on the preceding consonant as described above. As it happens, only two PS vowels, *i and *a, can be said to remain unchanged in Nganasan. The vowel *ä* mentioned above seems to represent a special development of PS *ä adjacent to the glide *j which may have been subsequently lost. In non-initial syllables, the low vowels exhibit changes *å > *u* and *ä

> (*ę* >) *į* (see the examples in the next paragraph) as if jumping one step further than the initial-syllable vowels. A notable exception to this process is the preservation of non-initial *å as a low vowel before a syllable-final *j followed by a syllable-final consonant, for instance, the genitive plural of *mənu* 'egg' would be *mənaq* (< PS *mênå-j-t).

Vowel harmony in Nganasan, while representing direct continuation from Proto-Samoyed and Proto-Uralic, no longer has any phonetic rationale because of the many vowel changes. The most frequent expression of PS non-initial *å : *ä (< PU *a : *ä) in Nganasan is *u* : *į*, as implied above, e. g. 3sg forms *bįnį-δį* : *bįnį* 'strap' (< PS *wenä-tä : *wenä [SW 175 *winä]) vs. *mənu-δu* : *mənu* 'egg' (< PS *mênå-tå : *mênå [SW 86]), but secondary changes have created allomorphs with *i* and *ü* as well (cf. Helimski 1998a: 490). An apparently related yet (at least to me) obscure alternation between a high and a low vowel appears in the basic verbal form of the gerund, producing no less than six allomorphs, of which those with *a* generally attach to back vocalic stems, e. g. *kotu-d'a* 'kill', *d'eru-sa* 'not know', *bənü-śa* 'lie (of an animal)', while *i* or *į* is found in front vocalic stems, e. g. *ńilį-d'i* 'live', *d'ud'i-sį* 'dream', *hid'i-śi* 'laugh, smile'. As for the background of vowel harmony, see section 4.

Nganasan has applied vowel harmony to a number of words that appeared disharmonic in Proto-Samoyed, as in Ng *bakunu* 'sturgeon' which would suggest PS **wękånå while the established PS reconstruction would be *wekånå [SW 174] > TN *yexana* ~ TE *bexana* (cf. Gusev 2022).

Besides the word for 'shaman' discussed above and several other basic words of unknown origin, Nganasan differs lexically from the other Samoyed languages in various ways. For instance, Nganasan has no reflex of (PU *koki- >) PS *ko- 'find' but the concept is covered by the polysemous verb *ŋətə-* 'see, find' which is cognate to Sk (№ 66) **atə-** 'see' < PS *ətə- < PU *itä- 'sprout'. Similarly, Ng *koq* (: 3sg *kuoδu* pointing to an original vowel sequence) 'pine, spruce' corresponds to two distinct tree names elsewhere in Samoyed, cf. Sk (№ 1435) **ćüə** 'pine' and (№ 1936) **qūt** 'spruce', presumably because the harsh climate allows few conifers to grow in the Nganasan homeland.

4.4.2 Enets

The Enets languages, Tundra Enets and Forest Enets, nowadays only have some tens of speakers altogether, as opposed to the Nenets languages with several thousands, but at the time of Russian contact these figures were much closer to each other, in the range of a couple of thousand on both sides, demonstrating the fact that language shifts and consequent demographic changes in the region have been rapid and far-reaching. The original territories of both Tundra Enets and Forest Enets were, however, located much further to the south of their current locations.

In comparison with the Nenets languages, whose phonemic systems differ noticeably, the Enets languages by and large share the same consonant and vowel phonemes, which has led to an interpretation that Tundra Enets and Forest Enets should

be regarded as being inherently closer to each other than Tundra Nenets and Forest Nenets. It would, however, seem that the Enets languages have diverged through a considerable number of innovations, in particular sound changes, that is comparable to those evident in the case of the Nenets languages. The difference is that many Nenets sound changes, notably the Tundra Nenets consonant shift and the restructuring of the Forest Nenets vowel system, have such a high frequency that a phonological difference can almost always be seen in cognates between the Nenets languages, whereas elements shared by both Enets languages, especially in their early records by Castrén, often appear indistinguishable. For instance, words such as TE *kiδe* ~ FE *kiδe* (later *kiδi*, see below) ~ TN *xidya* ~ FN *kĭtya* < PS *kị̈ta 'cup', TE *sopo-* ~ FE *sopo-* (later *sopu-*) ~ TN *səpə-* ~ FN *xăqpă-* < Pre-Enets-Nenets *sə̂t^{3}pə̂- 'chop off' show that the forms in question are, or at least have recently been, identical within Enets but with noticeable differences between Tundra Nenets and Forest Nenets. The correspondences between the Nenets languages, however, are regular and transparent, and reconstructing Proto-Nenets is a relatively straightforward matter, as in the above-mentioned words PN *kitya and PN *səqpə- respectively, while there are differences between the Enets languages that seem more drastic that anything found in Nenets and for the time making it impossible to come up with a decent reconstruction of Proto-Enets, despite the recurrent idea that the Enets languages are closer to each other. It is also to be noted that Tundra Enets and Forest Enets must have influenced each other because of their historical and geographic setting (cf. Khanina & Shluinsky 2022), while the Nenets languages have barely had contacts since the break-up of Proto-Nenets.

In what follows, Enets materials are mostly quoted after Helimski's extensive Tundra Enets manuscript, although Forest Enets forms are obviously included when necessary. This may appear surprising, because many more published sources are available on Forest Enets than on Tundra Enets. Helimski's corpus is, however, highly consistent and therefore of great practical value, not least because he has also added cognates from Forest Enets and Tundra Nenets, and Tundra Enets in itself can be seen as the more conservative of the Enets languages which serves a purpose in a comparative analysis. Furthermore, Forest Enets sound system has changed markedly since Castrén's records, to the extent that one can talk about early Forest Enets and late Forest Enets, a distinction that was already implied in a couple of examples above. Late Forest Enets records also tend to be rather varied because of varying competence among speakers and disparate notations among linguists.

The Enets consonant system consists of four nasals *m n ń ŋ*, five voiceless stops *p t t' k q* (= glottal stop), two sibilants *s ś*, three liquids *l l' r*, currently three approximants *δ j x* (= *γ*), and four voiced stops *b d d' g*, making the system much richer than what is reconstructed for Proto-Samoyed. A fourth approximant *w* seems to have been present until recently but already largely merged with *b* in both languages by the time of Helimski's fieldwork. A major phonotactic restriction involves the distinctly palatal consonants *ń t' ś l'* so that they only precede (original) high vowels (cf. PS *ńęrkå > TE *niga* 'willow'), which has led Alatalo to suggest an alternative phonemic analysis where adopting back-

front contrast among high vowels makes palatality of consonants an allophonic feature, for instance, Helimski's TE *tira* 'fist' and *t'iree* 'roe' or *tudo-* 'recognize' and *t'uro* 'staff' would be rendered as *tʲra* 'fist' vs. *tiree* and *tudo-* vs. *türo* in the alternative system. No phonological contrasts are compromised either way, and I follow Helimski for the sake of consistency. It must also be noted that at least Forest Enets allows palatal consonants to combine with non-high vowels in recent loanwords.

The main development behind the Enets consonant system is the simplification of all Proto-Samoyed consonant clusters into a single consonant in Enets, whereby intervocalic voiceless stops derive from PS clusters of two obstruents (e. g. PS *mâktut > TE *motuq* 'six') and voiced stops from PS clusters of a nasal and an obstruent (*muntâ > *mudo* 'piece'), while the original intervocalic stops have lenited into approximants (*mi̯tâ > *muδo* 'liver'). Consonant clusters are present in modern Enets but, barring recent developments in Forest Enets, they occur only infrequently in loanwords or otherwise secondary formations. Furthermore, such cases seem to have emerged independently in both languages, for instance in the case of TE *opto* 'smell', which can hardly be explained as anything but a borrowing from a neighbouring Samoyed language, while FE *oto* id. comes regularly from PS *əptə (> Proto-Nenets *ŋəptə).

A phonotactic consequence of the simplification process is that the approximants *δ j x* (and the pre-modern *w*) do not appear word-initially, while initial *d g* occur rarely as a special case of denasalization, cf. TE *gudoδi* 'length (of)', i. e. a word that invariably follows a genitive attribute ending in PS *-n. Initial *b d'*, by contrast, are common because of a fortition process *w > *b* (*wẹn > *buq* 'dog') and *j > *d'* (*jür > *d'uq* 'hundred'), shared by both Nganasan and, as it seems, Yurats.

While most Proto-Samoyed consonant clusters have the same reflexes in both Enets languages, there are two major differences between Tundra Enets and Forest Enets, hampering the reconstruction of Proto-Enets unless one assumes either that it still preserved (at least some) consonant clusters, or that it had two additional consonant phonemes to explain the discrepancies. First, a well-known correspondence between TE *d'* ~ FE *s*, already discussed above, derives from PS sequences of a nasal and a sibilant (whether original or < *k before front vowels), e. g. PS *märkä [SW 93 *merkä] > TE *med'e* ~ FE *mese* (later *mäsi*) 'wind'. Second, there is a rare but distinct correspondence TE *d* ~ FE *r*, deriving, notably, from PS *lt < *lč, e. g. PS *səlčə [SW 132 *sə̑lcə̑] > TE *sodeqei* ~ FE *soreqei* (originally augmentative derivatives, now lexicalized as) 'stump'.

Another striking innovation in Enets is that all Proto-Samoyed final consonants have merged into a glottal stop. This change is uniform across Enets, e. g. *buq* 'dog' < *wẹn [SW 173], *koq* 'pus' < *kət^{1} [SW 55 *kə̑t^{1}], *ńiq* 'shaft' < *nir [SW 102], or *ńiq* 'name' < *nim [SW 102 *nim ~ *nüm]. As shown by the respective plural forms *bunoq* < *wẹn-ə̑-t, *koδoq* < *kət^{1}-ə-t, *ńiroq* < *nir-ə-t, and *ńioq* < *nim-ə-t, the original consonant can usually be retrieved via internal reconstruction. Consequently, since canonical Enets wordforms, still prevalent in Tundra Enets and at least in early Forest Enets, do not contain consonant clusters, the only regularly occurring closed syllables are those final ones ending in a glottal stop.

Other major innovations involving consonants include the merger of intervocalic PS *l > *r*, e. g. PS *kålä > TE *kare* 'fish' (*l* being restored from original consonant clusters, e. g. PS *sât^{3}lå > TE *sola* 'idiot'), the loss of intervocalic *m and *ŋ, e. g. *mę̑må > *mua* 'action' or (via the vocalization of *j) *påjŋå [SW 115] > *pae* 'mattress', the merger of *w with *m after a sonorant, for instance, *rw > *m, e. g. *kârwå- > *koma-* 'want, wish' or (showing a cumulative effect of the loss of *m) *kajwå [SW 63 *käjwå] > TE *sea* ~ FE *see* 'spade, shovel', and the loss of *m and *w after a glottal stop, e. g. *tâlut^{3}må > *toruqa* 'covering' or *säjt^{3}wə [SW 139 *sejt^{3}wâ (? *sejkwâ)] > *seqo* 'seven'. These changes have taken place uniformly in both Tundra Enets and Forest Enets.

The Tundra Enets, and apparently also common Enets, vowel system seems rather untypical, as it contains three labial vowels, in Helimski's transcription followed here, *u ô o*, the circumflex marking a vowel intermediate between *u* and *o*, as opposed to the more usual triplet of illabial vowels *a e i*, although the small number of vowel phonemes in general is notable, given the 13 vowel phonemes *ə *ə *i *e *ä *a *å *o *u *į *ę *ö *ü reconstructed for Proto-Samoyed, which is to be explained as follows. The PS reduced vowels *ə *ə have become labialized in Enets with *o* as their reflex, pushing PS *o higher > *ô*, which also reflects *ö, e. g. PS *t^{1}ôpə > (*towo* >) *tobo* 'sand', *t^{1}oså- > *tôsa-* 'beware of', *töntä > *tôde* 'fox'. The PS illabial and labial high vowels have merged, respectively, i. e. PS *i / *e / *į > *i* and PS *u / *ü > *u*, potentially triggering the emergence of the Enets palatal consonants, e. g. (cf. above) *tį̑rå > *tira* 'fist', *tirämä > *t'iree* 'roe', *tumtə- > *tudo-* 'recognize', *čürə > (*türə >) *t'uro* 'staff'. Two full mergers finally reduce the system to six units, viz. PS *ä / *a > *e*, e. g. PS *mär-ə-n > *meroq* 'soon' (~ TN *myer°h*), PS *mat-ə-t > *meδoq* 'tents' (~ TN *myad°q*), and PS *ę > (*į >) *i*, e. g. PS *kęm > *kiq* 'blood'. Further conditional vowel changes have also taken place, as seen in the reflex *u* < PS *ę / *į when adjacent to a labial consonant, e. g. *mį̑tə > *muδo* 'liver'.

While the numerous changes in the vowel system mentioned above are shared by both Enets languages, the treatment of initial vowels shows marked differences, notably PS *a, cf. *aŋ [SW 20 *äŋ] > TE *eq* ~ FE *naq* 'mouth', and PS *i / *e / *ü, cf. *elä- [SW 27 *ilä- (ng, sk, tg) ~ *jilä-] > TE *ire-* ~ FE *d'ire-* 'live', (derivatives of) *üčä [SW 31] > TE *uδea* ~ FE *d'uδee* 'small'. The same goes with PS *Vj sequences, e. g. PS *åj 'foot, leg' > TE *ŋa* ~ FE *ŋo*, or PS *kåjkə [SW 51 *kə̑jkə̑] 'idol' > TE *kaxa* ~ FE *kixu* («шайтан»), as well as some instances of PS vowel sequences, for instance, *pua-ntə̑-ŋ > TE *puadoq* ~ FE *puedo*(*q*) 'for a long time'.

The retention of the original vowel sequences and the loss of intervocalic consonants have created three-vowel sequences, best preserved in Tundra Enets, cf. PS *kåə̑-må-tå > TE *kaaaδa* 'her/his death' vs. PS *kåjmå-tå > TE *kaaδa* 'her/his bone marrow' vs. PS *kåtå 'grandmother' > TE *kaδa* 'aunt'.

Non-initial syllables in Enets lack the "intermediate" round vowel *ô*, so that the result is a simple *u o a e i* system.

Among Enets lexical innovations, the word for 'tail' may be highlighted. All the other Samoyed language show reflexes of PS *tə̑jwå [SW 150 *t^{1}åjwå], but in Enets we

have *batuqo* ‘tail’ (with a Nganasan cognate meaning ‘coccyx’). As a potential explanation for the loss of the PS word in Enets, the erosion of the consonant cluster *jw may be offered, cf. the discussion on *kajwå ‘spade, shovel’ above.

4.4.2.1 Tundra Enets

The Tundra Enets language (TE *somatu nau*; also known as Madu after FE *madu* ‘Tundra Enets’) is currently spoken by very few people concentrated in the village of Vorontsovo (Воронцóво) on the Taimyr Peninsula, but its original territory was not only much larger but also more southerly, actually covering much of what later became Forest Enets territory. Few publications have been devoted to Tundra Enets, Labanauskas (2002) being a notable exception, while Sorokina & Bolina (2005) contains a handful of texts in Tundra Enets.

In addition to the features of Tundra Enets, either same as or different from those of Forest Enets mentioned above, only a few need to be listed here, because in most regards Tundra Enets has preserved a more archaic stage. Tundra Enets has, however, lost *j in most intervocalic contexts, like Nganasan as it happens but unlike Forest Enets or Nenets for that matter, cf. PS *sojå- [SW 142] ‘be born’ > TE *sôa-* (~ FE *sôja-* ~ TN *soya-*). Tundra Enets also exhibits a peculiar vowel reduction process, as shown by the second-syllable vowel in PS *nänäŋkə [SW 23 *nenäŋkâ] ‘mosquito’ > TE *nenogo* (~ FE *nenago*, later *nänag* ~ TN *nyenyaŋk°*; for Nganasan, curiously, variants corresponding to both TE *nenogo* and FE *nenago* seem to have been attested, viz. Ng *ńenəŋkə ~ ńeŋiŋkə*, respectively). Furthermore, Proto-Samoyed *k is labialized in specific contexts, e. g. PS *čuk [SW 34 *cuk (~ *cukâ)] ‘maggot’ > TE *tubo* (~ FE *tuxu* ~ TN *tux°*).

4.4.2.2 Forest Enets

The Forest Enets language (FE *onei baδa* or *bai baδa*; also known as Bai) is spoken by a very small number of people in and around the village of Potapovo (Пота́пово) on the Taimyr Peninsula, while the original territory of Forest Enets was located much further to the south, basically in the region where Taz Selkup in now spoken, enabling contacts with the Yeniseian language Ket, from which two of the personal pronouns of Forest Enets have been borrowed, creating another rather fundamental difference between the Enets languages. Two descriptive monographs have appeared on Forest Enets, a modern treatise focusing on syntax by Siegl (2013), and a more traditional account by Sorokina (2010), who together with the native linguist Bolina published two dictionaries and a text collection (Sorokina & Bolina 2001, 2005, 2009).

When it comes to early Forest Enets, mostly known through Castrén’s records, but, interestingly, also recorded by Helimski still in the 1970s from speakers of Tundra Enets who mastered Forest Enets as a second language, little needs to be added to what has

been covered above. What is generally attested in late Forest Enets, starting from the middle of the 20th century, shows marked innovations in the vowel system. The "intermediate" round vowel *ô* has merged with *o*, at least in the corpus compiled by Sorokina and Bolina with the former, but presumably with *u* in some idiolects, which has caused confusion when materials from several speakers have been studied and published indiscriminately. At the same time, the vowel corresponding to TE *e* appears as a low front vowel *ä*, while a "new" *e* has developed from an early Forest Enets vowel sequence *ue*, cf. (early) *perene* > (late) *pärni* 'kin' vs. *puena* > *pena* 'afterwards'. Furthermore, it seems that another vowel sequence has not only resulted as *e* contrasting with *ä*, but it has also caused secondary palatalization, cf. (early) *kie* > (late) *k'e* 'bone marrow'. As seen from the form *pärni* above, non-initial vowels have been subject to (at least phonetic) loss as well as other modifications, and the cumulative effects of these recent innovations may be considerable, e. g. (early) *mese* > (late) *mäsi* 'wind'.

4.4.3 Yurats

The Yurats language (Yr; the name 'Yurats' is an artificial variant of Yurak, a former exonym for Nenets) was presumably spoken between the lower reaches of Taz and Yenisei rivers in the south of the Gydan Peninsula. Yurats is documented through a single word list collected around 1740 that contains 250 translations of Latin words (Helimski 2000 [1976]: 50–55). The few inflected or derived forms in the list suggest, not surprisingly, shared morphological features with Enets and Nenets, and the material is substantial enough to show that phonologically and lexically Yurats was very much intermediate between these two sub-branches of Samoyed.

The Yurats vowel system appears quite archaic, as it probably retained the prosodic contrast between a reduced vowel *ə* (with notable allophonic variation) vs. full vowels like Tundra Nenets, cf. PS *lâptå > Yr ‹lópta› [≈ *ləbta*] ~ TE *lota* ~ TN *ləbta* 'plain', PS *mârå > Yr ‹márra› [≈ *məra*] ~ TE *mora* ~ TN *məra* 'sand', or Pre-Enets-Nenets *jâpto > Yr ‹jeptu› [≈ *jəbtu*] ~ TE *d'otu* ~ TN *yəbto* 'goose'. Yurats manifestly preserved PS vowel sequences like Nganasan, Enets, and Mator but unlike Nenets where they were replaced by single vowels, e. g. PS *tua > Yr ‹túa› ~ TE *tua* ~ TN *to* 'wing', PS *wüət > Yr ‹büu› [≈ *büuq*] ~ TE *biuq* ~ TN *yúq* 'ten'. Of the non-initial high vowels, *u seems to have merged with the reduced vowel, e. g. PS *mâktut > Yr ‹mat› or ‹mátta› [≈ *mətəq*] ~ TE *motuq* ~ TN *mət°q* 'six', while *i has been preserved, e. g. Pre-Enets-Nenets *mənti > Yr ‹modi› ~ TE *mod'i* ~ TN *məny°* 'I'.

The last example above reveals one of the most conspicuous sound changes linking Yurats more closely to Enets, i. e. the denasalization of PS nasal-stop clusters, e. g. Pre-Enets-Nenets *juntå > Yr ‹júda› ~ TE *d'uda* ~ TN *yuna* 'horse', PS *lempä > Yr ‹líbbe› ~ TE *l'ibe* ~ TN *lyimpya* 'eagle', PS *mâråŋkå > Yr ‹marága› ~ TE *moraga* ~ TN *məraŋka* 'cloudberry'. Other consonant clusters, however, are largely preserved, unlike in Enets. A specifically Yurats sound change *m > *b* is found in syllable-final positions, e. g. PS *um > Yr ‹ngub›

~ TE (analogical) *ŋuuo* ~ TN *ŋum* ‘grass’, PS *âmså > Yr ‹ngábsa› ~ TE *ud'a* ~ TN *ŋəmca* ‘meat’. Intervocalically, *p > *w*, *k > *x* (or rather *γ*), and *t > *δ*, e. g. PS *päpä > Yr ‹péwe› ~ TE *pebe* ~ TN *pyebya* ‘younger brother’.

4.4.4 Nenets

The Nenets languages, Tundra Nenets and Forest Nenets, derive from Proto-Samoyed through an intermediate proto-language, Proto-Nenets, which, unlike the prospective Proto-Enets, can be reconstructed in a rather straightforward manner. Employing Proto-Nenets extensively in a comparative context is also rather useful, as the sound changes that separate Tundra Nenets and Forest Nenets are numerous and frequent, yet highly transparent. Lehtisalo (1956) is a major dictionary containing material from both languages.

The Proto-Nenets consonant system is best described for non-palatalized and palatalized consonants separately, as the latter group represents an innovation known as translocation of palatality, whereby all (originally non-palatal) consonants are palatalized before all PS front vowels as well as before PS *a [SW *ä], e. g. PS *lempä [SW 83 *limpä] > PN *lyimpya ‘eagle’ or PS *mat [SW 90 *mät] > PN *myaq ‘tent’. Proto-Nenets non-palatalized consonants included three nasals *m *n *ŋ (PS *ń having been merged with the palatalized series), four stops *p *t *k *q (= glottal stop, which still in Proto-Nenets may have represented a non-contrastive neutralization of *t *k *s), one sibilant *s, two liquids *l *r (the rare PS *l' merging with *j), and two glides *w *j (the latter obviously an inherently palatal consonant). Proto-Nenets palatalized consonants were fewer, i. e. *my *ny *py *ty *sy *ly *ry *wy, because the PS velars *ŋ (including the prothetic *ŋ in original vowel-initial forms) changed to *ny and *k to *sy before front vowels.

Translocation of palatality combined with nasal prothesis discussed below has caused massive neutralizations in Nenets with regard to the PS initial sequences *ńa, *na, *ńå, and *a which have all resulted in PN *nya in their stead, cf. PS *ńa [SW 106 *ńä] > (Ng *ńa* ~) TN *nya* ‘friend’, PS *nakur [SW 99 *näkər] > (Ng *nagür* ~) TN *nyax°r* ‘three’, PS *ńåråmå > (Ng *ńorumu* ~) TN *nyarawa* ‘copper’, PS *aŋ [SW 20 *äŋ] > (Ng *ŋaŋ* ~) TN *nyah* ‘mouth’.

Another major innovation involving Proto-Nenets consonants was the progressive assimilation of dental sonorant-stop clusters, viz. PS *nt / *nč > PN *nn (> TN *n*), e. g. *kânčâ > *kənnə > *xən°* ‘sledge’, and PS *lt / *lč > PN *ll (> TN *l*), e. g. *səlčə > *səllə > *səl°* ‘stump’, with distinct reflexes of the PN geminates preserved in Forest Nenets. A new *nt cluster emerged from PS *ŋt, e. g. PN *nyantə ‘blade’ < PS *aŋtâ. Many Nenets *nt* clusters owe their existence to analogical developments, e. g. TN (*weh* ‘dog’ :) dat.sg *went°h* ‘to the dog’ instead of the presumably regular **wen°h* < PS *wẹn-tâ-ŋ, cf. TN (*ser* ‘ice, salt, white’ :) dat.sg *sert°h* < PS *sẹr-tâ-ŋ.

Proto-Nenets preserved the PS system of reduced vs. full vowels, but since the contrast between back and front vowels was neutralized because of the translocation of

palatality, both PS *ə̂ and *ə > PN *ə, but with a contrast in non-initial sequences *ə̂j (> TN *o*) vs. *əj (> TN *ye*) as explained above. The concomitant changes of the full back vowels besides PS *a are straightforward, i. e. PS *į > PN *i, *ę > *e, *å > *a, *o > *o, and *u > *u, as is the case with front rounded vowels, i. e. PS *ü > PN *yu and PS *ö > PN *yo. Because of the merger of PS *i and *e outside Nganasan, they both > PN *yi, which was followed by a chain change PS *ä > PN *ye and *a > (*ä >) *ya. In the final position, typically in the nominative singular of nouns, PS *ä > PN *ya instead, cf. *kålä > *kalya 'fish' as opposed to the regular development in verbal stems, e. g. *elä- > *jilye- 'live' because a verbal stem never occurs alone.

Besides the six PN vowels *ə *a *e *i *o *u resulting from PS single vowels as described above, Proto-Nenets acquired three new vowel phonemes *æ *í *ú from PS *Vj sequences, while *í *ú also reflect a number of PS vowel sequences. The reflex of the vowel *æ in modern Nenets languages is typically diphthongoid [ǎè], and the vowels *í and *ú simply differ from *i and *u in duration.

As for vowels in non-initial syllables, Proto-Nenets merged the Proto-Samoyed high vowels with *ə, and since a parallel development took place in Taz Selkup, another key language employed by Janhunen for the compilation of SW, high vowels are not included in its non-initial vowel inventory. Proto-Nenets acquired secondary non-initial high vowels from PS *Vj and vowel sequences.

Both Nenets languages possess a contrastive semi-suprasegmental (or in the case of Forest Nenets rather fully suprasegmental) feature tagged as "schwa" and marked with °, cf. FN *kămta-* 'pour' vs. *kăm°ta-* 'fell', where the main phonetic exponent of the contrast is the lengthening of the preceding consonant. It must be noted that ° acts as a nucleus of a syllable just like any (other) vowel despite its prosodic peculiarity, so that *kăm°ta-* 'fell' represents a trisyllabic verb, which is important because the number of syllables plays a crucial role both in Nenets phonology and morphology. When it comes to Proto-Nenets, however, it is still feasible to regard the "schwa" as an allophone of PN *ə, from which it is derived in unstressed positions. Both Tundra Nenets and Forest Nenets have developed exceptions to the predictable distribution of "schwa", yet in a different manner, manifesting its separate phonemization in the Nenets languages.

Both Nenets languages also exhibit secondary vowel sequences, which emerged through contraction after the original vowel sequences had become single vowel phonemes in Proto-Nenets. Most of the vowel sequences have developed separately in the Nenets languages, but a few shared examples exist such as PS *aŋə̂rå > Ng *ńaŋəru* ~ TE *eara* ~ FE *naara* ~ (PN *nyaəra >) TN *nya°ra* ~ FN *nya°ła* 'inner side of hide'. Similarly, substems with an epenthetic vowel of PS *ŋ stems regularly appear contracted, e. g. (*aŋ [SW 20 *äŋ] 'mouth' :) gen.sg *aŋ-ə̂-ŋ > TN *nya°h* ~ FN *nya°ŋ* 'of the mouth'.

Interestingly, the so-called aorist marker *-ŋå has been subject to contraction in both languages but under different conditions. In Tundra Nenets, *ŋ is only preserved when a suffix with an initial *x* follows, e. g. PN *jilyeŋakəŋ > TN *yilyeŋax°h* 'they [du] live' vs. *jilyeŋa > *yilye°* '(s)he lives' or *jilyeŋannə > *yilyeən°* 'you [sg] live' and *jilyeŋaq >

yilye°q ‘they [pl] live’. In Forest Nenets, by contrast, *ŋ is otherwise preserved except that it may be optionally lost when the marker *-ŋå occurs in the final syllable, which is the case in 3sg and 3pl forms, i. e. PN *jilyeŋa > FN *jĭłyiŋa ~ jĭłye°* ‘(s)he lives’ and *jilyeŋaq > FN *jĭłyiŋaq ~ jĭłye°q* ‘they [pl] live’, but always *jilyeŋannə > *jĭłyiŋan°* ‘you [sg] live’ etc. Another frequent source of vowel sequences is the unusual loss of *sy in the gerund marker (for a hypothetical explanation see the discussion on consonant gradation in section 6), e. g. PN *jilyesyə > TN *yilye°* ‘to live’, making it homonymous with the form ‘(s)he lives’ quoted above. Crucially, Nenets vowel sequences, irrespective of their secondary origin, are invariably contrastive (cf. Jalava 2015), compare, for instance, TN *yilye°q* ‘they [pl] live’ with the imperative form *yilyeq* ‘live!’.

Proto-Nenets went through numerous lexical changes, of which I mention a few. PS *mânå ‘egg’ (< PU *muna) was replaced, apparently after the word had acquired taboo connotations, with PN *saryəqnyu > TN *sar°qnyu* ~ FN *xały°qnyu*, which looks like a lexicalized compound yet with no readily identifiable segments. Nenets has no reflex of PS *kaså [SW 65 *käså] ‘tree bark’ but uses PN *syaptyə > TN *syabt°* ~ FN *syapty°*, deriving from PS *kapt[1]i ‘bark, scale’, cf. Mt (№ 446) ***käbti*** ‘husk’ (Helimski 1992–1993: 112). PS *tômå ‘mouse’ has also been lost in Nenets, but in this case the Nenets languages have different words, TN (presumably onomatopoeic) *pyísya* and FN (borrowed from Eastern Khanty) *jăŋkăł* ‘mouse’. Both Nenets languages have also replaced PS *kåjå [SW 58] ‘sun’ but, again, in a different way, so that TN *xayer°q* (< PN *kajarəq with no reflex in FN) represents a deverbal noun of *xayera-* (< PN *kajara- > FN *kajała-*) ‘clear up’, while Forest Nenets uses *jałya* ‘light, day, sun’, cf. Selkup below.

4.4.4.1 Tundra Nenets

The Tundra Nenets language (TN *nyeney° wada* in Siberian dialects and *nyenecyə°h wada* in European dialects) is spoken across a vast Arctic territory extending from the Kanin Peninsula in the west to the Yenisei Delta in the east by some 20,000 people.

Tundra Nenets is divided into three major dialect zones, viz. the western dialects to the west of the Pechora River, the central dialects between the Pechora River and the Ural Mountains, and the eastern dialects to the east of the Ural Mountains. The western and central dialects together can therefore be called the European dialects of Tundra Nenets, while the eastern dialects are synonymous with the Siberian dialects of the language. The western dialects are phonologically distinct, with innovations including denasalization (e. g. *lyimpya* > *lyibya* ‘eagle’) and retentions such as intervocalic *wy (> *by* in other dialects, e. g. *nyewya* pro *nyebya* ‘mother’). The major morphological and lexical isoglosses run along or close to the Urals, hence it is possible to talk about specifically European vs. Siberian features of Tundra Nenets, in lexicon often involving different loanwords from Russian and Northern Khanty, respectively, e. g. European *ŋow°cya* (< Russian *ovcá*) vs. Siberian *as° ~ ŋas°* (< Northern Khanty) ‘sheep’. In morphology, the most conspicuous difference is found in the subjective 1sg suffix, which in Siberian dia-

lects, quoted later, is *-d°m* < PN *-təm as opposed to European *-m°h* which presumably reflects an earlier suffix complex *-m-məŋ.

Tundra Nenets is rather well-studied for a minority language, with a modern grammar focusing on syntax (Nikolaeva 2014), and an excellent bilingual dictionary (Tereščenko 1965) as well as a slightly experimental morphological dictionary (Salminen 1998a).

The most conspicuous innovation of Tundra Nenets must be the so-called great consonant shift, with postvocalic weakening of PN *p > *b* and *t > *d*, postconsonant strengthening of PN *s > *c*, and the change PN *k > *x* both in intervocalic and initial positions. New intervocalic *p t k* emerged from PN *qC sequences, of which *qs yielded *c* thus giving affricates a contrastive status. These diachronic changes are still present in modern Tundra Nenets in the form of a complex sandhi system (Salminen 1997) because they are effective across word boundaries, e. g. *ŋarka to* becomes *ŋarka_do* 'big lake' via sandhi.

Other notable changes involving consonants are initial *wy > *y*, intervocalic *m > *w*, final *ŋ (itself representing all PS non-labial nasals) > *h*, a symbol for the so-called nasalizable glottal stop, phonetically identical with the non-nasalizable glottal stop, but with a distinct sandhi pattern, e. g. *nyeh xən°* becomes *nyeŋ_kən°* 'a woman's sledge', as opposed to the non-nasalizable glottal stop, whereby *nyeq xən°q* becomes *nye_kən°q* 'women's sledges'. Furthermore, Proto-Nenets palatalized consonants have largely become depalatalized before *ə, e. g. PS *jänti [SW 43 *jentâ] > PN *jennyə > TN *yen°* (~ FN *jeny°*) 'bowstring', except, curiously, when the first syllable vowel is *ə, e. g. PS *ətiw- > PN *ŋətyə- > TN *ŋədyə-* 'be visible' (for details, see Salminen 2012).

The Tundra Nenets vowel system, by contrast, may be considered conservative at least with respect to the initial syllable, where the only notable change on a systemic level would be the merger of PN *yæ > *yí* (i. e. a reflex of a PS front vowel plus a *Vj or vowel sequence). In non-initial syllables, PN sequences *əmə > (*əwə >) TN *uə* and *əjə > TN *iə*, e. g. PN *juqtəmə > TN *yutu°* 'stockfish', PN *patəjəta > TN *padi°da* '(s)he wrote them [pl]', a contraction interpreted as much older by Janhunen (1977a) as seen below. There are many more specifically Tundra Nenets contractions, e. g. PN *pəjer- > *pœər-* 'do', with further cases in dialects such as PN *ŋəmor- > (*ŋəwor- >*) *ŋoər-* 'be eating'.

Several Proto-Uralic and Proto-Samoyed words have been lost in Tundra Nenets alone, notable cases being TN *nya°myu* 'tongue', a lexicalized compound of *nyah* 'mouth' and *myu* 'insides', which has replaced (FN *syĕ* <) PN *sye < PS *käəj [SW 66 *keə̑j] < PU *käɣli, and TN *pəncyeq°* 'louse', with cognates referring to a particular large louse species in Forest Nenets and elsewhere, which has replaced (FN *ŋan°* <) PN *ŋannə < PS *åncu [SW 18 *åncə̑ (? ~ *åmcə̑)]. Another example of a Proto-Samoyed word without a reflex in Tundra Nenets would be *saču 'firebrand' (Helimski 1986: 120 *säcu; cf. TN *yataqma* id. derived from *yaq* 'coal'), which is curiously not found in SW despite its otherwise wide distribution, indicating the central role of Tundra Nenets in Samoyed comparative studies not only with regard to *Samojedischer Wortschatz* but also its predecessors, notably Lehtisalo (1927) and Collinder (1955).

4.4.4.2 Forest Nenets

The Forest Nenets language (FN *nye°qsya°ŋ wata*; a neologism "Nyeshang" has recently been coined on the basis of FN *nye°qsyaŋ* 'Forest Nenets; human being') is spoken across the watershed between the basin of the Pur River, draining to the Arctic Ocean, and several northern tributaries of the Ob' River by approximately 1,000 people, although the number of speakers is dwindling. Verbov (1973) and Sammallahti (1974) comprise brief but insightful descriptions of Forest Nenets phonology and morphology, while Koškareva (2005) constitutes a description of Forest Nenets syntax with a comprehensive bibliography.

In contrast with Tundra Nenets, Forest Nenets may be characterized as relatively archaic in its consonant system, while its vowel system has been restructured after the model of its most important contact language, Eastern Khanty.

As for non-palatalized consonants, the system of the most conservative dialect of Forest Nenets is nearly identical with Proto-Nenets, consisting of three nasals *m n ŋ*, four stops *p t k q* (= glottal stop), one sibilant *s*, two liquids *l r*, and two glides *w j*, plus a fricative *x* as the only addition to the phonemic inventory. Starting as an intervocalic allophone of *k in Proto-Nenets, *x* was phonemized in both Nenets languages after they acquired a secondary intervocalic *k*, but the processes were obviously different, as we would otherwise need to include *x already in PN inventory, cf. PN *qk > TN *k* as part of the great consonant shift, while in Forest Nenets (the rare) intervocalic *k* < *ŋk, e. g. *jiku* < *jaŋku 'there is not'. Because of the small number of consonant changes, Forest Nenets has much less morphophonology than Tundra Nenets, but the phonemization of *x* has resulted in a similar alternation found there, cf. loc.sg *măntal-kăna* 'in the herd' vs. *ŋŭta-xăna* 'in the hand'.

Forest Nenets retained all of Proto-Nenets palatalized consonants, i. e. *my ny py ty sy ly ry wy*, but innovated a series of palatalized velars *ŋy ky xy*, which, curiously, only contrast with *ŋ k x* before *i* (or secondarily shortened *ĭ*). FN *ŋyi kyi xyi* represent the original PN *ŋí *kí *xí sequences, while FN *ŋi ki xi* result from metaphony (see below), cf. *kímya > *kyimya* 'who' but *kasujə > *kisuj*° 'dry'.

The structural similarities between the Forest Nenets and Proto-Nenets consonant systems do not, however, tell the whole truth, especially when it comes to liquids, as Forest Nenets experienced a full merger of PN *l *ly > *r ry* (cf. Enets where PS *l > *r* only intervocalically), e. g. PN *ləmpa > FN *rămpa* 'ski', *kalya > *karya* 'fish'. This total loss of laterals was then reversed in specific contexts resulting in "relateralization" of some instances of both PN *l and *r, e. g. PN *luqkə > *lŭqk*° 'capercaillie', PN *seraqko > *xelaqku* 'white'. What is more, early Forest Nenets *r ry* have by now become lateral fricatives, or "fricolaterals", for which *ł ły* are the conventional symbols employed here, across the entire Forest Nenets language area. As a result, *r* is either absent or highly marginal depending on the particular dialect, for instance, a recent borrowing (originally from Russian *zdoróvo*) from TN *torowa* 'hello' may retain the consonant (*torawa*) or undergo sound substitution (*toława*), both variants being currently attested.

Other notable changes in the distribution of consonants include the PN initial *s > FN *x* (phonetically [h]), the dissimilation in geminate liquids *ll *lly > *nr nry* > *nł nły*, e. g. *səllə > *xănr°* > *xănł°* 'stump', the semivocalization of intervocalic *ny > *j* whereby *nny > *ny*, e. g. *tyenye- > *tyeje-* (~ TN *tyenye-*) 'remember' vs. *myinnye- > *myĭnye-* (~ TN *myinye-*) 'transport', and the (not yet fully implemented in all dialects) loss of syllable-final consonants after the "schwa", e. g. PN *toŋatəm > (nowadays rare) *toŋat°m* ~ (more common) *toŋat°* 'I came'.

While Tundra Nenets vowels have changed little from Proto-Nenets, the Forest Nenets vowel system looks markedly different, in its most archaic form consisting of six long (or full or tense) vowels *i e ä a o u* and four short (or reduced or lax) vowels *ĭ ă̆ ă ŭ*, and while the system of short vowels has been augmented by *ĕ ŏ* in contemporary Forest Nenets, making the number of long and short vowels equal, these late additions remain infrequent. This system is virtually identical with the inventory of cardinal vowels in Eastern Khanty, and since Eastern Khanty is also the donor of nearly all loanwords attested in Forest Nenets, there is little doubt that Eastern Khanty has influenced the vowel system as well, not only with regard to this restructuring of the general system as it happens, but also as a model for the reduction of vowel contrasts in unstressed syllables and the emergence of metaphony, discussed immediately below (see also Salminen 2007). The actual restructuring, the evidence for which derives from, among other things, morphophonological alternations concerning vowel length, has occurred as follows: FN *i e ä a o u* < PN *í *e *æ *a *o *ú and FN *ĭ ă̆ ă ŭ* < PN *i *æ *ə *a, respectively, *æ having experienced a phonemic split. The process is only valid for the initial syllable, cf. PN *pyiryepyita > FN *pyĭłyipyita* ~ TN *pyiryebyida* '(s)he is cooking it', with full *i* in the third syllable of the FN form.

The system of unstressed vowels is much smaller, only consisting of the "schwa" ° vs. maximally four full vowels *i ä a u*, of which *ä* has merged with *i* in most dialects. The original unstressed *e *o appear as *i u*, but can be recovered via internal reconstruction, cf. *syanaqku* 'toy' < PN *syanaqko : 3sg *syanaqkota > *syanaqkota*, the first syllable of any two-syllable foot being stressed.

Within Samoyed, metaphony, i. e. the change of stressed non-high vowels to high vowels before a following unstressed high vowel, e. g. *nyaŋu > *nyiŋu* 'chin', is a uniquely Forest Nenets phenomenon, again with a potential model in Eastern Khanty. Combined with reduction of unstressed vowel contrasts, it may result in neutralizations such as the accusative plural forms of 'boat' *ŋĭnu* (< PN *ŋənnu) and of 'bow' *ŋĭnu* (< PN *ŋinno).

The dialect differences within Forest Nenets are not numerous but they are quite conspicuous. Besides the issue with liquids, discussed above, the phoneme transcribed as *j* has three phonetically distinct realizations, viz. [j] in the modern Numto dialect as well as in the extinct Nazym and Lyamin dialects, [d'] in the Pur dialect spoken by the vast majority of the remaining speakers of Forest Nenets, and [t'] in the Agan dialect, heavily influenced by Eastern Khanty, where voiced obstruents are unheard of. More notably, the western dialects of Forest Nenets, i. e. Nazym, Lyamin, and Numto, as opposed to the eastern dialects of Pur and Agan, have undergone changes which are not only pho-

nologically innovative, but which have also had a major impact on morphology. First, denasalization of nasal-stop clusters has occurred after non-initial syllables in western Forest Nenets, e. g. *mălaka* < *mălaŋka* 'cloudberry' or *nyanatu* < *nyanantuŋ* 'with them'. Second, western Forest Nenets has generalized vowel stems in the nominative singular of nouns, e. g. *myat°* for *myăq* 'tent' or *xătät°* for *xătɨł* 'black grouse'.

As inconspicuous as the above-mentioned changes may appear on the surface, they result in levelling the system of grammatical cases in the western dialects, still present in eastern Forest Nenets. For example, the triplets of grammatical case forms consisting of the nominative, accusative, and genitive in the singular, as in PN *ŋuq 'trace' : *ŋutəm : *ŋutəŋ > FN *ŋŭq* : *ŋŭt°* : *ŋŭt°* (the last two from earlier *ŋŭt°m* : *ŋŭt°ŋ*) > western FN *ŋŭt°* : *ŋŭt°* : *ŋŭt°* (via vowel stem generalization), PN *mitə 'liver' : *mitəm : *mitəŋ > FN *mĭt°* : *mĭt°* : *mĭt°* (the last two from earlier *mĭt°m* : *mĭt°ŋ*) with no difference in western Forest Nenets, and PN *ŋuta 'hand, arm' : *ŋutam : *ŋutaŋ > FN *ŋŭta* : *ŋŭtam* : *ŋŭtaŋ* > western FN *ŋŭta* : *ŋŭta* : *ŋŭta* (via non-intial denasalization of nasal-stop clusters).

A unique grammatical innovation of Forest Nenets involves the andative 'go to do something' and venitive 'come to do something', e. g. FN *kata-soq-ŋa-m°* 'I went to kill it' and *kata-soja-ŋa-m°* 'I came to kill it', cf. *kata-ŋa-m°* 'I killed it'. While the andative in *-soq* was known from numerous examples in Lehtisalo (1956), the venitive in *-soja* was first discovered through fieldwork as reported by Volkov and Stenin (2019), a finding I was immediately able to confirm from my Forest Nenets colleagues, noticing that the form, while productive, was sufficiently rare to not have come up earlier. In my understanding, the markers are directly connected with the verbs FN *xoq-* (< PN *soq-) which besides the concrete meaning 'scoop, ladle' also means 'get something from somewhere', and FN *xoja-* (< PN *soja-) 'acquire, gain'.

Another form unique to Forest Nenets is what might be called the "decorative" in *-kăɬyiq*, whose only discernible function is to make the predicate, whether verbal or nominal, longer, especially for metric purposes in folklore, cf. *kata-xăɬyiq-tă-ŋa-m°* 'I am going to kill it' (vs. *kata-nă-ŋa-m°* id., *-tă* ~ *-nă* marking future) and *xoma-xăɬyiq* 'it is good' (vs. *xoma* id.).

Modern Forest Nenets is an exception to the general rule that the obj.sg verbal suffixes are identical with possessive suffixes. Notably, the PS 1sg possessive marker *-mə has largely been replaced with *-j°*, although the regular outcome *-m°* is still used with kinship terms in particular, cf. *xämă-j°* 'my eye' vs. *wyĭjă-m°* 'my son-in-law', both recorded from the same speaker. The only conceivable origin of FN nom.sg 1sg *-j°* would seem to be the oblique 1du suffix *-j°* which derives regularly from PS *-niŋ (cf. below) via the sound change PN *ny > FN *j* in intervocalic positions.

Among the lexical innovations of Forest Nenets, the word for 'needle' requires particular attention. FN *xătułăpsy°* 'needle' represents a productive instrumental noun derived from the intransitive verb *xătułă-* (~ TN *sædorə-* < PN *sætorə-) 'sew', and it could easily be viewed as direct inheritance from Proto-Nenets, if not for the fact that no such derivative is found in Tundra Nenets where 'needle' is *nyíbya* < PN *nyæmya < PS *äjmä [SW 22 *ejmä (mt) ~ *nejmä] < PU *äjmä. The regular outcome of PN *nyæmya in

Forest Nenets would be ****nyemya***, which, however, would be identical with FN *nyemya* 'mother' < PN *nyemya < PS *ämä [SW 23 *emä] < PU *emä, and avoidance of homophony appears to be a plausible explanation for the loss of Proto-Uralic word for 'needle' in Forest Nenets. Another ancient word with no reflex in Forest Nenets would be PU *l'i̯pśi̯ > PS *l'ępsə̑ (> TN *yebc°*; cf. above) 'cradle', for which Forest Nenets has *ŋĭpu* of unknown origin. The Eastern Khanty borrowing *jăŋkăł* 'mouse' has been mentioned earlier, but another loan from the same source would be *law°* 'horse', replacing the older Turkic loanword (> TN *yuna*) discussed below.

4.4.5 Selkup

Selkup (Sk, which specifically refers to "Common Selkup" as defined below) is a language complex originally spoken in a relatively compact territory between the upper reaches of the Ob and Yenisei rivers, mainly in the modern Tomsk province, close to the area where Proto-Samoyed was presumably spoken. Nowadays, however, virtually all speakers of Selkup, whose number probably does not exceed 1,000, are those of Taz Selkup, located much further north, mainly in the Krasnoselkup municipality in the Yamal Nenets district, because of a relatively recent migration. Of the other dialects groups, Narym Selkup has two speakers, Ket Selkup either one or none, while both Tym Selkup and Upper Ob Selkup are undoubtedly extinct. For the only comprehensive all-Selkup dictionary (SkWb), Alatalo was still able to salvage materials from Upper Ob Selkup and Ket Selkup, the latter being also documented in a specific dictionary (Alatalo 1998). Taz Selkup expectedly has by far the largest number of publications including a series of works by the so-called Moscow school of Selkup studies, notably Kuznecova, Xelimskij & Gruškina (1980).

Selkup has traditionally been described as a single language, which continues to be the view supported by leading authorities such as Helimski and Alatalo, but it could also be argued that the correct number of Selkup languages is either two, three, or five. What is clear is that there are five distinct dialect groups within Selkup, viz. Taz Selkup, Tym Selkup, Narym Selkup (also referred to as *h* dialect), Upper Ob Selkup, and Ket Selkup. Helimski actually claims that Taz Selkup and Upper Ob Selkup are not only mutually unintelligible but differ "approximately as much as Russian and Polish or Udmurt and Komi" (Helimski 1998b: 549), which, if literally true, would certainly render the single-language interpretation moot. With reservations concerning the Ivankino dialect of Narym Selkup, the Selkup language area can be neatly divided into two major subgroups, a northern-central one covering Taz Selkup, Tym Selkup, and Narym Selkup on the one hand, and a southern one consisting of Upper Ob Selkup and Ket Selkup on the other. The term "southern" may be tricky in the context, because all the other dialect groups besides Taz Selkup (also known as Northern Selkup) spoken in and around the Tomsk region, have occasionally been termed "southern dialects" in Russian-language sources. "Central Selkup", used as a cover term for Tym Selkup and Narym Selkup, is

based on geographic and pragmatic factors, but historically Taz Selkup is an offshoot of Tym Selkup, and they are therefore more closely related than Tym Selkup is to Narym Selkup.

In what follows, I mainly quote Selkup materials from three sources, Alatalo (SkWb and unpublished materials) for what may be termed Common Selkup, Helimski for Taz Selkup, and Alatalo (1998) for Ket Selkup. These materials are of the highest quality, while many sources for the other three dialect groups appear less reliable. This inequality in treating Selkup dialect groups is further compensated by the fact that Taz Selkup and Ket Selkup represent the linguistic and geographic extremities of the Selkup language area.

I quote Common Selkup materials with two modifications to the original. First, the long (or contracted) vowels in non-initial syllables appear here as **ē ā ō** for what Alatalo transcribes as **ä a o**, respectively, in order to make their special status transparent to non-specialist, myself included. Second, the uvular stop is given here as **q** instead of the less distinguishable diacritic marking **ķ** used in SkWb. For the voiced counterpart of **q**, which developed after Common Selkup outside the Taz dialect group, I use *g* as a conventional symbol. As for Taz Selkup, I modify the transcriptions in Kuznecova, Xelimskij & Gruškina (1980: 120) so that they better match Alatalo's practice, notably, the consonant characters *c š* are replaced with *ć ś*, the first-syllable vowel characters *ə y* with *ę į*, and the non-initial *y* with *ə*, while the secondary Taz vowels *ɔ ɛ į*, phonologically marginal at best, are merged with *ā e i*, respectively.

Alatalo's Common Selkup is perhaps best characterized as an underlying pan-dialectal representation from which attested Selkup materials can be straightforwardly derived. When it comes to elements descending from Proto-Samoyed and early borrowings, Common Selkup equals Proto-Selkup, and it can be employed parallelly to Proto-Nenets reconstructions, e. g. Ng *d'enti* ~ TE *d'edi* ~ PN *jennyə ~ Sk (№ 1513) **ćinti** < PS *jänti 'bowstring', Common Selkup being marked in boldface and based on records such as Taz *ćintə* ~ Tym *ćəndə* ~ Narym *ćəndə* ~ Upper Ob *kəddə* ~ Ket *kindi* id. In case of post-Proto-Selkup loanwords, Common Selkup functions like a comparative code, for example, Sk (№ 2465) **sammōk** (< Russian *zamók*) 'lock' has a geminate because Common Selkup intervocalic **m** would change to *w* or be lost in Narym, Upper Ob, and Ket, cf. Sk (№ 1633) **ńomā** > Ket *ńō* 'hare', while in 'lock' the *m* of the loan original is retained, because the word was borrowed after the weakening of the nasal.

The Common Selkup first-syllable vowel system looks very similar to the Proto-Samoyed one, especially with respect to its eight short vowels **i e a o u ę į ü**, and indeed, Selkup is the most conservative branch of Samoyed when it comes to preserving vowel quality. Common Selkup, however, also has no less that ten long vowels **ī ē ā̋ ā ō ū ę̄ į̄ ō̋ ū̋** as well as two diphthongs **uə̂ üə**, implying that the relationship between PS and Sk vowels is more complicated. It is true that, taking into account the widespread merger of PS *i and *e as well the possibly secondary status of PS *ö, most of the Sk short vowels reflect PS vowels directly, but the PS reduced vowels *ə̂ *ə > Sk **a**, while PS *a >

Sk **ā** and PS *å > Sk **uə̑** (although other reflexes are found in certain contexts), e. g. PS *səlčə [SW 132 *sə̑lcə̑] > Sk (№ 2722) **salčə** ‘stump’, *amtə̑ [SW 20 *ämtə̑] > (№ 47) **āmtə** ‘horn’, PS *kålä [SW 59] > Sk (№ 2362) **quə̑lə** ‘fish’. At the same time, Helimski’s law, aptly named by Alatalo, states that PS high vowels yield Sk short high vowels if and only if the second-syllable vowel was a full vowel, notably PS *å *ä, but are lengthened when the second-syllable vowel was a reduced vowel, e. g. PS *sị̊rå [SW 140] > Sk (№ 2689) **sị̄rə** ‘snow’, *mị̊tə̑ [SW 93] > (№ 757) **mī̯tə** ‘liver’. The remaining long vowels result from contractions of PS vowel and vowel-glide sequences, e. g. PS *wüət [SW 177 *wüt] > Sk (№ 1965) **kȭt** ‘ten’, *uj- [SW 29 *u-] > (№ 4) **ū-** ‘swim’.

The inventory of non-initial vowels is much smaller, as it only contains three short vowels **i ə u**, and three long (or contracted) vowels **ē ā ō** (for the notation, see above). Of these, the non-initial **i** has merged with **ə** outside Ket Selkup, and Taz and Tym Selkup have also lost the contrast with **u**, resulting in a system of a single short vowel *ə* (written as *y* in most sources on Taz Selkup) in non-initial positions. Furthermore, all the most common PS non-initial vowels, and the only ones already reconstructed in SW, viz. *ə̑ *å *ä, merge in Sk **ə**, although the contrast between original reduced vowels vs. full low vowels is still seen in Selkup morphophonology so that a suffix containing an original reduced vowel triggers what Alatalo calls a shift stem with a long vowel instead of **ə**, cf., for instance, possessive genitive forms of **qatə** ‘fingernail’ : 2sg **qatōntə** ‘of your [sg] fingernail’ < PS *kə̑tå-ntə̑ : 3sg **qatəntə** ‘of her/his fingernail’ < PS *kə̑tå-ntå. The occurrence of the shift stem does not depend on the original stem vowel, so that a form like PS *kə̑nčə̑-ntə̑ > **qančōntə** ‘of your [sg] sledge’ would still show it.

A unique type of vowel harmony is found in connection with the shift stem triggered by the vowel of the following syllable. As opposed to the above examples with **ō** in back-vocalic stems, **ē** occurs in front-vocalic ones, cf. locative forms Sk **pitə** ‘nest’ : **pitēqən** ‘in the nest’ < PS *petä-kənä vs. **qatōqən** ‘in the fingernail’ < PS *kə̑tå-kə̑nå (as seen in the locatives, the expected final vowel is under certain conditions lost in Selkup). Taz and Tym Selkup have lost this vowel harmony by generalizing **ō** throughout inflectional paradigms, but relics such as *śitēqə* ‘the two’ (< PS *kitä-kə including a dual marker) appear in Taz and Tym as well.

The Common Selkup consonant system includes four nasals **m n ń ŋ**, four voiceless stops **p t k q** (= uvular stop), two affricates **č ć**, two sibilants **s ś**, and three liquids **l l' r** (**l'** being largely restricted to final positions where it reflects *j), plus a glide **j** found only in a couple of instances. The Selkup consonants not present in Proto-Samoyed derive from sound changes **q** < *k whereby **k** < *w, and **ć** < *j, whereas Sk **j** is mainly recorded as a reflex of PS *jm as in **saji** ‘eye’ < *səjmä [SW 132 *sə̑jmä]. Common Selkup had no ***w**, but Narym Selkup, Upper Ob Selkup, and Ket Selkup have developed a secondary *w* through lenition, cf. (№ 1037) **timə** > Taz *timə* ~ Tym *təmə* ~ Narym *təwə* ~ Upper Ob *təwə* ~ Ket *tiwwə* ‘tooth’, while in Taz Selkup *w* is restricted to recent loanwords, unless one regards instances of initial **uə̑** having been reanalyzed as *wə* as in (№ 162) **uə̑ćə** > Taz (?) *wɐ̨ćə* ‘meat, body’ (Upper Taz Selkup subdialects show an initial *m* in such cases which seems to support the idea of reanalysis).

Besides *jm, several other PS consonant clusters have undergone changes in Selkup, most often simple assimilations as in **apsə** 'food' < PS *əmså, but more drastic innovations, occasionally affecting adjacent vowels, are also noted, for instance, (№ 2741) **sēlću** 'seven' < *säjt^{3}wə [SW 139 *sejt^{3}wə (? *sejkwə)], (№ 1798) **kōptə** 'bed' < *wåt^{3}wə [SW 173], (№ 218) **āŋku** 'sleep' < *äŋwå [SW 24 *eŋwå]. In particular contexts, Proto-Samoyed initial glides merge with the following vowel, i. e. *wi- / *we- / *jü- > Sk **ü-**, e. g. *wet [SW 176 *wit] > (№ 107) **üt** 'water', PS *jür [SW 50] Sk (№ 284) **ür** 'fat, oil'.

The following are among the most prominent features separating northern-central Selkup (Taz, Tym, and Narym) and southern Selkup (Upper Ob and Ket), the first three points representing northern-central and the last three southern innovations.

A merger of final stops and nasals into a relatively free variation between stops and nasals has taken place in northern-central Selkup, so that Sk **āŋ** 'mouth' > Taz, Tym, and Narym *āk* (~ *āŋ*), cf. also the most frequent case markers, accusative **-m** and genitive **-n**, which frequently appear as *-p* and *-t*, respectively, in the northern-central dialect groups. Since final consonants often alternate with internal ones, various sources for analogy have been present in the history of Selkup, and the overall picture is complex.

The consonant cluster **ŋq** has assimilated to *qq* in northern-central Selkup, e. g. Sk (№ 585) **poŋqə** > Taz *poqqə* 'net' (~ Ket *poŋgə*).

A small but significant layer of Khanty loanwords only appears in northern-central Selkup. A revealing example is the word for 'Russian', recorded as Sk (№ 2886) **rūś** borrowed from Khanty, with reflexes in Taz, Tym, and Narym (including Ivankino), as opposed to (№ 2172) **qasāq**, an apparent Turkic loan, found in Upper Ob and Ket.

A merger of Sk **ś** with *s* has taken place in southern Selkup, e. g. Sk (№ 2423) **śū̄** > Ket *sū̄* 'snake', Sk (№ 2420) **śē** > Ket *sē* 'tongue' (~ Taz *śū̄*, *śē*).

In southern Selkup, Sk **ć** merges with *k* before front vowels, for instance, Sk **ćinti** yields Upper Ob *kəddə* ~ Ket *kindi* 'bowstring', with cognates with initial *ć* in Taz and Narym as well, or Sk (№ 1433) **ćū̄** > Ket *kū̄* 'belt' (~ Taz *ćū̄*, *ćintə*). Before back vowels, the sound value of Sk **ć** is rather *t'* in southern Selkup.

The original number suffixes have been replaced with a collective marker **-lā** in southern Selkup, although it has spread not only to the Ivankino dialect discussed below but also to the southernmost parts of Narym Selkup proper, also known as the *h* dialect.

Without repeating the above points, each of the five dialect groups may be briefly characterized as follows.

(1) Taz Selkup, an offshoot of Tym Selkup, was only established after the Russian conquest of western Siberia in a territory whose earlier inhabitants spoke Forest Enets, with the effect that Forest Enets was pushed further north as explained above. Taz Selkup is in many ways conservative, in particular because it lacks voicing of intervocalic obstruents and subsequent simplification of geminate obstruents, typical of the other dialect groups, and preserves intervocalic **m ń ŋ** from lenition to *w j γ* (or ultimate loss), respectively, again found in most other parts of the Selkup language area. Taz Selkup has, however, merged **č** with *t*, although this appears to be a rather recent change. Another reductive development is the merger of diphthongs **uə̂** > *ǭ* and **üə** > *ǟ*

or *ə̄* (apparently depending on the preceding consonant), e. g. Sk (№ 1772) **qüə** > Taz *qə̄* 'birch', (№ 960) **tüə** > *tə̄* 'birchbark', (№ 1435) **ćüə** > *ćə̄* 'pine'.

(2) Tym Selkup was the first Selkup dialect group to become extinct and its documentation relies heavily on Donner's fieldwork, the results of which are found in Donner, Sirelius & Alatalo (2004). From the perspective of historical phonology, Tym Selkup shows a more checkered picture, as both preserved and lenited examples of intervocalic **m ń ŋ** are found there. Several features that Tym Selkup nowadays shares with the neighbouring Narym Selkup may be based on recent contacts. The elimination of Selkup vowel harmony, as discussed above, remains an innovation shared by Tym Selkup and Taz Selkup.

(3) Narym Selkup is also referred to as *h* Selkup because of its characteristic sound change *s > *h* (which in internal positions has often led to the loss of the consonant). The actual distribution of this change is, however, more complex, which is why I have used the traditional name Narym Selkup instead of *h* Selkup preferred by Alatalo. Narym Selkup consists of three subunits as follows. (i) A north-western dialect (TyP in SkWb) is otherwise similar to the *h* dialect, showing vowel harmony unlike Tym Selkup, except that it notably lacks the very *s > *h* development. (ii) Narym Selkup proper equals the *h* dialect in the strict sense and covers the bulk of the Narym Selkup area. A large portion of Narym Selkup speakers were forcibly displaced in the 1930s, when those parts of the Tomsk region became a hub of the Soviet concentration camp system. Local linguists interviewing them have not always distinguished between their place of origin and their contemporary location, which has led to certain confusion. (iii) A south-eastern dialect in Ivankino and Kolpashevo has also retained *s*, yet in most other respects it coincides with *h* Selkup as opposed to southern Selkup as discussed above. Southern influences include some instances of points (v), Sk **ć** > *k* before front vowels, and (vi), the collective marker **-lā**, which, however, extend to parts of Narym Selkup beyond Ivankino.

(4) Upper Ob Selkup is the most heterogenous of Selkup dialect groups but, fortunately, it has been thoroughly described by Alatalo (1994). As a whole, Upper Ob Selkup may be defined only negatively, representing those southern Selkup dialects that lack consonant gradation. The dialect zones can be identified on the basis of a single but frequent sound change, viz. denasalization of clusters as follows. (i) A northern group has no denasalization but does exhibit a number of similarities to Narym Selkup, the Ivankino dialect in particular. (ii) A central group has the characteristic denasalization process, e. g. PS *jänti > Sk (№ 1513) **ćinti** > Upper Ob *kiddə* (~ Taz *ćintə* ~ Narym *ćendə* ~ Ket *kindi*) 'bowstring'. Denasalization does not, however, define the border with Ket Selkup in the way gradation does, because it has spread across the main dialect boundary to the lower Ket region. (iii) A southern group, located in the southernmost area where Selkup was spoken, again lacks denasalization. Its typical features include, presumably under Russian influence, secondary palatalization and phonetic diphthongization of long mid vowels, as well as an apparently Turkic influenced variation between initial *l* and *n*, reflecting both Sk **l** and **n**. The Chulym subdialect exhibits the same merger of **č** with *t* as Taz Selkup, which is obviously an example of a convergent sound change.

(5) Ket Selkup is readily recognizable because of its characteristic consonant gradation whereby geminates in the strong grade alternate with single (in the case of obstruents voiced) consonants in the weak grade, e. g. PS *kâtå 'fingernail' : gen.sg *kâtå-n : 1sg *kâtå-mə > Sk **qatə** : **qatən** : **qatōmə** > Ket *qattə* : *qadən* : *qattōm*, with the weak grade in the beginning of closed syllables if and only if a short vowel follows (incidentally, only Ket Selkup keeps Sk non-initial **i u** distinct from **ə**). The strong grade is regular before a long vowel, cf. Sk (№ 2172) **qasāq** 'Russian' > Ket *qassāq*, creating a marked difference from Upper Ob *qazāq* id. Reverse gradation works as in Nganasan (and Finnic), e. g. Sk **čopər** : **čopərən** > Ket *čobər* : *čoppərən* 'berry', as does rhythmic gradation, cf. the comitative Sk **qatəsē** > Ket *qattəzē* 'with the fingernail'. Unique features of Ket gradation are, first, that it covers all consonants, cf. Sk **timə** : **timən** > Ket *tiwwə* : *tiwən* 'tooth', and, second, that it does not cover any consonant clusters.

Thanks to its central position within Samoyed, one might expect that Selkup would also be lexically conservative. The number of core words restricted to Selkup is, however, relatively high, for instance, (№ 94) **ɐ̈čі** 'word, language' vs. the widespread PS *wåât[1]å (> Ng *buəδu* etc.). As opposed to the other Samoyed languages which have a Turkic borrowing as their numeral for 'hundred', Selkup employs (№ 1111) **tōn** probably reflecting an original meaning 'number, digit' as explained below. Selkup alone lacks a reflex of PS *åj 'foot, leg', the concept being expressed by Sk (№ 977) **topə** 'foot, leg' < PS *topå [SW 166] 'hoof', while PS *kâtå [SW 55] 'fingernail' has broadened its meaning to Sk (№ 1876) **qatə** 'fingernail, hoof'. The reflex of PS *kåjå [SW 58] 'sun' has shifted its meaning to Selkup (№ 2054) **quâćə** 'heat', while Sk (№ 1618) **ćēlə** 'day, sun' < PS *jalä [SW 40 *jälä] 'light, day, sun' is the only expression for 'sun' in Selkup. PS *mânå 'egg' has experienced a semantic shift to (№ 772) **manə** 'penis' throughout the Selkup language area, yet, curiously, there is no common Selkup word for 'egg', cf. Taz *eŋ ~ ēŋ* borrowed from Yeniseian, (№ 1630) **ńāp-ī** lexicalized from the meaning 'duckling' found in Taz, Narym, and Ket, and the Upper Ob (notably including Ivankino) word (№ 2123) **kį̈kaj** apparently of Turkic origin.

In its grammatical structure, Selkup is decidedly less complex than the northern Samoyed languages. For instance, Selkup has no reflexive conjugation or destinative declension, although traces of both may remain, and there is little doubt that both categories derive from Proto-Samoyed. Selkup does preserve the contrast between subjective and objective conjugations, but only has one set of personal suffixes irrespective of the number of the object, which itself is left unmarked as in northern Samoyed, cf. Sk **qoŋāl** which is cognate to FN *koŋał°* < PS *ko-ŋå-rə 'you found it' but also fulfils the functions of FN *kojăt°* < PS *ko-jə-tə 'you found them [pl]' and FN *koŋaxăjot°* < PS (?) *ko-ŋå-kə́ńəj-tə 'you found them [du]'. Similarly, Selkup has no distinct sets of possessive suffixes depending on the number of possessed as in FN *kănăt°* < PS *kənčə-j-tə 'your sledge' vs. *kănot°* < PS *kənčə-j-tə 'your sledges', although there is a uniquely Selkup marker of plural possessed forms, cf. Taz Selkup (Kuznecova, Kazakevič, Gruškina & Xelimskij 2002: 121–123) *qopām* 'my fur' : *qopōqim*(*ə*) 'my furs [du]' (with a dual marker *-qi*) : *qopəīm*(*ə*) 'my furs [pl]' (with a plural marker *-ī*, the historical background of which

may stretch to Proto-Uralic, yet problems regarding its exact reconstruction remain). Selkup also has few traces of the so-called plural oblique stem, reflecting PU and PS postvocalic *j in non-initial syllables, cf. Sk **amər-** ~ Ng *ŋəmur-* ~ TE *ôôr-* ~ TN *ŋəwor-* < PS *ə̑mə̑jr- 'be eating'. A prospective explanation for the relative simplicity of Selkup grammar has been offered on the basis that Selkup formerly served as a lingua franca among speakers of several other languages such as Eastern Khanty, Evenki, and Ket (Helimski 1996).

4.4.6 Kamas

The Kamas language (Km; ethnonym *kaŋmāžə*) was last recorded in the village of Abalakovo (Абала́ково) in the upper Yenisei basin but spoken earlier in the surrounding mountain region. Kamas survived into the 20th century, losing its last native speaker in 1989. Kamas has been recorded and studied in several distinct stages, starting with early explorers such as Müller in the 18th century who also collected word lists from the other, soon to be extinct dialect Koibal, followed by Castrén in the 19th century when Kamas was still used as a community language, by Donner in the early 20th century at the time when only a handful of speakers remained, by Künnap in the late 20th century after the unexpected discovery of two people still able to speak Kamas, and after the extinction of the language especially by Klumpp, who has been able to create a comprehensive picture of Kamas out of the highly uneven records. Not unlike in the case of Forest Enets, it is possible to talk about early Kamas and late Kamas, or in Klumpp's terms, pre-shift and post-shift Kamas, referring to the status of the language before and after the time when the speakers stopped communicating in Kamas and only remembered the language thereafter. Both the consonant and vowel systems seem to be affected by this shift, and I quote here Klumpp's analysis with a focus on late Kamas. Donner's materials are published in his posthumous dictionary (Donner 1944), while Künnap used his collections for a two-volume monograph on Kamas inflection (Künnap 1971, 1978), and Klumpp (2002) studied Kamas aspect markers that had been grammaticalized from verbs. In his grammatical sketch of Kamas, Klumpp (2022) was forced to use IPA, manifestly unsuited for the purpose, but, fortunately, much of his data is also available at <www.infuse.finnougristik.uni-muenchen.de/e-learning/kamas/o1_kamas.pdf> in proper transcription. Janurik (2021) provides an updated summary of Kamas lexicon.

The Kamas consonant system consists of four nasals *m n ń ŋ*, four fortis stops *p t t' k* plus *ʾ* representing (preconsonantal) glottal(ization), four lenis stops and glides *b d j g*, two sibilants *s š* and a further fricative *h*, and two liquids *l r*, disregarding secondarily palatalized consonants which reflect PS *Vj sequences. The basic lenis stops derive from PS intervocalic stops, while in syllable-final position they appear preglottalized, e. g. PS *petä > Km *pide* 'nest' vs. *mat > *maʾd* (~ *maʾn* ~ *maʾ*) 'tent, house'. The cacuminal sibilant *š* corresponds to the palatal sibilant *ś* elsewhere in Samoyed, deriving from *k before front vowels (cf. Selkup materials which mostly employ *š* for what Alatalo

transcribes as *ś* since it is historically primary). PS initial glides have two reflexes, in the general case *w > *b* and *j > *t'*, but *w > *m* and *j > *n* when a nasal follows, e. g. *wäsa [SW 175 *wesä] > *basa* [baza], *jåå [SW 36] > *t'ü*, *wenä [SW 175 *winä] > (*mine* >) *müne* 'strap', *jåmpâw [SW 37 *jåmpâ] > *numo* / *numu* 'long'. Other conspicuous consonant changes include PS *ń > Km *n*, e. g. *ńåc > *noʾd* 'grass', PS *lč / *lt > Km *ll*, e. g. *səlčə > *sällə* 'stump', PS *nč / *nt > Km *n*, e. g. *wånčå > *mona* 'root', PS *mt > Km *mn*, e. g. *tumtâ- > *tĭmnə-* 'recognize', and PS *t > Km *š* when followed by *j, e. g. *tojmå > *šōmi* 'larch'.

The Kamas vowel system, at least in its later stages, was modeled after its most important contact language, viz. Khakas, before Russian became dominant in the region. It consisted at minimum of four full front vowels *i ü e ö*, three full back vowels *u o a*, to which may be added at least *į*, often found in loanwords, and two reduced vowels, front and back, *ĭ ă*, neutralized as *ə* in non-initial syllables. Two more full vowels, front *ä* and back *å*, may also have been contrastive, and the system is augmented by a number of (long) contracted vowels, apparently of recent origin. Diachronically the most conspicuous changes were the loss of PS *į *ę which merged with respective front vowels, and the conditional PS *u > *ü*, with a concomitant chain change *o > *u*, and *å > *o* (or *u*), e. g. *tumtâ- > *tĭmnə-* 'recognize', *kopå > *kuba* 'fur', *kålä > *kola* 'fish'. Kamas has developed a purely phonetic vowel harmony, again on the model offered by Khakas and parallelly to Mator, cf. PS *sįŕå > Mt (№ 891) *sirä* 'snow', with the plural *sirä-zäŋ* (Simoncsics 1998: 582; cf. below).

The most characteristic consonant change of the long-since extinct Koibal dialect was the merger of *š* with *s* (for further details, see Mikola 2004: 90), while any vowel changes would remain ambiguous because of the scanty material. There appear to be, however, more archival materials of Koibal than known in general publications, so that new discoveries are still possible.

As mentioned above, Kamas has innovated a comitative case in *-seʾ*, similar to Selkup **-sē**, while the plural marker *-sAŋ* is unique to Kamas. In verbal inflection, the most striking innovation is the virtually complete loss of the so-called aorist marker PS *-ŋå, mainly replaced with a *li* element in the function of the present tense.

4.4.7 Mator

The Mator language (Mt; formerly also known as "Motor"; ethnonym probably **modor*, cf. MS 63) was spoken in the northern slopes of the Sayan Mountains until around 1840, shortly before Castrén visited the area and tried to locate surviving speakers. Thanks to Helimski's efforts, however, we have a more comprehensive picture of Mator than probably of any language only attested in wordlists. The dialects of Mator, i. e. Mator proper, Taigi, and Karagas, are close, but Mator proper is characterized by the depalatalization *ń > ***n**, and Karagas, uniquely in the Uralic language family, adopted loanwords directly from Mongolic, more exactly Buryat. As mentioned earlier, "Abakan Samoyed" repre-

sents a conglomeration of Kamas and Mator materials (Klumpp 2019), while "Soyot" constitutes a synonym of Karagas.

As presented, or rather reconstructed, by Helimski (1997: 70) the Mator sound system appears quite archaic both for consonants and vowels. Among vowels, in particular, few marked innovations can be pointed out, but several interesting details can be recovered in the case of consonants. As mentioned above, Mator not only preserves *k from assibilation, but also exhibits the typologically unusual shift PS *s > Mt ***k*** before front vowels, e. g. PS *sejə > Mt (№ 455) ***keje*** 'heart'. Another conspicuous change, which happens to be reminiscent of Nganasan, is PS *p > Mt ***h*** in initial and intervocalic positions, except that PS initial *p > Mt ***b*** when the following consonant is *s, which is also the development found in preconsonantal positions, e. g. PS *påjŋå [SW 115] > Mt ***hajŋa*** 'mattress', *kopå [SW 73] > ***koha*** 'fur', but *posåmå > ***bosama*** 'rotten', cf. also *jåptå [SW 38 *jåptå (? *jåâptå ~ ssm *jäptå)] > [***čabta***] 'thin'. As in Kamas, PS initial glides have two reflexes in Mator, in the general case *w > ***b*** and *j > ***ć***, but *w > ***m*** and *j > ***ń*** in case of a following nasal, e. g. *wäsa [SW 175 *wesä] > ***besä***, *jåâ [SW 36] > ***ča*** (? ***čā***), *wenä [SW 175 *winä] 'strap' > ***minE***, *jåmpâw [SW 37 *jåmpâ] > ***ńambuh*** 'long'. As seen in the last example, Mator exhibits a segmental reflex of PS final *w (Gusev 2008), cf. the Nganasan section above. Preglottalized consonants appear much like in Kamas, e. g. PS *mat > Mt ***mä***[ˀ]***t*** ~ ***mäˀn*** 'tent, house'.

4.5 Topics in Samoyed phonology

Proto-Samoyed vowel harmony, preserved in an altered form in Nganasan, clearly represents a continuation of an earlier Proto-Uralic system. Unlike in Proto-Uralic, where vowel harmony is presumed to have been exceptionless, with only back or front vowels in a single word-form, making it factually a phonetic phenomenon, Proto-Samoyed had a few disharmonic roots, notably as a consequence of a sound change concerning second-syllable PU *a > PS *ä after laterals and spirants. Compare PU *kala > PS *kålä [SW 59] 'fish', whose reflexes might still be called disharmonic not only in Nganasan (Ng *koḷi*) but also in Nenets (TN *xalya*), where original frontness is expressed in consonant palatalization, in this case present only in the second syllable while the great majority of Nenets stems have either non-palatalized or palatalized throughout the word-form. Nenets also shows direct relics of back vs. front allomorphs of suffixes in the same manner, cf., for instance, PS *matkånå 'in a tent' > PN *myaqkəna > TN *myak°na* vs. *wetkənä 'in water' PN *wyiqkənya > TN *yik°nya*, although this kind of "consonant harmony" has become optional in Tundra Nenets where (historically disharmonic) *yik°na* is equally grammatical and has been virtually eliminated in Forest Nenets where only *wyĭqk°na* is nowadays used.

Nganasan consonant gradation is to some extent described from a synchronic point of view above, but it needs to be added that it has a close parallel in the system of gra-

dation found in many (but not all) Finnic languages, which has led many specialists, notably Helimski, to assume a common Uralic origin for the phenomenon. While the principle of consonant gradation including the two types of gradation mentioned above holds good in Finnic, there are also major differences in its scope, notably with regard to consonant clusters, where Nganasan often shows no gradation when Finnic does, cf., for instance, Ng *turku* 'lake' : pl *turkuq*, and vice versa, Nganasan has gradation in sibilants which is not the case in Finnic. What is more, Finnic gradation has features that makes it more likely that it only emerged dialectally in Proto-Finnic and eventually covered, albeit it slightly different forms, in much of Finnic, as well as most, but again not all Saami languages, where it would in this scenario derive from Finnic influence, leaving Nganasan gradation as a parallel but unrelated innovation. That said, there may well be relics both in Enets and Nenets of rhythmic, or suffixal, gradation in particular, which point to a considerable age of at least this type of gradation. One of them is the loss of the consonant of the gerund marker, Pre-Enets-Nenets *ki, in polysyllabic stems, e. g. TN *xo-sy*° 'to find' vs. *xada-*° 'to kill', and another the loss of the nasal after an unstressed syllable in certain suffixes such as the durative marker, cf. TN *xo-mpə-* 'be finding' vs. *xada-bə-* 'be killing'. It would also be tempting to connect the development of PU *t > (*δ >) PS *r in the nominative 2sg possessive suffixes to gradation, since these suffixes most frequently occur after bisyllabic stems, e. g. *kopå-rə̂ 'your fur', as opposed to, notably, the retention of PU *t in the ablative suffix, commonly found after a stressed syllable, as in *kopå-kə̂-tə̂ 'from fur'. These are obviously very tentative ideas, and the firmest conclusion would still be that the Nganasan consonant gradation arose independently, albeit under similar prosodic premises.

Nasal prothesis is restricted to the northern areal of Samoyed languages, yet it required a detailed description both as a historically and typologically interesting innovation and as proposed evidence for a distinctly northern Samoyed subgroup. The modern reflexes of nasal prothesis are as follows. (i) Western Tundra Nenets has lost the Proto-Nenets initial *ŋ under the influence of contact languages, i. e. Komi and Russian; the loss of ŋ is also found other positions, e. g. *aqo* vs. *ŋaqŋo* 'long-tailed duck' in other dialects; the palatalized prothetic *ny* has been preserved in western Tundra Nenets. (ii) Central Tundra Nenets is the only variety of Samoyed without vowel-initial word-forms, so that even recent loanwords will add *ŋ* to prevent a phonotactically impossible sequence, e. g. *ŋan°tyena* 'antenna' (< Russian *anténna*). (iii) Eastern Tundra Nenets has a small number of vowel-initial word-forms, viz. recent loanwords and a few pronominal forms. (iv) Forest Nenets also has a small number of vowel-initial word-forms, but their distribution is different from eastern Tundra Nenets, and the group also includes basic words such as FN *apa* 'older sister' and *ïja* 'older brother'. (v) In Enets, nasal prothesis seems to be restricted to monosyllabic words and their forms and derivatives, making it the most distinct group within (northern) Samoyed, the Enets languages being notably uniform in this respect. (vi) In Nganasan, the situation is similar to Forest Nenets, i. e. the number of vowel-initial word-forms is small, but still notable. The words in question, however, typically do not have cognates in Forest Nenets or elsewhere.

In a handful of cases, a palatal glide appears in positions where nasal prothesis would be expected in either Nganasan or Nenets, and the material appears contradictory, cf. the Nganasan reflex in PS *utå [SW 30] > Ng *d'utü* ~ TE *uδa* ~ Yr ‹múde› ≈ ***ŋuδa*** ~ TN *ŋuda* ~ Sk (№ 87) **utə** ~ Km *uda* ~ Mt (№ 1109) ***uda*** 'hand, arm', and the Tundra Nenets reflex in PS *elä- [SW 27 *ilä- (ng, sk, tg) ~ *jilä-] > Ng *ńilį-* ~ TE *ire-* ~ TN *yilye-* ~ Sk (№ 344) **ilə-** 'live'. Other currently unaccountable developments include the apparent metathesis in Nganasan in the word PS *äjsä [SW 22 *ejsä] > Ng *d'esį* ~ TE *ese* ~ TN *nyísya* ~ Sk **ɐsə** 'father'.

4.6 Samoyed word classes

The term "noun" is used in this chapter in a broad sense, covering (some) adjectives as well as numerals and (non-personal) pronouns, insofar as their morphological characteristics coincide with the prototypical "substantive" nouns. Furthermore, several categories of adverbs and postpositions show some inflectional features with nouns. Effectively, verbs and nouns represent the only open word classes, and while in older literature there have appeared claims that the borderline between verbs and nouns is somehow vague in the Samoyed languages, this is not the case (Salminen 1993 and the literature therein). Notably, while Samoyed verbs and nouns share two kinds of suffixes, personal and possessive, the morphological patterns are always distinct, since suffixes marking person are not attached directly to a basic verbal stem, as seen in examples below.

Adjectives, as a predetermined entity, are a divided category in Samoyed, insofar as both morphological verbs and nouns are present in a particular semantic field. Notably, Samoyed colour terms tend to involve both morphological nouns and morphological verbs, cf. TN *serako* 'white' and *tasyexey°* 'yellow' vs. *pəryidye-* 'be black' and *nyar°ya-* 'be red', cf. *serako-d°m* 'I am white' vs. *pəryidye-ə-d°m* 'I am black' where the personal suffix does not attach directly to the basic stem of a verb. In negation, *serako nyí ŋaq* 'it is not white' requires the connegative form of the copula (*ŋœ-* 'be' :) *ŋaq* because nouns have no connegatives while verbs do, cf. *nyí pəryidye-q* 'it is not black'. As an attribute, verbal stems occur as participles, cf. *pəryidyenya ti* 'black reindeer' vs. *serako ti* 'white reindeer'. A similar pattern is repeated in other semantic fields, notably with regard to temperature, the extremes being expressed by verbs and the moderate options by nouns, i. e. TN *yiba* 'warm' and *xənco* 'cool' as opposed to *yadyempə-* 'be hot' and *tyecyə-* 'be cold', cf. also TN *yabta* 'thin' vs. *nyaŋoq-* 'be thick', which may give credence to a hypothesis that the somehow "stronger" qualities are expressed by verbs and those deemed "weaker" by nouns. Adjectives, whether morphologically nouns or verbs, may share derivational morphology, e. g. TN *yabtarka* 'thinner' vs. *nyaŋosərka-* 'be thicker', with a comparative suffix *-rka*, yet the derived forms remain nouns and verbs, respectively.

Samoyed numerals have attracted considerable interest because they are rather different from those of the other Uralic languages, collectively called "Finno-Ugric" from a traditional viewpoint, that share numerals up to 'six' or 'seven'. In short, only three numerals appear to be common to all of Uralic, viz. PU *käktä > PS *kitä 'two', PU *wiγti 'five' > PS *wüət 'ten', and PU *śäjśimä (Aikio 2022) > PS *säjt^{3}wə 'seven', yet each of these comparisons presents challenges, none of them, however, compromising the cognacy of the words themselves. The simplest reconstruction for 'one' in Proto-Samoyed is *op, a rare *p stem as such and not an immediate cognate of forms found in the other Uralic languages, which are not unified in this respect either. There is no doubt that PS *kitä 'two' derives from the PU source indicated above, but the first-syllable vowel is irregular, or rather it belongs to a small but distinct group of words where a PS high vowel reflects a PU low vowel, notable examples of which are also found among kinship terms and words denoting body parts, e. g. PU *wäŋiw > PS *wiŋiw 'son-in-law'. As for PS *nakur [SW 99 *näkâr] 'three', a superficial and far from conclusive resemblance of the second syllable *-kur to the Hungarian and Mansi 'three' (Hungarian *három*) may be noted. PS *tät^{2}tə [SW 159 *tet^{2}tâ (? *tettâ)] 'four' must be a Samoyed innovation because of its consonant cluster alone, as PU geminates would not have survived in Proto-Samoyed as such. PS *səmpəlaŋkə 'five' offers a major reason to believe that the Samoyed numeral for 'ten' discussed below represents a semantic shift, insofar as it manifests a derivative consisting of four syllables and therefore patently a secondary development, as pointed out by Janurik long ago. PS *mâktut 'six' constitutes another Samoyed innovation, yet it must be mentioned that Jaakko Sarvela once suggested that PU *kuγtį (or perhaps *koγtį) 'six' could be seen as a derivative of the numeral 'two' in the same manner as PU *wiγti 'five' would be derived from 'one', cf. Erzya (Mordvin) *vejke* 'one', a process that must have already taken place in Proto-Uralic. PS *säjt^{3}wə 'seven' has always been recognized an Indo-European borrowing, specifically from Tocharian, and as such parallel to many other similar forms in the other Uralic languages, yet Aikio (2022) emphasizes its direct cognacy with the Finnic word for 'seven', of which I remain uncertain. The PS numeral 'eight' invariably combines the numerals 'two' plus 'four', e. g. TN *syid°ntyet°* (cf. TN *syidya* 'two' and TN *tyet°* 'four'), but an exact reconstruction is difficult to come by. PS numerals for 'nine' vary greatly, cf. Klumpp (2005). That the PS *wüət [SW 177 *wüt] 'ten' is a regular reflex, with a semantic shift related to the development of the Samoyed numeral system as a whole, of PU *wiγti 'five' should not be doubted. This has been questioned on the basis of an assumption that most intermediate numerals were lacking, along with misguided ideas about the chronology of Proto-Uralic and Proto-Samoyed. The original vowel sequence has also not always been understood correctly, and Kuznecova, Xelimskij & Gruškina (1980) and several other sources miss the crucial vowel length with records such as **köt* instead of the correct *kȭt* in Taz Selkup. Unlike the other Uralic languages, which have borrowed the word for 'hundred', albeit perhaps parallelly into the different branches, from Indo-European, the Samoyed word PS *jür [SW 50] > Ng *d'ir* ~ TE *d'uq* ~ TN *yur* ~ FN *jŏł* ~ *jŭł* ~ Km *t'ür* (> *t'üs*) ~ Mt ***čür*** (> ***čüs***) represents a Turkic loan; the final *s* in the later Kamas and Mator forms was

established through secondary contact with so-called *z* Turkic languages, while the common Samoyed form was borrowed from some *r* Turkic language; Selkup has no reflex of PS *jür but Sk (№ 1111) **tōn** 'hundred' probably meant 'number, digit' originally.

Personal pronouns constitute an entirely separate word class, although there are numerous differences among the languages, including the fact that in southern Samoyed the third person pronouns are replaced with demonstratives inflected like nouns, so that only the first and second person pronouns are included in a special word class. Only the first person pronouns are shared by all languages, and even their correspondences are highly irregular, cf. Ng *mənə* vs. TE *mod'i* ~ FE *mod'i* ~ Yr ‹modi› [≈ *mədi*] ~ TN *məny°* ~ FN *măny°* vs. Sk **man** ~ Km *măn* ~ Mt ***mən*** 'I'. The shortest forms, found in southern Samoyed, are at least superficially cognate (< PS *mən) with each other and perhaps with some forms in other Uralic languages, which also show many irregularities, presumably because of the ultimate origin of the Uralic personal pronouns as Indo-European borrowings. The Enets, Yurats, and Nenets forms represent a shared innovation, deriving from Pre-Enets-Nenets *mənti. In the second person, Nganasan and southern Samoyed have rhyming words, cf. Ng *tənə* vs. TE *tod'i* vs. FE *uu* vs. TN *pidər°* vs. FN *pĭt°* vs. Sk **tan** ~ Km *tăn* ~ Mt ***tən*** 'you [sg]', but only the Tundra Enets forms derives from *tənti, while both this and the following Forest Enets pronoun were actually borrowed from Ket (Siegl 2008), and the Nenets second and third person pronouns derive from the stem of a noun meaning 'body' that was first grammaticalized as a reflexive pronoun. The third person pronouns Ng *sʲiʲi* vs. TE *ńiitoda* vs. FE *buu* vs. TN *pida* ~ FN *pĭta* '(s)he' show again major differences among northern Samoyed languages, although here, unlike in the second person pronoun, the Nenets forms are exact cognates. The dual and plural personal pronouns are generally formed by replacing the singular possessive suffix contained by the singular pronouns accordingly, cf. dual Ng *sʲi-ti* vs. TE *ńiito-diq* vs. TN *pi-dyih* : plural Ng *sʲi-tʲiŋ* vs. TE *ńiito-duq* vs. TN *pi-doh*; for more details, see Siegl (2008). A major difference between Nganasan and the other Samoyed languages is that Nganasan has no distinct accusative forms of personal pronouns but relies on verbal indexing, while at least in Enets, Nenets, and Selkup accusatives are formed from a stem grammaticalized from PS *ket [SW 70 *kit[1]] 'likeness, image' < PU *keti (Aikio 2006), e. g. TN *syiqm°* 'me'. Curiously, the above-mentioned Nganasan pronouns *sʲiʲi* : *sʲiti* : *sʲiʲiŋ* represent grammaticalizations of the same stem in a different function.

Other minor word classes include adverbs and postpositions that exhibit some inflectional features of nouns. The paradigm of local adverbs consists of four case forms, e. g. based on the interrogative pronominal stem PS *ku- [SW 75] Tundra Nenets has the following forms with the basic meaning 'where': TN dat. *xuh* (< PS *ku-ŋ) : loc. *xuna* (< PS *ku-nå) : abl. *xud°* (< PS *ku-tə̑) : pros. *xumna* (< PS *ku-mnå), i. e. with single local case suffixes unlike in the nominal inflection where the case markers are preceded by a co-affix as explained below. The short prosecutive marker *-mnå is only found in Nenets, while elsewhere the longer marker *-mə̑nå is employed both in adverbial and nominal inflection; for example, TE *kuone* (< Pre-Enets *ku-mə̑nå) is semantically equivalent with TN *xumna* (< Pre-Nenets *ku-mnå). It is not immediately obvious whether

Nenets innovated the single-syllable marker or the other languages generalized the nominal prosecutive marker into the adverbial paradigm. What is more, the Nganasan adverbial prosecutive marker is generally recorded as *-məənu*, i. e. with an unetymological vowel sequence, which at least on the surface looks like a hypercorrection.

Local postpositions are inflected not only for (local) case but also for person, meaning that their paradigm contains altogether 40 units. There are thus four absolute forms and four times nine personal forms, e. g. TN *myad°h nyana* (loc.) 'at the tent' (< PS *mat-ə-n na-nå) vs. *nyanan°* 'at me' (< *na-nå-nə), with forms of the basic postpositional stem TN *nya-* (< PS *na- [SW 99 *nä-]), which also fills the position of local case forms of personal pronouns. Random examples of other local postpositions include TN *ŋil°nta* (dat.) 'under it' (< PS *i̯lə-n-tå : *ŋilə-* < *i̯lə- [SW 24] 'under'), *nyid°nyih* (abl.) 'from top of us [du]' (< PS *i-tə-niŋ : *nyi-* < *i- [SW 26 i(-) (sk) ~ *ni] 'on'), and *yəxah myumnya* (pros.) 'through the river' (*jəkå-n mü-mnä : *myu-* < *mü- [SW 96 *mü (? *müə)] 'in'). While the relics of vowel harmony in Nenets are mostly optional as discussed above, the locative and prosecutive forms of *nyi-* and *myu-* always show palatal consonants, i. e. *nyinya* : *myunya* and *nyimnya* : *myumnya*, respectively.

Three additional minor word classes can be illustrated using Tundra Nenets material (cf. Salminen 1997). The paradigm of non-local postpositions consists of an absolute form and nine personal forms as in *xadanaq xaw°na* 'except our grandmother' vs. *xaw°nan°* 'except me'. Referential or connective adverbs have exactly nine forms in their paradigm, because they invariably agree in person with the subject of the sentence, e. g. *ŋədy°byan°* 'therefore I' : *ŋədy°byan°* 'therefore you [sg]' etc. Finally, many adverbs are uninflected, and may be called particles instead, such as *ŋulyiq* 'very', although most members of this word class derive from earlier complex forms, including *təmna* 'yet, still', a lexicalized prosecutive of a pronominal stem, and *məlye°* 'already', a lexicalized gerund of a verb meaning 'finish'. Furthermore, the internal structure of particles such as *tyedah* 'now' is shown by 'omnibased' derivatives, e. g. *tyeda-ryi-h* 'for the time being', with the limitative suffix *-ryi* 'only' (see subsection 9.6.). On the basis of both internal and external reconstruction, it may also be claimed that Samoyed possesses particles that derive from so-called Uralic *s latives, cf. TN *nyer°q* 'beforehand' ~ Finnish *edes* 'even' (Salminen 2014).

4.7 Samoyed stem types

The fundamental distinction in the Samoyed stem types is between vowel and consonant stems. As mentioned in the introduction, Proto-Uralic did not possess a phonotactic structure *(C)VC in major word classes, i. e. verbs and nouns (in the broad sense), which means that this distinction by and large represents an innovation in Samoyed. There probably were polysyllabic consonant stems of the structure *(C)VC(C)VC in Proto-Uralic, but in comparison with the PU canonical *(C)VC(C)V pattern they were few

in number and, with the exception of the significant group of *w stems discussed in some length below, left little trace in Samoyed. In any case, three morphophonological processes connected with consonant stems can be recognized for Proto-Samoyed and its descendants: (i) epenthesis, or the addition of a reduced vowel to a consonant stem when a suffix consisting of a single consonant is added, (ii) truncation, or the loss of the initial consonant of a suffix-initial consonant cluster, and (iii) the replacement of initial *r with *l when a suffix is attached to a consonant stem, e. g. (PU *weti : gen.sg *weti-n >) PS *wet 'water' : gen.sg *wet-ə-n : dat.sg *wet-tə-ŋ : *wet-lə 'your water', as opposed to the parallel forms of a vowel stem such as (PU *künti : gen.sg *künti-n >) PS *küntə 'smoke' : *küntə-n : *küntə-ntə-ŋ : *küntə-rə. The resulting alternations are still largely valid for the attested Samoyed languages, cf. TN *yiq* : *yid-°-h* : *yi-t°h* : *yiq-l°* vs. *syun°* : *syun°-h* : *syunə-n°h* : *syunə-r°* via regular sound changes. In Selkup and Kamas, however, the possessive suffixes with initial *l* have been generalized to vowel stems as well. In Nenets and Kamas, because of the convergent loss of the stop in *nt clusters, an alternation between a nasal and a stop has emerged, as in the TN dative forms *syunə-n°h* vs. *yi-t°h* above. Exceptions to the conditions of epenthesis vs. truncation are possible, for instance, the comparative and iterative suffixes (TN *-rka* and *-ŋkə*, respectively) trigger epenthesis, e. g. TN *xœw°rka* 'shorter' derived from *xœm* 'short' or *səl°ŋkə-* 'keep returning' from *səl-* 'return' show epenthesis instead of truncation sanctioned by the general rule.

The most conspicuous alternations affecting vowel stems, preserved in northern Samoyed as opposed to Selkup and Kamas, where they have largely been eliminated via analogy, derive from the widespread contraction of non-initial syllable-final *Vj sequences. Given that the accusative plural marker originally consisted of *j alone, the secondary substem based on it could simply be called the accusative plural stem in Enets and Nenets, cf. the Tundra Nenets accusative plural forms *wadyi* of *wada* 'word' or *xale* of *xalya* 'fish' with similar vowel changes in denominal verbs such as translatives *səqlyim-* 'become an idiot' (*səqla* 'idiot') or *yalem-* 'lighten, dawn' (*yalya* 'light, day'). This terminology does not, however, work for Nganasan, where PS final *j is preserved, as in the corresponding accusative plurals *buəδuj* (: *buəδu* 'word') and *kolij* (: *kolį* 'fish'). When followed by a consonant, however, *j in non-initial syllables is lost in Nganasan as well, triggering a parallel vowel alternation as in the corresponding genitive plurals *buəδaq* and *kolaq* as well as the translative derivative *d'alam-* (~ TN *yalem-* < PS *jalåjm-) of *d'alį* (~ TN *yalya* < PS *jalä) 'light, day'. In Nganasan studies, the substem represented by *d'ala-* etc. is simply called the third stem (the second stem being the vowel stem created through epenthesis as explained above), but since even in Nganasan it is found in the genitive plural as well as in possessive plural forms of all grammatical cases, I will follow Alatalo in referring to it as the "plural oblique stem".

Another canonical word structure innovated by Samoyed is *(C)V, which emerged through the loss of intervocalic *γ or *k, e. g. PU *kokį- > PS *ko- 'find', *piγi- > *pi- 'cook' (intr.). The latter PU reconstruction is not supported by Aikio (2022), but it is clear that PU patterns *(C)Vji or *(C)Vwi do not belong here, as the glides in question are invariably

preserved in Proto-Samoyed as seen below. In any case, monosyllabic vowel stems were not common in PS but a few existed. Among nouns, notable examples include PS *pi 'night' > TE *pi* ~ Yr ‹pi› ~ TN *pyi* ~ Sk **pi** ~ Km *pĭ* ~ Mt ***hi*** (Ng *hii* 'night' represents a derivative of *pi, apparently < PS *pi-jə > TN *pyiy*° 'nightly'), PS *nä[1] [SW *ne] 'woman' > Ng *ŋi̯* ~ TE *ne* ~ TN *nye* ~ Sk **ne** ~ Km *ne* ~ Mt ***nä*** (under the current understanding Ng alone reflects *ne while the others rather derive from *nä, yet the correspondence could also be regular in this phonotactically unique position), PS *ńa [SW *ńä] 'friend' > Ng *ńa* ~ TN *ŋya* ~ Sk **ńā** ~ Mt ***ä*** (the inclusion of the Mator cognate presupposes a reanalysis of the initial nasal as the genitive marker). Among verbs, uncontroversial examples include PS *ju- 'melt, warm up' > TE *d'u-* ~ TN *yu-* ~ Sk **ću-** ~ Km *ťü-* with polysemy rather than homonymy, cf. Sk (№ 1427) **ću-** 'melt' and (№ 1429) **ću-** 'warm up' (with a cross reference), cf. also Ng *d'utü-* 'melt' [both transitive and intransitive] < PS *jutå- 'melt' [transitive] > TN *yuda-* id., PS *pi- 'cook' [intr.] > TE *pi-* ~ TN *pyi-* ~ Km *pĭ-* (cf. Ng *hiri-* ~ Sk **pirə-** < PS *pirä- 'cook' [tr.] > TN *pyirye-* id.), PS *ko- 'find' > TE *kô-* ~ TN *xo-* ~ Sk **qo-** ~ Km *ku-* ~ Mt ***ko-*** 'see, find'.

Interestingly, Samoyed stems with the structure *(C)VV, i. e. involving a vowel sequence, appear to be more common than those with *(C)V, yet few can be traced to Proto-Uralic. Among them the most celebrated cases would be the verb 'die', as in PS *kåə̂- [SW 56] 'die' > Ng *kuo-* ~ TE *kaa-* ~ TN *xa-* ~ Sk (№ 1759) **qū-** ~ Km *ku-* ~ Mt ***kā-***, yet the recent PU reconstruction *kali- by Aikio (2022) does not seem to fit; cf. Janhunen (1981) and Sammallahti (1988) where PU *γ is, in my view correctly, suggested. Another widespread word with a vowel sequence ending in a reduced vowel would be PS *ńuə̂ [SW 111 *ńuə̂ (? *ńuə̂j)] '(own) child' > Ng *ńuə* ~ TE *nio* ~ TN *nyú* ~ Sk **ńū** 'small' ~ Mt (№ 779) ***ńuh*** [? = *ńuγ*]. As mentioned in section 2, several other words in Tundra Enets, Yurats, and Mator in particular suggest that the second element was a low vowel, cf. PS *tua [SW *tuə̂j, MS *tua] 'feather, wing' > Ng *t'üə* ~ TE *tua* ~ FE *tue* ~ Yr ‹túa› ~ TN *to* ~ FN *tŏ* ~ Sk **tū** ~ Mt (№ 1070) ***tua***, ***tuga*** (? ***toga***) [? = *tuγa*]. In my view, words such as PS *tęa [SW 155 *t[1]ęə̂, MS *cęз] 'domestic reindeer' > Ng *taa* ~ TE *tia* ~ TN *ti* : *te-* ~ Mt (№ 1001) ***tęgä*** (? ***tę̄gä***) [? = *tęγa*] and PS *lęa [SW 82 *lę, MS id.] 'bone' > Ng *laa* 'ring' ~ TN *li* : *le-* ~ Sk **lę**, nom **lį̄** ~ Km *le* ~ Mt 618 ***lE*** (? ***lę̄***), would also represent the pattern *(C)Va, although details are yet to be worked out. The Nenets vowel alternation of the type TN *ti* : *te- and li* : *le-* is conditioned by syllable structure so that the high vowels only occur in monosyllabic word forms. As for the latter word, Tundra Enets has *laa* 'ring' without a cognate in Forest Enets, so that, although inter-Samoyed loans have occasionally been used as an *ad hoc* solution to explain apparent irregularities, here the distribution, the phonological shape and the meaning of the Tundra Enets word all point to a borrowing from Nganasan. Otherwise, Enets only has reflexes of a correlative derivative, cf. FE *liδi* 'bone' ~ Ng *latəə* id. ~ TN *ledi*° 'spine, skeleton'.

There are also several examples of monosyllabic glide stems, i. e. with the structure *(C)Vj or *(C)Vw. Syllable-final glides are mainly preserved in Nganasan and Mator, cf. PS *åj [SW 17 *åj (? *åə̂j)] 'foot, leg' > Ng *ŋoj* ~ TE *ŋa* ~ FE *ŋo* ~ TN *ŋæ* ~ Mt ***aj*** (: 3s ***ašta*** via a regular sound change *jt > *št*). Lack of cognates in these languages may therefore make

vowel sequences difficult to distinguish from vowel-glide sequences discussed above, cf. PU *uji- > PS *uj- [SW 29 *u-] which only has attested reflexes in Forest Nenets and Selkup, cf. FN *ŋu-* (< PN *ŋú-) ~ Sk (№ 4) **ū-** 'swim', yet the vowel length (PN *ú ~ Sk **ū**) is enough to show that the PS reconstruction in SW and the consequent PU reconstruction *uγi- in Sammallahti (1988: 536 PU *uxɨ-) are inaccurate. Monosyllabic *w stems are less common than *j stems, but a rather ingenious Uralic etymology was proposed by Aikio (2002), regarded as uncertain by him but in my view entirely reliable, viz. PU *lewi- > PS *jew- > Sk (№ 1434) **ćŭ-** 'shoot'. This word is not in SW because the verb as such is only found in Selkup, although Alatalo connects it with (the initial part of) the TN verb *yenyer-* id. In any case, this comparison shows that syllable-final *w is not lost in Samoyed as suggested by Aikio.

There are also genuine examples of stems with a glide following a vowel sequence, mainly evident in Nganasan material, cf. PS *juâj > Ng *juəj* 'fish dam' ~ TN *yú* ~ Mt (№ 234) ***čuj***, or PS (?) *wuaj [SW 177 *woâj] 'island' > Ng *ŋüaj* ~ TE *ńue* ~ TN *ŋo* ~ Sk (№ 1743) **ko** 'island, mound'. In my opinion (see section 2), Ng *śiəd'ə* 'tongue' represents a generalized vowel stem *käəj-ə- of PS *käəj [SW 66 *keâj], cf. TE (?) *sioro* (potentially deriving from an older vowel stem *käəl-ə- in a similar manner) ~ FN *syĕ* ~ Sk (№ 2420) **śē** ~ Mt (№ 450) ***kEj*** (? ***kăj***, ? ***kēj***, ? ***kij***), and *käγli may still be preferred as its Proto-Uralic reconstruction rather than Aikio's *käli. Other examples of similar generalizations of vowel stems in Nganasan and Enets would seem to be PS *n^{1}iəj [SW 102 *n^{1}i (? *n^{1}iâ) (nsm, mt-kg) ~ *ji (? *jiw) (sk, km-kb)] > Ng (?) *ŋiəd'ə* ~ *ńiəd'ə* ~ TE *niojo* ~ FE *ńiejo* ~ TN *ŋyí* ~ Sk **ćŭ** ~ Km *t'i* ~ Mt ***Ni*** [!] 'belt', and PS *tüəj [SW 167 *tüâj] > Ng *t'iid'ə* ~ TE *tiojo* ~ FE *t'iejo* ~ TN *tyú* ~ Sk (derivative) **tŭnaŋ** ~ **tünaŋ** ~ Km *tü* 'sleeve'.

Most polysyllabic vowel stems end in *â / *ə or *å / *ä, which, with the exception of the front reduced vowel *ə, introduced by Helimski (1993), were the only options in the SW framework. While many stems ending in *â in SW are now understood as having had a high vowel in its stead, *â stems remain numerous in current reconstructions, e. g. PS *mi̯tâ [SW 93] 'liver' > Ng *mitə* ~ *muδo* ~ TN *mid°* ~ Sk (№ 757) **mi̯tə** ~ Km *mĭt*, or *čürə [SW 34 *cürâ] 'staff' > Ng *t'irə* ~ TE *t'uro* ~ TN *tyur°* 'driving pole' ~ Sk (№ 1291) **čŭrə**, where the presence of the final *ə is regulated by vowel harmony.

Proto-Samoyed *å stems and *ä stems mostly continue to be reconstructed in the same manner as in SW, e. g. PS *kâtå [SW 55] 'fingernail' > Ng *kətu* ~ TE *koδa* ~ TN *xəda* ~ Sk (№ 1876) **qatə** 'fingernail, hoof' ~ Km *kăda* ~ Mt (№ 377) ***kada***, or PS *petä [SW 126 *pitä] 'nest' > Ng *hi̯ti̯* ~ TE *piδe* ~ FE (*piδe* >) *piδi* ~ TN *pyidya* ~ Sk (№ 463) **pitə** ~ Km *pide* ~ Mt (№ 294) ***hide***, so that they are largely unproblematic both diachronically and synchronically. It is important to notice, however, that, contrary to Janhunen (1998: 468), stem-final *a and *ä did not alternate morphophonologically with the reduced vowel in Proto-Samoyed. In the case of nouns, Janhunen assumed that accusative plurals such as TN *tyon°* of *tyonya* 'fox' would derive from PS **töntə, yet FN acc.pl *tyony°* already points to the correct reconstruction PS *tönti. Similarly, TN acc.pl *xob°* of *xoba* 'fur' < PS *kopu, cf. TE *kôbuδa* ~ TN *xob°da* 'her/his furs' < PS *koputå, demonstrating that the final vowels of the Nenets plural oblique stem derive from PS high vowels, which regularly

merged with PN *ə as described above. In the case of verbs, what Janhunen analyzed as a complex suffix sequence as in PS **kåjə-jəj- 'remain', a stative verb derived from PS *kåjä- 'leave', falls under current *w stems instead (Gusev 2008), PS *kåjaw- being the apparent reconstruction for 'remain', cf. also PS *kontaw- [SW 73 *kontə-jəj-] > Ng *kundå-* (~ TN *xonyo-*) 'sleep' where Nganasan weak grade *nd* shows that the second syllable must have been closed in Proto-Samoyed. From a comparative point of view, the elements **kåjə- and **kontə- in SW are products of a particular, now obsolete, reconstruction technique, and even in that framework they were not supposed to occur independently.

A small but significant corpus of Proto-Samoyed *a stems, obviously missing in SW since no PS *a was recognized at the time, may be established in words such as PS *wota [SW 177 *wotɜ] 'berry' > Ng *ŋuta* ~ TE *ôδe* ~ Yr ‹ngóde› ≈ ***ŋoδä*** ~ TN *ŋodya* ~ Sk (№ 1917) **kotə** 'bog bilberry' as well as PS *kɨ̯ta 'cup' > Ng *kita* ~ TE *kiδe* ~ TN *xidya*, resulting in disharmonic stems in Nenets. On the basis of Nganasan evidence, the historically important word PS *wäsa [SW 175 *wesä] 'iron' > Ng *basa* ~ TE *bese* ~ Yr ‹jésse› ~ TN *yesya* ~ Sk (№ 2185) **kuəsə** would belong to PS *a stems as well.

Polysyllabic vowel stems ending in a high vowel, however, present a more complex picture. In the case of illabial high vowels, there appears little doubt that they existed in word-final position, e. g. PS *jänti [SW 43 *jentə] > Ng *d'enti* ~ TE *d'edi* ~ TN *yen°* ~ FN *jeny°* (< PN *jennyə) ~ Sk **ćinti** 'bowstring', where the original *i is evident even in the palatalization of the Forest Nenets consonant (Salminen 2007). In the case of verbs, Enets data is particularly relevant, because both the imperfective participle suffix *-ntå and the future marker *-ntə continue to have two allomorphs postvocalically, viz. TE *-da* and *-do* after an original vowel but *-δa* and *-δo* when a PS stem-final glide is to be reconstructed, deriving from truncated PS allomorphs *-tå and *-tə, respectively, cf. TE *beri-* ~ TN *yerə-* ~ FN *wyełyă-* (< PN *wyeryə-) < PS *wäri- 'rip', with TE future stem *beridо-* pointing to an original vowel stem, as opposed to a glide stem such as TE *oδi-* ~ TN *ŋədyə-* ~ FN *ŋătyă-* (< PN *ŋətyə-) < PS *ətiw- 'be visible', with TE future stem *oδiδo-* marking the distinction. When it comes to polysyllabic vowel stems ending in a high labial vowel, however, a question arises whether most or all of those stems having a final *u* in Enets and Selkup are better analyzed as *əw stems, as discussed below.

Besides the morphophonological processes connected with consonant stems in general presented above, it may be added that in Selkup the so-called aorist marker PS *-ŋå appears in the secondary form **-nə** after a consonant.

Proto-Samoyed *m stems are generally highly stable across the languages. PS *-ŋå results in *-ma* in Enets and Nenets, e. g. TE *omaδa* ~ TN *ŋəmada* < PS *əm-ŋå-tå > Sk **amnətə** '(s)he ate it'.

In Nenets, PS *n *ń *ŋ stem nouns are only distinguishable on the basis of prevocalic allomorphs, while in Nganasan, final *n > *ŋ*, *ń > *j*, cf. PS *węn [SW 173] 'dog' : nom.pl *węn-ə-t : acc.pl *węn-ə-j > TN *weh* (~ Ng *baŋ*) : *wen°q* (~ Ng *banəq*) : *weno*, PS *poəń 'gap' : *poəń-ə-t : *poəń-ə-j > TN *poh* (~ Ng *huoj* 'border, juncture') : *poy°q* (~ Ng *huońəq*) : *poyo*, PS *aŋ [SW 20 *äŋ] 'mouth' : *aŋ-ə-t : *aŋ-ə-j > TN *nyah* (~ Ng *ŋaŋ*) : *nya°q* (~ Ng *ŋaŋəq*) : *nyaŋo*, etc. A remarkable feature of Enets is that PS *n *ń *ŋ stem

verbs have analogically shifted into *r stems, e. g. TE *piśir-* ‘laugh’, cf., for instance, FN *pyĭsyăn-* < PS *pisiń- id. In Nenets, by contrast, PS *n *ŋ stem verbs have adopted the pattern of PS *ń stems so that, for instance, the connegative of PS *men- [SW 94 *min-] ‘go’ in Forest Nenets is *myĭj°q* instead of **myĭn°q* that would presumably be the regular reflex of PS *men-ə-k, cf. FN *pyĭsyăj°q* < PS *pisiń-ə-k.

Proto-Samoyed *p stems were few, and all known examples are nouns, one of them, however, of high frequency, i. e. TN *ŋob* ‘one’ < PS *op [SW 28 *op (? *oâp)].

Proto-Samoyed *r stems are numerous, especially in the case of frequentative *r stem verbs, while *l stems appear rare or secondary because of the syllable-final sound change PU *l > PS *j.

In the same way as uncovering the correct nasal in PS *n *ń *ŋ stems in northern Samoyed, PS *t *č *s *k stems often reveal the actual obstruent through prevocalic allomorphs, cf. PS *mat [SW 90 *mät] ‘tent’ : nom.pl *mat-ə̂-t : acc.pl *mat-ə̂-j > TN *myaq* : *myad°q* : *myado*, PS *wə̂js [SW 169] ‘half’ : *wə̂js-ə̂-t : *wə̂js-ə̂-j > TN *wœq* : *wœs°q* : *wœso*, while *t (and *č) stem verbs have largely shifted to *s* stems, e. g. PS *jämpit- [SW 88 *jempə̂t- ? ~ *jempə̂jt-] > TN *yempəq-* ‘dress’ (~ FN *jempyăq-* < PN **yempyəq-*) ~ Ng *d'embiq-* while the respective connegatives TN *yempəs°q* ~ Ng *d'eŋhid'əq* would seem to reflect PS **jämpis-ə-k, yet the correct reconstruction *jämpit-ə-k is revealed by internal reconstruction on the basis of derivatives such as PS *jämpitäw- > TN *yemp°dyo-* ‘be dressed’.

Original *k stems were few, and in Enets and Nenets, *k stem nouns have often shifted to *kə̂ stems analogically, e. g. PS *čuk [SW 34 *cuk (~ *cukə̂)] > TE *tubo* ~ FE *tuxu* ~ TN *tux°* ~ Sk (№ 1188) **čuk** ‘maggot’, while *k stem verbs have followed *t stem verbs in becoming *s* stems, e. g. *jok- [SW 46 *jokə̂- (? *jok-)] > Ng *d'uk-* ~ TE *d'ôq-* ~ TN *yoq-* ‘lose’.

This leaves us with (polysyllabic) glide stems, of which those with PS final *j were ubiquitous in SW, while those with final *w were introduced by Gusev (2008). Without aiming at comprehensiveness, the following makeshift stem types need to be discussed in this context. A curious fact seems to be that no *ə̂j stems, while numerous in SW, can actually be reconstructed in Proto-Samoyed, even though there appears to be no rationale for such a lacuna. This can still be seen in modern Nganasan, where no stems of the pattern CV(C)Cəj are attested, which can only be due to a lack of a historical source for them, because phonotactically they would be fine, cf. accusative plural forms such as PS *mat-ə̂-j > *matəj* ‘tents’, and final *əj* is also found in the dual marker *-gəj* (< PS *-kə̂ń), plus there are many other stems ending in *j* preceded by other vowels as seen below. In any case, final *j is preserved in Nganasan and Mator, while the same languages have a segmental reflex of final *w, viz. old Nganasan (Castrén) ‹ŋ›, which has since been lost but its effect can be seen through gradation, and Mator ***h*** [? = *y*].

It would also seem that there was no contrast between non-initial *ə̂w and *uw sequences in Proto-Samoyed, but I use *ə̂w here as the more conventional option. In any case, the most celebrated case of a Proto-Samoyed *ə̂w stems would probably be PS *kə̂čə̂w [SW 57 *kə̂cə̂] ‘blizzard’ > Ng (Castrén ‹koduŋ› >) *koδu* ~ TE *kaδu* ~ FE *kaδu* ~ Yr ‹cháda› ≈ ***kaδə*** ~ TN *xad°* ~ FN *kat°* ‘snowfall’ ~ Sk **qoču** ~ Km (derivative) *kadaŋ* ‘freshly fallen snow (in autumn)’ ~ Mt ***kaduh***, especially because Nganasan gradation

immediately points to a final consonant. The final consonant is also reflected in Nganasan inflection, cf. pl *kotuq* < PS **kåčâw-ə̂-t* or 3sg *kotutu* with the allomorph *-tu* (rather than *-δu* found in corresponding vowel stems). Compare also the Nenets reflexes of the translative derivative PS **kåčâw(ə̂)jm-* > TN *xadum-* ~ FN *kitum-* < PN **katum-* where the second-syllable high vowel excludes the possibility of PS ***kåčəjm-* which would have resulted in PN **katom-* instead. Another widely attested word belonging here would be PS **jåmpəw* [SW 37 **jåmpə̂*] 'long' > TE *d'abu* ~ TN *yamp°* ~ Sk **ćumpu** ~ Km (Castrén ‹numu› >) *numo* ~ Mt ***ńambuh***, although it has no cognate in Nganasan. The reflexes of the word for 'north' (in Nganasan 'north-east wind') [SW 22 **ęrmə̂*] > Ng *ŋarmi* ~ *ŋarmü* ~ TE *umu* ~ TN *ŋerm°* ~ Sk ***ęrma** are more problematic, because the Selkup word appears to be based on correlative derivation, and the 3sg form in modern Nganasan is actually recorded with the allomorph *-δü*, which would reflect an original vowel stem, although it may well be an analogical creation.

Two ancient kinship terms are the best candidates for Proto-Samoyed **əw* stems, viz. (PU **wäŋiw* >) PS **wiŋəw* [SW 176 **wiŋə̂*] 'son-in-law' > Ng *biŋi* ~ TE *bii* ~ TN *yiy°* ~ Mt ***mijüh***, and PS (PU **käliw* >) **käləw* [SW 67 **kelə̂*] 'brother-in-law' > Ng (Castrén) ‹śaluŋ› ~ TE *seri* ~ TN *syel°* ~ Sk **śelu**, for which the Common Selkup form in Donner, Sirelius & Alatalo (2004 № 3017b) was ***śelə** based on Taz and Narym data alone, but Ket *sellu* (Alatalo 1998) shows the original final vowel. Again, a potentially analogical 3sg form *biŋiδi* has been recorded for Nganasan.

There is no reason why stem-final **w* could not have been preceded by a high vowel, and the reconstruction PS **ətiw-* [SW 16 der. **ə̂tə̂-- (? *ə̂tə̂jɜ-)*] 'be visible' would account for Ng *ŋəδu-* ~ TE *oδi-* ~ TN *ŋədyə-* ~ Mt (№ 7) [?] ***adə-*** [? = *adi-* on the basis of the original record]. Morphophonological features including Nganasan gerund in *-sa* and Tundra Nenets participle *ŋədy°da* (pro **ŋədy°na*) also point to a final consonant.

Proto-Samoyed **åw* stems can also be demonstrated with a Proto-Uralic kinship term, i. e. (PU **nataw* >) PS **nåtåw* [SW 98 **nåt¹ə̂- (? *nåt¹əj)*] 'brother-in-law' > Ng *noδu* ~ TN *nado* ~ Km *nado*, the Nganasan cognate having only been discovered after SW. A verb showing the same correspondences would be PS **t¹åŋåw-* [SW 151 **tåŋə̂-j(-)*] 'be worn out' > TN *taŋo-* ~ Km *taŋo-*, a stative derivative of PS **t¹åŋå-* [SW 151] 'wear out'.

Proto-Samoyed **aw* stems are represented by several nouns of the type PS **śåraw* [SW 135–136 **śårə̂-jə̂j*] > Ng (Castrén ‹soruaŋ› >) *sorå* ~ TN *saryo* 'rain', derived from PS **śårå-* [SW 135] 'rain', as well as numerous stative verbs such as PS **kə̂ptaw-* [SW 54 **kə̂ptə̂-jə̂j-*] 'be extinguished, go out' > Ng *kəbtå-* ~ TE *koti-* ~ TN *xəbtyo-* ~ Sk **qapti-** ~ Km *ku'bdo-*, derived from PS **kə̂ptå-* [SW 54] 'extinguish, put out'. Again, morphophonological features such as Nganasan gerund in *-sa* and Tundra Nenets future in *-də-* point to a final consonant. Compare also **åmtaw-* [SW 18 **åmtə̂-jə̂j-*] 'sit' > Ng *ŋomtü-* ~ TE *adu-* ~ TN *ŋamtyo-*, as well as (with an irregular vowel correspondence in the first syllable) Sk (№ 49) **āmti-**, derived from a reflexive verb **åmtə̂-* 'sit down'.

Proto-Samoyed **äw* stems may be demonstrated by PS **wätäw* [SW 175 **wetə̂-jə̂j*] 'bowel' > Ng (Castrén ‹beatuŋ› >) *bətu* ~ TE *beδe* ~ Yr ‹bédu› ≈ ***beδu*** ~ TN *yedyo* ~ Sk **ketu** ~ Km *bedö* ~ Mt ***bedüh*** ~ ***bedöh***, on the condition that the strong grade in Ng *bətu* rep-

resents analogical levelling on the basis of the (in the case of this word very frequent) plural form *bətuq* < PS *wätäw-ə-t.

As seen from the SW reconstructions above, Janhunen analyzed *w stems with a final low vowel as (complex) *j stems, while Helimski in MS went to the opposite direction by reconstructing final *o *ö where Nenets has *o yo* (and Kamas actually *o ö*), regarding Mator ***h*** as the result of "h-Erweiterung", i. e. a secondary addition after labial vowels (Gusev 2008). It is not obvious to what the *aw stems owe their final vowel, but a similar vowel change is also found before *j, cf. PS *utå [SW 30] 'hand' > TN *ŋuda* : acc.pl *ŋudyi* < PS *uta-j, here obviously more easily explained as simple assimilation.

The inflection of *w stem verbs in Nenets indicates a rather unusual contraction insofar as we assume that they too originally had the so-called aorist marker PS *-ŋå attached to the stem, cf. PS *kontaw-ŋå > PN *koni > TN *xoni* ~ FN *kuni* '(s)he sleeps'. Regular sources for PN *-i would be PS *-iâ or even *-iŋâ, e. g. TE *kaδio* ~ TN *xadi* < *kåtiâ 'spruce', so *-awŋå > *-i obviously seems a bit of a stretch. Enets does not shed light on the matter, cf. TN *xonyo-* : 3sg *xoni* ~ TE *kôdu-* : *kôdua*, where *kôdua* looks like a plausible reflex of PS *kontaw-ŋå. The connegative of *w stem verbs in Nenets does not seem regular either, as it has a second-syllable high vowel instead of the expected mid vowel, e. g. PS *kontaw-k > PN *konyuq > TN *xonyuq* ~ FN *kunyuq* (cf. Ng *kundåq* ~ TE *kôduq* which would derive from the PS form regularly).

While many polysyllabic words reconstructed with final *j in SW are now better understood as *w stems, several examples of *j stems remain, cf. *čamčåj [SW 32 *cämcâjâ] 'frog' > Sk (№ 1018) **čāmčē** which is cognate to the root of TN *tyamteq* id. (< PN *tyamtæq). In Tundra Nenets, original *j stems are often reflected as analogically extended forms, cf. TN *sirey°* (pro **sire*) ~ TE *sire* 'two-year old reindeer cow' < PS *sị̈råj as opposed to PS *sị̈råjâ > *sirey°* ~ TE *sireo* 'hibernal' (the tentative reconstructions for these words in Salminen 2012 are obviously outdated). The distinction between the Tundra Nenets words is still shown in inflection, cf. the respective prosecutives *sirew°na* < PS *sị̈råj-månå vs. *sirey°wəna* < PS *sị̈råjâ-månå.

Two widespread words need to reconstructed as clumsy-looking Proto-Samoyed *âwj stems in the current framework, viz. PS *jâŋkâwj [SW 35 *jâŋkâj] 'trap' > Ng *d'eŋguj* ~ TE *d'ogu* ~ TN *yəŋko* ~ Sk **ćaŋku** and PS *ântâwj [SW 15 *ântâj] 'boat' > Ng *ŋənduj* ~ TE *odu* ~ Yr ‹ngáddu› [≈ *ŋədu*] ~ TN *ŋəno* ~ Sk **antu** ~ Km *eńi* ~ *äńi* ~ Mt ***ondoj***, because PS final *âj should result in Ng **əj* rather than the attested **uj*, and Nenets also has dialectal accusative plural forms such as FN *ŋănuju* with both a labial vowel and a glide.

4.8 Basic aspects of Samoyed inflectional morphology

4.8.1 Nominal inflection

Virtually all elements of Uralic nominal inflection (Janhunen 1982: 29–33) have been preserved in Samoyed, and the functions of the inherited suffixes have also remained largely intact. The nominative (or absolute) singular form is expectedly unmarked, and the other grammatical cases have simple markers, while the local cases exhibit multiple suffixation, involving so-called co-affixes, in nominal inflection, simple markers being only found in forms of adverbs and postpositions (see section 6).

Nominal forms in Samoyed are generally marked for number, case, and person in this order, unlike in neighbouring Mansi and Khanty branches, where many case suffixes have developed from postpositions and possessive suffixes precede case markers. In Selkup, the latter order is found in the so-called pseudo-cases, notable the comitative in Sk **-sē**, for the same reason.

The Uralic number and case system was asymmetrical, with three grammatical cases in the singular, i. e. the unmarked nominative (or absolute case), the accusative in *-m, and the genitive in *-n, as opposed to a two-way contrast between the plural nominative in *-t and the plural conjunctive (or oblique case) in *-j, which covered both accusative and genitive functions in the absolute (or non-possessive) declension and was also used as the plural marker in the possessive nominative forms. The reconstructions of these four basic suffixes consisting of a single consonant are identical for both Proto-Uralic and Proto-Samoyed, the only notable difference being that the marker *-j became restricted in the plural accusative function in Samoyed, while the genitive plural as attested in northern Samoyed was marked with a double suffix *-j-t. For example, the three grammatical case forms of 'fur' would be PU *kopa : *kopa-m : *kopa-n > PS *kopå : *kopå-m : *kopå-n > TN *xoba* : *xobam* : *xobah* ~ Sk **qopə** : **qopəm** : **qopən**, with regular cognates in the other Samoyed languages.

The nominative plural can be demonstrated similarly with PU *kopa-t > PS *kopå-t > TN *xobaq* : Sk **qopət**, with the caveat that this suffix and the dual suffix discussed below have been replaced with what Alatalo calls collective markers in southern Selkup, where the suffix is **-lā**, and with the suffix *-sAŋ* in Kamas. The plural forms of the local cases do not constitute a uniform system among the Samoyed languages, pointing to their late origin. The other plural marker *-j has been preserved as such only in Nganasan, e. g. PS *kopå-j > Ng *kubuj*, while in Enets and Nenets this suffix like any other postvocalic *j has created a complex array of morphological alternations referred to as the plural oblique stem. In Selkup and Kamas, by contrast, PU *-j has by and large been eliminated in favour of agglutinative structures based on analogical levelling, cf. Sk nom.pl **qopət** : acc.pl **qopətəm** : gen.pl **qopətən** vs. Ng *kuhu* : *kubuj* : *kubaq* (< PS *kopå-j-t).

The basic dual marker in Proto-Samoyed was probably a single suffix *-kâ(ń), where (most dialects of) Selkup shows no final consonant but the northern Samoyed languages suggest the longer variant, for which wider Uralic cognates may also be found, cf. nom.du Ng *kəntəgəj* ~ TE *kodoxoq* ~ TN *xən°xəh* < PS *kânčâ-kâń vs. Sk **qančōqə** < Pre-Selkup *kânčâ-kâ of PS *kânčâ 'sledge' (> Ng *kəntə* ~ TE *kodo* ~ TN *xən°* ~ Sk **qančə**). In Nenets, the same suffix *-xəh* is used in all grammatical cases, while in Nganasan, the accusative–genitive corresponding to the nominative *kəntəgəj* would be *kəntəgi* instead, and the Enets dual marker is recorded as *-xiq*, the vowel of Ng *-gi* and Enets *-xiq* being apparently secondary. In any case, dual forms of local cases were not present in Proto-Samoyed, postpositional phrases being used in northern Samoyed and secondary agglutinative forms in southern Samoyed.

Besides the three grammatical cases (nominative, accusative, and genitive), there are four local cases (dative, locative, ablative, and prosecutive) in the basic Samoyed case system. While the Uralic background of the dative (PU *-ŋ > PS *-ŋ), the locative (PU *-na > PS *-nå), and the ablative (PU *-ta > PS *-tâ) has been obvious for a long time, perhaps with the exception of the development of the vowel in the ablative marker which seems an outcome of a conditioned sound change, the prosecutive has traditionally been regarded as a uniquely Samoyed innovation. Recently, however, Jalava (2022) has suggested a Uralic postpositional origin for the prosecutive, which seems a highly plausible idea, even if I would probably put the two potential postpositions suggested by her in another order of likelihood.

As already noted, the above-mentioned simple markers for the local cases only appear in forms of adverbs and postpositions. In the case of nouns, i. e. in the vast majority of forms involving local case markers, they are preceded by another element, commonly called a co-affix. There are two such co-affixes, of which *-ntâ is invariably used in the absolute form of the dative case (*-ntâ-ŋ; cf. the status of the possessive forms immediately below) and *-kâ is always used in the ablative (*-kâ-tâ), while in the locative, as discussed earlier, Nganasan uniquely employs the first co-affix *-ntâ-nå (> Ng *-ntənu* with allomorphs) as opposed to *-kâ-nå in the other languages (> TN *-xəna* etc.). Plural markers follow the co-affix, e. g. Pre-Nenets loc.pl *-kâ-t-nå (> TN *-xəqna* etc.). Yet here again there is a difference within northern Samoyed, in this case the isogloss uniting Enets and Nganasan, which use the other plural marker *-j where Nenets has the above-mentioned *-t, cf. loc.pl of 'fur' Ng *kubutinü* (< *kopå-ntâ-j-nå) vs. TE *kôbaxine* (< *kopå-kâ-j-nå) vs. TN *xobaxəqna* (< *kopå-kâ-t-nå).

Local case suffixes did not combine with the dual marker in Proto-Samoyed either, which is expected given the late origin of the plural local case forms. The attested northern Samoyed languages, however, preserve the situation by employing postpositional phrases in this function, for instance, TN *nyax°yuta nyad°* involving gen.du3sg of *nya* 'friend' (< *ńa) and abl. of (the etymologically unrelated postposition) *nya-* (< *na) fulfils the functions of both case forms (TN sg *nyaxəd°* < *na-kâ-tâ : pl *nyaxət°* < *na-kâ-t-tâ) and and postpositional phrases (TN sg *nyanta nyad°* : pl *nyíta nyad°*) in the other numbers. In Selkup, secondary agglutinative structures involving the dual marker, similar to the

Sk nom.pl **qopət** : acc.pl **qopətəm** : gen.pl **qopətən** mentioned above, are available, cf. the so-called pseudo-case comitative, in Taz Selkup sg *qopəsē* : du *qopōqəsē* : pl *qopəssē* (where *ss* < *-t-s).

A related difference between Nganasan and the other Samoyed languages occurs in the possessive forms of the dative case, where Nganasan expectedly employs *-ntâ as in Ng *kəδutətu* dat.sg of 'fingernail' < Pre-Nganasan *kâtå-ntâ-ntå, while Enets, Nenets, and Selkup exhibit a different suffix PS *-kâ, obviously identical with the above-mentioned co-affix, as in TE *koδaxoda* ~ TN *xədaxənta* ~ Sk **qatōqəntə** < PS *kâtå-kâ-ntå id.

The system of possessive forms in Samoyed is rich yet highly regular. Originally agglutinated from personal pronouns, the possessive suffixes have developed from Proto-Uralic to Proto-Samoyed in a direct way. They form several series, already present in PU, cf. PS nom.sg 1sg *-mâ : 2sg *-râ : 3sg *-tå as opposed to acc.sg 1sg *-mâ : 2sg *-m-tâ : 3sg *-m-tå and gen.sg 1sg *-nâ : 2sg *-n-tâ : 3sg *-n-tå, the genitive series being also used in the possessive inflection of the local cases. The *r in the nom. 2sg *-râ goes back to Uralic *t, which may be another potential example of early rhythmic, or suffixal, gradation (see section 5). The non-agglutinative character of the acc. 1sg *-mâ and the gen. 1sg *-nâ, by contrast, exhibits a rare instance of reconstructable morphophonology already present in Proto-Uralic. At least in Nenets, there are further two series of possessive forms for non-singular forms, one for the nominative and accusative, and another for the genitive and (in plural) for the local cases. The full list of PS nominative possessive suffixes would be 1sg *-mâ : 2sg *-râ : 3sg *-tå : 1du *-miŋ : 2du *-riŋ : 3du *-tiŋ : 1pl *-måt : 2pl *-råt : 3pl (?) *-tåwn (cf. Alatalo 1999: 16), of which all but the 3pl (?) *-tåwn appear as expected on the basis of what we know of the other Uralic languages. The five series of possessive suffixes, evident in Nenets, can be illustrated by TN 2sg nom.sg *xədar°* : acc.sg *xədamt°* : loc.sg *xədaxənant°* : nom.–acc.pl *xədyid°* : loc.pl *xədaxəqnat°* which would mechanically derive from Pre-Nenets *kâtå-râ : *kâtå-m-tâ : *kâtå-kâ-nå-n-tâ : *kâta-j-tâ : *kâtå-kâ-t-nå-t-tâ, but, as mentioned above, Enets and Nganasan use the plural marker *-j instead of the Nenets *-t in the inflection of the local cases, indicating that these forms became established only after Proto-Samoyed.

With few relics in southern Samoyed, destinative, with the functions of 'for' and '(becoming) to' constitutes a distinct category in the northern Samoyed languages, where it undoubtedly represents a retention. The PS destinative marker is *-tâ, which is still sometimes confused with the 3sg possessive suffix *-tå, although they have nothing in common either synchronically or diachronically. In fact, the apparent PU origin of the destinative marker is the PU translative case suffix *-ksi (for the regularity of the sound changes, cf. PU *suksi̯ > PS *tutâ > Ng *tutə*), as opposed to the 3sg *-tå which derives from PU *-sa (for details, see Salminen 2014; cf. also Ylikoski 2017 and the literature therein). Starting as a case suffix, the destinative marker has expanded by being able to attach the three series of possessive suffixes mentioned in the previous paragraph, so that there are distinct forms of nominative, accusative, and genitive destinative in different syntactic functions. It seems that destinative can be combined with the category of number

in Enets and Nganasan, while no plural destinative forms have been reliably attested in Nenets (Burkova 2010).

For the predicative forms of nouns, also known as the nominal conjugation, see the following section.

4.8.2 Finite verbal inflection

The bulk of Samoyed verbal inflection is also inherited from Proto-Uralic, yet there appear a few instances where detecting cognates for verbal suffixes in the other Uralic languages is not obvious. A feature shared with the other Siberian Uralic branches, i. e. Khanty and Mansi, and to a slightly lesser degree with Hungarian, is the contrast between subjective vs. objective conjugations in the inflection of transitive verbs. In these languages, the subjective forms are used when the object represents new information, or, in other words, has focus, as opposed to the objective forms used when the object is known from the previous context and thus lacks focus. In the latter case, the object can also be omitted in the sentence, which is not possible if a subjective form of the verb is present (see section 10). The objective forms of transitive verbs, in a system preserved in northern Samoyed, obligatorily mark the number of the object. Furthermore, intransitive verbs in the broad sense are divided into intransitive verbs in the narrow sense versus reflexive verbs as described below, with marked morphological differences. At least in Nenets, finite verb forms exhibit four series of personal suffixes, illustrated by TN 1sg subj. *podyerŋad°m* 'I harnessed' (requiring an immediately preceding object noun phrase as explained above) : obj.sg *podyerŋaw°* 'I harnessed it' : obj.pl *podyerÿən°* 'I harnessed them [pl]' : refl. (semantically unusual but not entirely impossible) *podyerÿəw°q* 'I got harnessed' reflecting the mechanical but plausible PS reconstructions *potar-ŋå-tə̑m : *potar-ŋå-mə̑ : *potar-jə̑-nə̑ : *potar-jə̑-mə̑t, respectively. The corresponding dual object forms share the suffix series with obj.pl, e. g. TN obj.du *podyerŋax°yun°* 'I harnessed them [du]', but the marker of the dual object (PN *-kəjo) is not directly reconstructable on the Proto-Samoyed level.

The Proto-Samoyed personal suffixes in the subjective conjugation, viz. 1sg *(-tə̑)-m (with other poorly understood variants in several Samoyed languages) : 2sg *-ntə̑ : 3sg ∅ : 1du *-niŋ : 2du *-tiŋ : 3du *-kə̑(ń) : 1pl *-mə̑t : 2pl *-tə̑t : 3pl *-t, show that in the third person they are identical with plain numerus markers of nouns as explained above. Since personal suffixes are not usually attached directly to verbal stems, the morphological patterns remain distinct, cf. FN *pəraŋaq* 'they [pl] burn' < PS *pə̑rå-ŋå-t vs. FN *məraq* 'dunes' < PS *mə̑rå-t. That said, there is an unusual FN 1sg personal suffix *-m°q* which does attach directly to a verbal stem, e. g. *tom°q* 'I came', which, however, is fully equivalent to the more regular from a comparative point of view form *toŋat°(m)* < PS *toj-ŋå-tə̑-m.

Most reflexes of the Proto-Samoyed personal suffixes appear as expected in the attested languages, bar a few instances of analogical levelling such as the replacement

of the 2sg *-ntə̂ with *-l* in Kamas, taken over from the objective paradigm. The 2sg suffix in Nganasan is *-ŋ*, which is not the regular reflex of *-ntə̂, yet, as pointed out by Alatalo, it can be explained as an idiosyncratic development in a frequently occurring suffix. The 3pl suffix in Selkup is *-tət* both in subjective and objective conjugations, and while it obviously does not reflect either PS *-t or (?) *-tåwn, it can be tentatively explained as a secondary combination of two suffixes. The most problematic person is, as indicated above, the 1sg, whose suffix is reconstructed as *(-tə̂)-m, yet only the Siberian dialects of Tundra Nenets, Forest Nenets (with the reservation mentioned above), and both Enets languages reflect PS *-tə̂-m. In Nganasan and Kamas, the suffix is simply *-m*, while the European dialects of Tundra Nenets use *-məh* deriving from an earlier suffix complex *-m-məh, the latter part of which may be identified with the 1sg personal pronoun. Furthermore, the 1sg subjective suffix in Selkup is *-ŋ*, with no apparent connection to PS *(-tə̂)-m.

The 1du above suffix is reconstructed above with *n while the respective 1pl suffix has *m, which reflects the Nenets state of affairs, i. e. only 1pl of the subjective conjugation has a suffix identical to the corresponding suffix in the objective conjugation (cf. Trosterud 2006). In Nganasan, also the 1du suffix reflects *m, making it indistinguishable from the respective objective counterpart.

This series of personal suffixes is the only one found in the inflection of intransitive verbs, but it can also be attached to nominal stems, in their cases directly without *-ŋå or any other suffix, e. g. FN *katan*° 'you are a grandmother' < PS *kåtå-ntə̂, a structure known either as the nominal conjugation or the predicative forms of nouns.

The suffixes found in the objective conjugation are the very same possessive suffixes of nouns found in the nominative series. Again, unlike the possessive suffixes, they cannot be directly attached to the verbal stem, but a distinct suffix, most commonly PS *-ŋå, must augment the stem, e. g. FN *kata-* 'kill' : 2sg *kataŋał*° 'you [sg] killed it' < PS *kåtå- : *kåtå-ŋå-rə̂ (as opposed to a homonymous nominal stem FN *kata* 'grandmother' : 2sg *katał*° 'your [sg] grandmother' < PS *kåtå- : *kåtå-rə̂). Another series of suffixes identical with the respective possessive series is used in non-singular forms, typically with a different augmenting suffix, e. g. FN *katajăt*° 'you [sg] killed them' < PS *kåtå-jə̂-tə̂.

A third type of conjugation besides subjective and objective, known as the reflexive conjugation, is productive in northern Samoyed but only has uncertain relics in southern Samoyed. Nevertheless, it presumably represents another retention rather than a shared innovation found in Nganasan, Enets, and Nenets. The reflexive conjugation has a particular marker, reconstructable on the basis of Nenets evidence as PS *-jə̂, which is thus identical with the marker used in plural objective forms such as FN *katajăt*° quoted above. The personal suffixes, which constitute a mix of uniquely reflexive ones and those shared with other series, however, differentiate reflexive forms, for instance, a 2nd person form of FN *jamta-* 'get on the way' would be *jamtajăt*° < PS *jåmtå-jə̂-ntə̂. It has been argued that a better term for this category would be "middle" conjugation (cf. Körtvély 2005), and while she has a point, there are also instances of a broadly reflexive meaning such as 'dress oneself' or 'wash oneself'.

The Proto-Samoyed mood and tense system appears to be best preserved in Selkup, while the other languages show numerous innovations in this area. The aorist marker *-ŋå contrasted with the preterite suffix *-så (of Uralic origin), which developed into an interrogative mood marker in Forest Enets and Nenets (see section 3). In Nganasan, the common preterite marker *-suə* derives from PS *-så combined with an adjectival marker *-jə̂. Since the aorist of a perfective verb denotes immediate past, the function of the preterite of a perfective verb would be remote past, cf. Sk **amnām** 'I ate it (right now)' (< PS *ə̂m-ŋå-mə̂) : **apsām** 'I ate it (earlier)' (< PS *ə̂m-så-mə̂).

Besides the imperative, the only old and widespread mood marker would be that of the conjunctive, with **-nē** in Selkup and *-yi* (~ *-nyi*) in Tundra Nenets. In Nenets, the aorist conjunctive expresses a request as in *syiqm° ŋətyeyin°* 'please wait for me' while the preterite conjunctive has a conditional meaning, cf. *syit° ŋətyeyidəmcy°* 'I would wait for you'. There are numerous differences among the Samoyed languages with regard to their mood systems, Tundra Nenets probably coming close to a world record when it comes to the number of moods, cf. Salminen (1998b: 530).

The second person imperative has its own morphological patterns, including *-k in the subjective and *-tə̂ in the objective conjugation with a singular object, cf. TN *xalya ŋəw°q* 'eat the fish' (< PS *kålä ə̂m-ə̂-k) : *ŋəmt°* 'eat it' (< PS *ə̂m-tə̂). The first person imperative, or the (ad)hortative, has the suffix *-ku, and the third person imperative, or the optative, *-ŋə̂, which I had earlier, as pointed out by Alatalo, confused with the aorist marker *-ŋå (Salminen 1997, 1998b). The Nenets suffix combination indicating a non-singular object would be PN *-n-ə-q as in *ŋəwən°q* 'eat them [pl]!', and PN *-t-ə-q in the reflexive conjugation, cf. *sələd°q* 'return!'.

There is even less uniformity among the Samoyed languages when it comes to marking the future tense, which, however, typically does exist as a distinct morphological category, unlike in many other Uralic languages. In Enets and Nenets, the future markers can be formally characterized as derivational rather than inflectional, because they combine with particular inflectional categories such as the connegative in the same manner as other, unquestionably derivational elements, as in TN *tolaŋkuw°* 'I am going to read it' vs. *nyíw° tolaŋkuq* 'I am not going to read it' where the future marker behaves the same way as, for instance, the durative marker, cf. *tolabyiw°* 'I am reading it' vs. *nyíw° tolabyuq* 'I will be reading it', as opposed to a mood marker such as the necessitative *tolabcuw°* 'I shall read it' vs. *nyíbcuw° tolaq* 'I shall not read it'. The morphotactic position of the future marker in Enets and Nenets makes it possible to combine it with the preterite, e. g. TN *tolaŋkuwəsy°* 'I was going to read it' or *nyíwəsy° tolaŋkuq* 'I was not going to read it'. The future in Selkup, by contrast, is expressed by a mood marker **-lə** (whose cognate in Kamas denotes future on its own) and a particle **-si** which is added to the personal suffix, i. e. in the same way as the preterite marker in Enets and Nenets, and Alatalo regards a historical connection possible despite the semantic mismatch. He also discusses the suffix *-pså and other suffixes with initial *ps under the term "future array", whose direct reflexes would be lexicalized such as PS *ə̂m- 'eat' : *ə̂m-så 'food' (> TN *ŋəmca* 'meat'), but both the Nganasan future marker

-(C)*sutə* and the Nenets obligative suffix (TN *-bcaki°*) would seem to be based on such a primary element.

Most of the Samoyed languages also have a formal expression of habitual action. In Nenets the habitual suffix (TN *-syətə*) lies on the border of inflectional and derivational morphology even more so than the future marker discussed above. The two patterns *nyi-sy°ti yeryebyer°q* (the habitual marker attached to the negative auxiliary, which suggests we are dealing with a habitual mood within inflection) and *nyí yeryebyer-cy°tu-q* (the habitual marker in the connegative of the main verb, making it formally a derivational suffix), both mean broadly 'it does not usually happen' although potentially with subtle semantic differences. Nganasan also has a fully productive habitual marker (*-muŋφa* with allomorphs), although it is obviously unrelated to the Nenets suffix.

The expression of evidentiality in Samoyed would require a more detailed discussion, but in Nenets a mood which I have called "narrative" has evidentiality as one of its core functions, cf. TN *xum-pancyiq to-we-waq* 'we came in vain (since it turned out that our expectations were not met)', with the narrative marker TN *-mi°* < PN *-məjə. The narrative marker is historically not identical with the perfective participle, also ending in TN *-mi°*, cf. FN (perfective participle) *tomï* < PN *to-məj vs. (narrative) *tomăj°* < PN *to-məjə, subject to the sound change *əjə > TN *iə* having a neutralizing effect in this case. A formal but functionally perhaps slightly different cognate of Nenets narrative has been called "perfect" by Siegl (2013) and "inferential" by Urmanchieva (2016).

4.8.3 Non-finite verbal inflection

Besides the finite forms dealt with above, the Samoyed languages have a large number of non-finite categories. The gerund (also known as "modal gerund" and rather misleadingly as "infinitive") is the most commonly cited verb form as a head word in dictionaries for most of the Samoyed languages, yet it has no apparent cognate elsewhere in Uralic and, what is more, reconstructing a Proto-Samoyed suffix turns out to be a major challenge. The most basic allomorphs in northern Samoyed seem to be Ng *-sa* vs. Enets (?) *-se vs. PN *-syə, as opposed to Sk **-qo**. In Tundra Enets, polysyllabic vowel stems show an allomorph without a consonant as in *sedoe* of *sedo-* 'be smoked' similar to Tundra Nenets *syamtə°* of the cognate verb *syamtə-* id., while Forest Enets gerunds apparently always have a sibilant as in late FE *säduš* 'to be smoked', Forest Nenets showing variation in this respect. In all of these languages, however, monosyllabic stems preserve the consonant, cf. TE *kôse* ~ TN *xosy°* 'to find' (see section 5 for a potential connection with consonant gradation). Another poorly analyzable form is the supine (also known as "final gerund") which is mainly know from Nenets, with PN suffix *-mənsyə.

The Proto-Samoyed infinitive in *-må, by contrast, derives directly from PU *-ma. It is sometimes called a "verbal noun" instead, which is rather misleading as regards its argument structure. Notably, it governs an accusative object, which makes it undoubtedly part of the non-finite verbal inflection rather than deverbal derivation.

The imperfective participle has the suffix *-ntå is, in my view, cognate with the widely attested Uralic suffix *-ntA, found in similar functions in Saami, Finnic, and Permic (cf. Sammallahti 1998: 86).

The perfective participle appears to have two competing suffixes since the time of Proto-Samoyed, i. e. *-på, usually followed by an adjectival marker *-jə̑, the modern reflexes of which tend to be lexicalized, while the other suffix *-mə̑j generally appears more productive, yet at least in the European dialects of Tundra Nenets both still seem to be in widespread use.

The subordinative in *-put^3 is restricted to northern Samoyed, cf. Ng *kotubütü* ~ TE *kaδabuta* ~ TN *xadab°ta* < PS *kåtå-put^3-tå 'if (s)he killed'. The participles are also used in subordinate constructions, e. g. TN *yadertaxənt°* < PS *jåtåjr-tå-kə̑-n-tå 'when you [sg] walk' (where the first *-tå is the truncated allomorph of imperfective participle marker *-ntå and second one the 3sg possessive suffix).

The connegative in *-k (< PU *-k), i. e. the verbal form that invariably follows a negative auxiliary carrying all the inflectional markers, is another central form in the grammar of the northern Samoyed languages in particular, the negative auxiliary having been lost in Selkup. The connegative, where attested, is always homonymous with the respective 2nd person subjective imperative form, and this seems to have been the case since Proto-Uralic.

4.9 An overview of Samoyed derivational morphology

4.9.1 Preliminary notions

The Samoyed languages are abundantly rich in derivational morphology. Much of it is productive and much of it represents Uralic legacy, yet the topic can only be treated superficially here. Lehtisalo (1936) remains not only the most detailed description of Uralic derivatives but is also heavily based on Samoyed material.

A curious tradition in Uralic historical linguistics, of which Lehtisalo's work is a prime example, is that derivational suffixes, or suffixes in general, are not supposed to contain consonant clusters, in other words, their canonical shapes should be CV or C and all instances of what appear to constitute suffixes of the type CCV need to be explained away by identifying them as compound formations of two originally distinct elements. I would agree that many suffixes containing a consonant cluster are secondary, yet extending that principle to every suffix seems both implausible from a typologically point of view and unfounded by historical evidence. For instance, the PS imperfective participle suffix *-ntå discussed above or what may be called the PS future suffix *-pså discussed below, cannot easily be regarded as secondary compound suffixes.

Janhunen's recent idea that the widely attested Proto-Uralic comparative marker *-mpå, reflected in Samoyed in the form a TN moderative suffix as in *ŋarka* 'big' : *ŋarkampoy°* 'rather big', would consist of the translative suffix *-jm and the perfective participle suffix *-på, ingenious as it is, suffers not only from the above-mentioned tacit assumption but also from morphophonological issues derived from the glide in the translative marker *-jm, cf. TN *yəna* 'slow' : (moderative) *yənampoy°* 'rather slow' : (translative) *yəŋyim-* 'become slow'. Analogical levelling may obviously be a factor here, but from a Samoyed perspective, the suffixes do not seem immediately connected.

4.9.2 Deverbal verbs

The most common feature regulated by deverbal verb derivatives would be the aspectual class, i. e. the relationship between imperfective and perfective verbs. For instance, quoting Tundra Nenets data, when a perfective verb needs to be made imperfective, available derivational means are those known as durative, cf. TN *xada-* 'kill' : *xadabə-* 'be killing', frequentative, cf. *ŋəm-* 'eat' : *ŋəwor-* 'be eating' or *to-* 'come' : *túr-* 'be coming', and simply imperfective, cf. *pyi-* 'cook' [intr.] : *pyinə-* 'be cooking'. On the other hand, an imperfective verb can be made perfective as an inchoative, cf. *nyoda-* 'follow' : *nyodal-* 'start following'. Curiously, however, even the morphologically conservative northern Samoyed languages employ different suffixes for the purpose, cf. FE *kaδa-* 'kill' : *kaδagu-* 'be killing' or Ng *kotu-* : *koδutə-* id. Both the Enets suffix *-gu* and the Nganasan suffix *-tə* have exact cognates in Nenets, yet their functions differ, as TN *-ŋko* (corresponding to the Enets *-gu*) could mark either intransitivity as in *tola-* 'read' : *tolaŋko-* 'do reading' : 3sg *tolaŋku* '(s)he does reading', or future tense as in *tolaŋkuda* '(s)he will read it', while TN *-nə ~ -tə* (corresponding to the Nganasan *-tə* < PS *-ntə̂) appears in a similar function (as in *pyinə-* 'be cooking' above) but not in the case of transitive verbs plus it is used as another future marker in Nenets. What is more, the cognate of the Nenets *-mpə (presumably < PS *-mpiw) in Selkup can be attached to verbs of both aspectual classes, cf. Sk (№ 344 durative of **ilə-** 'live') **iləmpi-** whose expected cognate in Nenets, say, TN ***yilyebə-*, would be fundamentally ungrammatical. Such discrepancies among the Samoyed languages remain unexplained for the time being.

Similarly, the strict subcategorization between transitive and intransitive as well as reflexive verbs typical of the Samoyed languages is overcome by derivation, as in TN intr. *yu-* : tr. *yuda-* 'melt, warm up', intr. *tirabta-* : tr. *tira-* 'dry up', refl. *xona-* 'go to bed' : intr. *xonyo-* 'sleep', or tr. *tola-* 'read' : intr. *tolaŋko-* 'do reading'. In the last two cases, the aspectual class is changed from perfective to imperfective, but these effects operate independently, cf. the stative verbs (*wada-* 'grow' :) *wadyo- and* (*ŋida-* 'hang' :) *ŋidyo-* where the former remains perfective, while the latter becomes imperfective, a distinction based on so-called event classes.

Other deverbal verbs found in Nenets would be iteratives, e. g. TN (*siər-* 'look' : inchoative *si°lə-* 'glance' :) *si°ləŋkə-* 'peep', attenuatives, e. g. (*nú-* 'stand' :) *núy°btye-*

'stand for a while', momentatives, e. g. (*tesə-* 'drip' :) *tes°xəl-* 'drop', and passives, e. g. (*xada-* 'kill' :) *xadara-* [refl.] 'get killed'. For a special study of Nganasan derivational morphology, see Wagner-Nagy (2001).

4.9.3 Denominal verbs

The most notable type of denominal verbs would be translatives ('become something') in PS *-jm, e. g. PS *jalå-jm- [SW 40 *jälə̂m- (~ *jälə̂jm-, *jälə̂jə̂m-)] 'lighten, dawn' > Ng *d'alam-* ~ TE *d'erim-* ~ TN *yalem-* ~ Sk **ćēləm-** 'appear, be born' (from PS *jalä 'light, day' > Ng *d'alị* ~ TE *d'ere* ~ TN *yalya* ~ Sk **ćēlə**). Similar though not identical translatives are found in the other Uralic languages.

Odoratives, expressing the notion 'smell like something', may be reconstructed with the suffix (?) *-ńiw in PS, e. g. PS *kålä-ńiw- 'smell like fish' > TN *xalyayə-* ~ Sk **quə̂ləńi-** (from PS *kålä 'fish' > TN *xalya* ~ Sk **quə̂lə**).

Captatives express the notion of catching or consuming something, with the PS suffix *-jń, e. g. PS *kålå-jń- > TN *xaleh-* 'eat fish' ~ Sk **quə̂ləj-** 'fish'.

Caritives present a complex array of interrelated suffixes, but the most basic PS suffix may have been PS (?) *-kij forming morphological verbs, which, however, are typically recorded as participles, e. g. TN *sœw°syəda* ~ Sk **sajikita** < PS (?) *səjmä-kij-tå 'blind (person)', cf. Mt (№ 1000) ***tergəstä*** [? = *tergištä* on the basis of the original record ‹dergíschtä›] ~ TN *tərcy°da* 'hairless'.

Factitives in the broad sense comprise several derivational processes, but a widespread one would seem to have ended in PS *jt-, e. g. Ng *kəδas-* ~ TE *koδes-* ~ TN *xədyeq-* ~ FN *kătyaq-* (< PN *kətyaq-) ~ Sk **qatət-** < PS (?) *kə̂tajt- 'scratch' (from PS *kə̂tå 'fingernail' > Ng *kətu* ~ TE *koδa* ~ TN *xəda* ~ Sk **qatə** 'fingernail, hoof'). Another PS factitive suffix would have been *-tå, e. g. Ng *ńimti-* ~ TE *ńide-* ~ TN *nyumtye-* ~ *nyimtye-* ~ Sk **nimtə-** < PS *nimtä- 'name' (from PS *nim 'name' > Ng *ńim* ~ TE *ńiq* ~ TN *nyum* ~ *nyim* ~ Sk **nim**).

4.9.4 Deverbal nouns

Instrumental nouns of the type TN *pad°nəbcy°h* 'pen' (< PS *påtə̂-ntə̂-psin) derived from *pad°nə-* 'be writing' are common in all Samoyed languages, and the derivative suffix in the current example can undoubtedly be reconstructed with PS *-psin, yet there appears another variant as in TN *tíncyah* 'lasso' (< PS [?] *tẹa-jń-san) derived from *tíh-* 'catch reindeer' where the suffix would rather derive from PS *-psan. In Nganasan, the alternation between high and low vowels parallels that found in the gerund, yet nothing similar appears in the other Samoyed languages.

In Alatalo's extensive analysis, the instrumental marker belongs to a wider array of suffixes with a broadly future reference, cf. Sk **apsōt** 'food' ~ TN *ŋəmcod°* 'what is to be eaten' < PS *ə̂m-så-w-tə̂ (Salminen 2014: 294).

Locative nouns of the type TE *ôôlaa* ~ TN *ŋəwoləwa* ~ FN *ŋămunłăma* (< PN *ŋəmolləma) 'dining room, canteen' (from PS *âmâjr- 'be eating' > TE *ôôr-* ~ TN *ŋəwor-* ~ FN *ŋămoł-*) are most common in Enets and Nenets. In Nganasan, locative nouns, such as *kəturəmu* 'pasture', show no indication of a consonant cluster, which, however, is evident in Enets and Nenets, which may suggest a PS suffix (?) *rtâmå.

4.9.5 Denominal nouns

Diminutives are used very frequently in all Samoyed languages, but their markers vary significantly. The most common suffixes involve PS *kk, and while many diminutives in the other Uralic languages also exhibit a similar element, the connection is indirect at best, because clusters of PU *k and an obstruent were eliminated in Proto-Samoyed. The most common marker can be illustrated by TN *xalyako* ~ Sk **quâləkka** 'small fish', which at first sight would seem to reflect PS **kålä-kkåw. In Nganasan, however, the plural oblique stem (also known as the third stem) appears before the suffix, as in *kolaqku* (from *kolį* 'fish' : gen.pl *kolaq*), suggesting that the full suffix was rather PS *-jkkåw, so that the correct reconstruction would be *kålå-jkkåw instead. Relics in Nenets confirm this conclusion, cf., for instance, TN *munocyiyeko* < *muntåwtsin-ə-jkkåw 'small beard' (from *munocy°h* 'beard' : gen.pl *munocyiyeq*).

Comitatives or possessives are derivatives not found in Nganasan nor attested in Mator but common and productive in the remaining languages, with the suffix PS *-såmåj > TE *-sae* ~ TN *-sawey°* ~ Sk **-səməl'** ~ Km *-səbi* in the general meaning 'having something', e. g. TN *yurcawey° xalya* ~ Sk **ürsəməl' quâlə** 'oily fish' (cf. TN *yur* ~ Sk **ür** < PS *jür 'fat, oil', hence PS *jür-såmåj 'fatty, oily').

Relationals or connective-reciprocals, with the suffix PU *-kśi > PS *-sâ > Ng *-sə* ~ TE *-so* ~ TN *-sə* ~ Sk **-sə** ~ Km *-sə*, refer to a mutual relationship among people, such as 'sisters' or 'brothers-in-law'. For obvious semantic reasons, the suffix is usually attached to a numerus marker but appears in the singular when the derivative is preceded by a numeral, e. g. TN pl *syeləs°q* 'sisters-in-law' < PS *käläw-sə-t (cf. Finnish *kälykset* id.; Salminen 2014), or TN *syidya nyas°* 'two friends' (cf. TN *nya* 'friend').

Derivatives denoting a property of an adjective such as TN *yampəd°* 'length' (cf. TN *yamp°* 'long') are commonly used in Enets, Nenets, and Selkup, but not in Nganasan or Kamas. The Tundra Enets cognate of TN *yampəd°* is TE *d'aboδi* which suggests PS *jåmpâtį [SW 37 *jåmpâtâ] instead of Janhunen's original reconstruction. The Selkup cognate of TN *pyirəd°* 'height' is, however, Sk (№ 677) **pīrēt** which would seem to support PS (?) *pirətə rather than *pirəti which would be expected on the basis of Enets data. Janhunen (1981: 238) reconstructs PU **piδiw for TN *pyir°* 'height' (~ Ng *hirə* ~ TE *piro* ~ Sk **pīrə** < PS *pirə [SW 125 *pirâ]) where the relevant reflexes show no sign of a final *w, which casts further doubt on the validity of his reconstruction (already with a question mark) *pidiwti for TN *pyirəd°* etc.

4.9.6 Other derivational processes

A Samoyed specialty would seem to be what I have called "omnibased" derivatives, applicable to both nouns and verbs, and among them those that carry kind of particle-like semantics and are therefore notoriously difficult to gloss, viz. TN (limitative) *-ryi* 'only', (simulative) *-rəxa* 'like, resembling', (concessive) *-xərt°* 'not even', (affirmative) *-xəwa* 'even', e. g. (denominal) *noxaryi* 'only an Arctic fox', *xoy°rəxa* 'like an hill', (deverbal) *pərompəryi°* "only hurrying", 'gradually', *ŋœdarəxa* 'seems to be'. Despite their vague meanings, they are prone to lexicalizations just like any other derivatives, cf. TN *pyel°xa* 'flounder, flatfish' from *pyeh* 'palm' or European TN *xíbyaryi* 'human being' from *xíbya* 'who'.

4.10 Notes on Samoyed syntax

In what follows, I concentrate on what could be called micro-level syntactic features and refer to major works such as Siegl (2013), Nikolaeva (2014), and Wagner-Nagy (2019) for wider perspectives on the topic. Furthermore, I deliberately focus on Tundra Nenets, both for practical and historical reasons, hoping to avoid errors thanks to my familiarity of the data but also because I am able to give illustrative examples, given that Tundra Nenets is a syntactically conservative language within an already conservative group of languages.

As noted in the introduction, the word order in the Samoyed languages is predicate-final, with all of its typologically cogent consequences. A regular transitive sentence, with preferred slots for adverbials, appears as Time adverbial–Subject–Place adverbial–Object–Manner adverbial–Predicate verb. Any focused constituent may be placed in preverbal position, but otherwise the order is quite rigid, only heavy emphasis, or rather afterthought, resulting in a postverbal constituent. Intensive contacts with Russian (or Komi), however, have often led to a markedly freer word order, notably in western Tundra Nenets and parts of the Selkup language area. In Nganasan, curiously, it is not uncommon to have post-verbal adverbial phrases, although SVO structures still remain rare.

In accordance with the overall SOV pattern, a modifier always precedes its head. The unmarked word order in a maximally long noun phrase would be Demonstative–Genitive–Numeral–Adjective–Relative–Noun. Again, permutations are possible when required by focus considerations.

Question words notably do not cause changes in the word order. For instance, a question such as TN *nyar° xə-nyana yilye°?* ("your.friend where lives?") 'where does your friend live?' has the same structure as an affirmative sentence *nyar° Ya-mal°xəna yilye°* ("your.friend in.Yamal lives") 'your friend lives in Yamal'.

In sentence negation, the two final word-forms are, in this order, the negative auxiliary, which carries all inflectional markers, and the main verb in the connegative, e. g. TN *Ya-mal°xəna nyíq yilyeq* 'they [pl] do not live in Yamal' negates *Ya-mal°xəna yilye°q* 'they [pl] live in Yamal', or *nyír° xadaq* 'you did not kill it' vs. *xadaər°* 'you killed it'. This system is preserved in northern Samoyed, while Selkup only employs an uninflected sentence negator Sk (№ 229) **aśa** (with variants), and Kamas exhibits a split system with either a negative auxiliary or a negative particle depending on the mood and tense of the verb (Klumpp 2022: 830). Since nouns do not have a connegative, a negative nominal clause ends with the negative auxiliary followed by the connegative of the copula, cf. TN *yuryor° lúca nyí ŋaq* ("your.partner Russian not is") 'your partner is not Russian' vs. *yuryor° lúca* 'your partner is Russian', or *lúcadaq nyídaq ŋaq* 'you [pl] are not Russian' vs. *lúcadaq* 'you [pl] are Russian'.

The nominative is obviously the case of grammatical subjects, but it is also used, at least in syntactically conservative varieties, for grammatical objects in subjectless constructions, notably in sentences with the second person imperative, which have the object in the nominative, cf. TN *xalya taq!* 'bring a fish!' vs. *xalyam ta°q* 'they [pl] brough a fish'. This would seem to be an ancient Uralic feature, while an alternative system involving nominative as the case of an indefinite object may be attributed to the influence of neighbouring languages. It must be further noted that negative imperative clauses at least in Nenets, involving negative forms that have indicative rather than imperative personal suffixes, take an accusative object, e. g. FN *kałyam nyĭnyun° kataq!* 'do not kill the fish!' vs. *kałya kataq!* 'kill the fish!'.

With the above-mentioned exception, the accusative marks the object of a transitive verb. Personal pronouns, however, differ from nouns in that they appear in the accusative even in imperative sentences, e. g. TN *syid°naq nyadaq!* 'help us [pl]!' (not **mənyaq nyadaq!*) with the form *syid°naq* 'us' being the same as in *syid°naq nyadaən°* 'you [sg] helped us [pl]!', cf. *yuryor° nyadaq!* 'help your partner!' vs. *yuryomt° nyadaən°* 'you [sg] helped your partner', where *yuryor°* and *yuryomt°* are the nominative and accusative 2sg possessive forms, respectively.

In addition to possession and similar attributive functions, e. g. TN *xalyah ŋəbt°* 'smell of a fish', *xalyah paŋo* 'massive occurrence of fish, fish season', *ser°h pad°* 'bag of salt', the genitive (of nouns but not personal pronouns, see immediately below) is used in postpositional phrases, e. g. TN *waqw°h nyinya* 'on the bed', and in embedded clauses, i. e. for a subordinate subject, e. g. TN *myercyanta tyebəb°q* 'when the wind blows' with the 3sg gen.sg form (instead of the nom.sg *myercyada*) followed by a subordinative form of the verb *tyebə-* 'hit'. The genitive is also commonly used in time adverbials as in TN *pyih* 'at night'.

The Samoyed languages are typical "pro-drop" languages where the subject is indexed on the verb and personal pronouns appear only for emphasis, e. g. TN *mənyaq Ya-mal°xəna yilye°waq* 'it is us [pl] that live in Yamal'. In all functions of the (nominal) genitive listed above, personal pronouns appear in the nominative, e. g. TN *mənyaq ŋesixənanaq* 'in our camp [unlike in yours]', or *mənyaq nya°naq* 'to us [rather than to

you]'. At least in Enets and Nenets, however, personal pronouns do have a genitive form, but its use is restricted to a few instances of obligatorily non-possessive phrases where personal markers are not allowed in the main word, e. g. TN *syid°qnaq nyana* 'around us' or *syid°qnaq toql°xa* 'resembling us'.

The local cases are not unexpectedly multifunctional, which is occasionally reflected in terminology as well. The dative has also been called "lative" or "dative-lative", cf. TN *ŋəcyeken°h taəw°* 'I gave it to the child' but *pedaran°h xəyad°m* 'I went to the forest', with the dative forms of *ŋəcyeki°* 'child' and *pedara* 'forest' in the respective functions. The locative is often termed "locative-instrumental" because of forms such as *xər°xəna* 'with the knife'. The prosecutive (sometimes referred to as "prolative") is perhaps most frequently used for adjectival nouns in an adverbial function, e. g. TN *nyencyelməna yilə°da* '(s)he lifted it easily', cf. *nyencyel mənc°ya* 'easy task'. Local cases may obviously be governed by a verb, cf. TN *ŋəmken°h xərwaən°?* 'what do you want?' or *nosyiqməna xanye°waq* 'we hunt Arctic foxes?', where *ŋəmken°h* is the dative of *ŋəmke* 'what' and *nosyiqməna* the prosecutive plural of *noxa* 'Arctic fox', as required by the verbs *xərwa-* 'want' and *xanye-* 'hunt', respectively.

The category of comparison has no morphological marking in Samoyed as such, but the compared item appears in the ablative, cf. TN *tobak° sæd°wa pyíbyi sæd°waxəd° nyencyel* (literally "socks sewing boots from-sewing easy") 'sewing socks is easier than sewing boots'. The same pattern holds good when the adjective is morphologically a verb as in TN *yiryiw° nyísyaxədən° munocyiyeqŋa* (literally "my-grandfather from-my-father is-bearded") 'my grandfather is more bearded than my father'. A superlative meaning can only be attained through specific attributes such as TN *sya-myan°xət°* (literally "of-which-measures" or the like), e. g. *sya-myan°xət° ŋarka* 'the biggest' and *sya-myan°xət° myeryecy°* 'the fastest' (with the gerund of the verb *myerye-* 'be fast').

Postpositions have been defined above as a minor word class on morphological grounds, but there are also what could be called syntactic postpositions, i. e. local case forms of otherwise fully inflected nouns function in the same patterns as the morphological postpositions, notably not allowing intervening elements within postpositional phrases. In Tundra Nenets, four nouns are employed in this manner, viz. *ŋesoh* 'joint', *yolcy°h* 'time', *xœw°* 'side', and *poh* 'gap', for example, *yəxah ŋesont°h* [dat.] 'up to the river', *tə°h yolcy°ŋkəna* [loc.] 'during the summer', *tuh xœw°xəd°* [abl.] 'from beside the fire', and *myadoqnaq pom°na* [pros.] 'through between our tents'.

As seen in several examples above, a predicate, whether a verb or a noun, agrees with the subject in person and number. A predicate verb also agrees in number with the object if it is in the objective conjugation, e. g. TN *pyiryeəw°* 'I cooked it' : *pyiryeŋax°yun°* 'I cooked them [du]' : *pyiryeyən°* 'I cooked them [pl]'. The choice of conjugation is sensitive to focus, so that when introduced as new information, the object is placed immediately before the verb which appears in the subjective conjugation, e. g. TN *yam pyiryeəd°m* 'I cooked soup'. When non-focused, the object may appear apart from the verb, e. g. *yam pyíx°nya pyiryeəw°* 'I cooked soup outdoors'. The object can also be com-

pletely omitted, as in *pyiryeəw°* etc. above, the verb being obligatorily in the objective conjugation. Notably, clauses with subjective forms such as **yam pyíx°nya pyiryeəd°m* or **pyiryeəd°m* are ungrammatical in the absence of a preverbal object, while it is possible to have an objective form preceded by an object as in *yam pyiryeəw°* 'I cooked the soup', with the implication that the object is known from the immediate context. An example of the contrast between the subjective and objective conjugations that came up spontaneously in a discussion with the late Tundra Nenets scholar Helena Susoi was about how to say, 'wipe the table!': *toly° tæs°q!* if "a task arose" but *toly° tæt°!* if "there was already a plan", with the respective subjective and objective 2sg imperative forms of the verb *tæq-* 'wipe'. Once again, personal pronouns make an exception to the rule, as in their presence as an object of the clause the predicate verb must be in the subjective conjugation, cf. *syit° ŋaxət° mənеqŋa* '(s)he saw you from afar' with the 3sg subjective form *mənеqŋa* (instead of the objective *mənеqŋada*) despite the presence of an adverbial between the pronominal object *syit°* and the verb.

The habitive construction in Samoyed typically involves an existential verb in a sentence with the subject noun phrase marked for person, e. g. TN *ŋənor° tənya°* 'you [sg] have a boat' vs. *ŋənor° yaŋku* 'you [sg] do not have a boat'.

The passive in Samoyed may involve derivational morphology, e. g. TN *xada-* 'kill' vs. *xadara-* 'be killed', a process notably more productive in Nganasan, or what might be called syntactic passive, cf. TN *tyiki° xən° myeləd°h syertawi°* 'this sledge is made by a master', with the perfective participle of *syerta-* 'make' as the nominal predicate preceded by the genitive of *myel°q* 'master'. Relevant features concerning Nganasan and the other Samoyed languages are thoroughly discussed by Leisiö (2006).

Within a noun phrase, an attribute does not usually agree with its head in case, but optional agreement in number is possible. For example, the normal way to say 'big rocks' in Tundra Nenets *ŋarka pœq* with *ŋarka* 'big' in the singular, but *ŋarkaq pœq* with the attribute in the plural would add definiteness as in 'the big rocks'. Notably, the place name *Ŋarkaq pœq* 'the Urals' typically has both elements in the plural. In a special form of agreement, an attribute may duplicate the possessive suffix of its head, cf. *xiy°waq yawaq* 'our surroundings', literally "our.surrounding our.place". Because of the structure of the possessive suffixes, this occasionally leads to agreement in case as well, e. g. *syidyamta ŋudamta* 'her/his two hands' in the accusative, cf. *syidya ŋuda* 'two hands', although such examples may be restricted to Tundra Nenets folklore language and serve metric rather than grammatical purposes.

As seen in the previous example, a numeral is followed by a singular noun, which is another ancient Uralic feature. Paired body parts and other regular pairs of objects appear typically in the plural rather than in the dual, which is again a Uralic feature as opposed to the traditional Indo-European usage, e. g. TN *ŋudaq* 'hands [pl]' : *ŋudyin°* 'my hands [pl]', whereas a dual form such as *ŋudax°h* 'hands [du]' would be used in the marked context when, say, two hands belonging to different individuals are referred to. A peculiar function of the dual, shared with Mansi and Khanty, the other Uralic branches that have nominal dual, is coordination. In this construction, both of the coordinated

elements, while themselves single, are marked for dual, e. g. TN *nyebyax°h nyísyax°h* (literally "mothers [du] fathers [du]") 'mother and father; parents'.

Definiteness has no specific morphological expression in Samoyed, but the possessive declension is widely used to mark noun phrases as definite. In the case of natural phenomena, 3sg forms are employed, e. g. TN *myercyada* 'the wind', but 2sg forms appear commonly to refer to an entity that is presumed known to the addressed person, for instance, if a form such as *nyúnyar°* which literally means 'your loon' is found in Tundra Nenets material, the likeliest interpretation is 'the loon [that was discussed before]'.

As already referred to above, an unmarked nominal predicative clause has no copula, unless negation is involved, cf. TN *wœwa* 'it is bad' vs. *wœwa nyí ŋaq* 'it is not bad'. In Enets and Nenets, one-word expressions such as TN *wœwasy°* 'it was bad' are available because of the morphotactics of the past tense marker, cf. *səqlad°m* 'I am stupid' vs. *səqladəmcy°* 'I was stupid', or even *nyar°* '(s)he is your friend' vs. *nyarəsy°* '(s)he was your friend' (but two personal suffixes are not allowed in a word-form so 'I am your friend' etc. would involve a personal pronoun). Other exceptions to the lack of copula in Nenets, in particular, are found when another tense or a mood needs to be expressed, e. g. *wœwa ŋœŋku* 'it will be bad' with future or *syencən° ŋaq!* 'be healthy!' with imperative marked in the copula, in TN *ŋœ-* 'be'.

The destinative declension, as noted in the morphology section, combines the destinative marker, reflecting the Uralic translative case suffix, with the three grammatical cases thanks to the presence of three respective series of possessive suffixes. In Nenets, the nominative destinative appears to be restricted to imperative sentences following the general case marking rule explained above, e. g. TN *pyadər° pœəs°q* 'chop wood (for yourself)', where the form *pyadər°* contains the destinative marker *-də-* followed by the nominative 2sg possessive suffix, whereas the semantically nearly equivalent *pyadəmt° pœ°qÿin°* employing the conjunctive mood, expressing a more polite command than the imperative, has the destinative marker preceding the accusative 2sg possessive suffix. Nganasan and Enets, by contrast, have a much wider use of the nominative destinative, extensively documented by Leisiö (2014). As for Nenets, another example TN *yamp° xər°dəmta yadə°* '(s)he forged her/himself a long knife' shows the typical use of the destinative form as an object, i. e. adding the meaning 'for oneself' to the whole. The third case form in the destinative declension, while formally genitive, has also been called determinative. It represents the direct continuation of the Uralic translative, given than all non-grammatical cases employ the genitive series of the possessive suffixes. A classic example would be TN *nyed°nta me°da* 'he took her as his wife', with a 3sg destinative form of *nye* 'woman, wife' and an 3sg objective form (in the absence of a grammatical object, see above) of *me-* 'take'. There is also a destinative form unmarked for person, but, incidentally, it is only used after a genitive attribute, e. g. *xasawanta maly°cyad° sœd°byi* 'she is sewing a coat for her husband', when the beneficiary, here *xasawa* 'man, husband', is not coreferential with the agent, and appears in the 3sg possessive genitive followed by the absolute destinative form *maly°cyad°* of *maly°cya* 'coat'. As an independent form, the destinative is at least in Nenets replaced with the essive, cf. *səwa*

yerw°ŋe° toli '(s)he is considered a good boss', with the recently, and unevenly among the northern Samoyed languages, grammaticalized essive marker *-ŋe°* in *yerw°ŋe°* 'as, for a boss'.

Subordination is expressed by non-finite constructions, which is also the case of relative clauses. Simple parataxis often serves for co-ordination, but various connective adverbs are also available, for instance, TN *ŋanyih* 'again' can be glossed 'but' when it appears in the second of two connected clauses, but since it is not restricted to such use, it would rather be classified as an adverb. In more traditional varieties of the Samoyed languages, there are therefore no conjunctions, but recent language contacts have introduced various loan conjunctions. In Nenets, there is one inherited element, TN *yib°q* 'although', that does not occur in isolated clauses and should strictly speaking be regarded as the one and only conjunction present, but historically it is a petrified subordinative form.

Yes-no questions are expressed by an interrogative mood when available (with past reference in Enets and Nenets as explained above), but, in most cases, a distinct question intonation is used.

Compounding is not a widespread feature of Samoyed, but it is found in specific contexts. In Tundra Nenets, compounds can be identified in cases where two elements keep their stress patterns but are joined by obligatory sandhi. Two examples were already mentioned in the beginning of this section, for instance, TN *xə-nyana* 'where' cannot be pronounced without sandhi (**xəh nyana*) but the first syllable of *nyana* continues to be stressed. *Ya-mal°* 'Yamal' follows the pattern when used as a place name, yet it is possible to say *yah mal°* 'land's end' in an appellative sense, although the sandhi variant *ya_mal°* is predominant here as well. The locative expression *Ya-mal°xəna* shows the stress placement, as the locative of a hypothetical **Yamal°* would be **Yamaləx°na* instead. That compounds are phonologically contrastive in Tundra Nenets can be seen in minimal pairs such as *tə-nyana* 'there' vs. *tənyana* 'existing'. Incidentally, compounds represent the one single occasion where Nikolaeva (2014) made a substantial and, unfortunately, detrimental change to the phonological transcription, as she simply deleted the hyphen used in this function in the available raw materials, notably the Tundra Nenets sample sentence corpus <www.mv.helsinki.fi/home/tasalmin/tn_corpus.html> based on Tereščenko (1965).

4.11 In place of a conclusion

This chapter has focused, certainly too much, on Samoyed historical phonology, so finally here I address some more general points of ethnohistory and related topics. As mentioned above, Samoyed emerged as a lexically innovative branch of the Uralic family if the sheer quantity of shared lexicon is under scrutiny. Yet it would not seem that Samoyed differs that much from another geographic outlier, viz. Hungarian, or, for

that matter, Saami, which constitutes another branch with a presumably high incidence of substrate lexicon. All in all, Samoyed would seem to parallel the other groups consisting of closely-related languages within Uralic, i. e. Saami, Finnic, Mordvin, Mari, Permic, Hungarian, Mansi, and Khanty, and the idea of these branches making up an intermediate "Finno-Ugric" subgroup as opposed to Samoyed seems unwarranted. In the same fashion, the final disintegration of Proto-Uralic would have taken place considerably later than assumed earlier, i. e. towards the end of the second millennium BC.

Accordingly, the time when Proto-Samoyed was spoken was roughly the first millennium BC. As mentioned in the introduction, the most notable language contacts at the Proto-Samoyed period, or soon thereafter, took place with speakers of Indo-European (at least Indo-Iranian but potentially also Tocharian) and Turkic (*r* Turkic, also known as Bolgar Turkic), as well as possibly Mongolic, Tungusic, Yeniseian, and Yukagir. Furthermore, Janhunen (1998: 477) points out that words for 'sky, heaven, god' (TN *num*) and 'dead body' (TN *xalmyer°*) could have their origin in an early form of Khanty.

Proto-Uralic already had a small but significant layer of Indo-European borrowings, and a number of them were inherited by Samoyed, notably PU *weti > PS *wet > TN *yiq* 'water' and *nimi > *nim > *nyim ~ nyum* 'name'. Both of these words have been used as evidence for long-range hypotheses such as "Indo-Uralic" and "Nostratic", yet the similarities between the loan originals and the borrowings alone speak for language contact, given that the contemporaneous early forms of Proto-Indo-European and Proto-Uralic would have been grammatically and typologically highly different. What is more, while the word for 'water' can be reconstructed in Proto-Uralic, it is not strictly speaking common Uralic but only covers part of the Uralic language area, the original PU word *śäčä 'water' being retained in Saami in its original meaning and in Khanty as 'flood', pushed from its earlier sense by a derivative of 'ice'. Several other words could be discussed, but in our context the main conclusion is that Samoyed does not differ substantially from the other Uralic branches when it comes to early Indo-European contacts. That said, there are several words of Indo-European origin that are widespread in Uralic but absent in Samoyed, for instance the PU *śorwa 'horn' which was borrowed from an Indo-European dialect often dubbed Pre-Indo-Iranian because it exhibits some but not all Indo-Iranian developments (Aikio 2022 reconstructs PU **śarwi which, however, is based on secondary reflexes in Finnic). The PS word *amtə̂ > TN *nyamt°* 'horn' could in fact be a retention of an earlier Uralic word, replaced by a prestigious loanword elsewhere. In the case of the word for 'hundred', Samoyed alone exhibits a Turkic borrowing as discussed below. By contrast, the "Finno-Ugric" word appears to be a genuine Indo-Iranian loanword which spread relatively late across most of the Uralic language area. As for the word for 'metal', PS *wäsa 'iron' > TN *yesya* 'iron, money', I agree with what seems to be the consensus that it was borrowed from an Indo-European source at a relatively late stage, yet it is of considerable age, the only reason for not deriving PS *wäsa from PU *wäśkä directly being the rare second-syllable *a, but it rather represents typical variation found in loanwords that have been adopted by an already dialectally diversified proto-language. The word for 'seven', PS *säjt^{3}wə > TN

syíqw°, is obviously of Indo-European origin in all Uralic languages, yet it seems clear that the word was either borrowed parallelly several times or spread secondarily across branches. I would also add the word for 'salt', PS *sẹr > TN *ser*, to this group of relatively early but parallel Indo-European borrowings. Often treated as homophony rather than polysemy, PS *sẹr also has meanings 'ice' and 'white' in several Samoyed languages, and in Selkup in particular, 'white' is the only meaning of Sk (№ 2697) **sẹrə** (a generalized vowel stem), cf. a separate Iranian borrowing Sk (№ 2697) **śāq** 'salt', and Sk (№ 321) **ūlqō** 'ice' of potentially Yeniseian origin (in Taz Selkup also recorded as *ulqā ~ qulqā* pointing to variation typical of relatively recent loanwords).

A specifically Tocharian origin has been proposed for the above-discussed Proto-Samoyed 'seven' and 'metal', perhaps without conclusive evidence so far, but the topic is being actively pursued by several colleagues in Leiden and beyond. In any case, a few loanwords specific to Samoyed have been attributed to Indo-Iranian or simply Iranian influence, as discussed in a wider context by Holopainen (2019 and the literature therein). The word for 'bridge', with reflexes such as Ng *hula* vs. TE *pule* vs. TN *pul°* vs. Sk (№ 708) **pi̯li**, would be characteristic of this small layer of borrowings insofar as it is, first, difficult to establish the exact Iranian loan original, and second, impossible to reconstruct a Proto-Samoyed form, *pulâ ~ *pi̯lâ [SW 131] being notably able to account for the Nenets word alone. A similar case is Ng *ŋarka* vs. TE *bogo* vs. TN *wərk°* vs. Sk (№ 2281) [northern-central] **qorqə** ~ [southern] **quârqə** 'bear', apparently deriving from an Indo-Iranian word for 'wolf', where the suggested PS reconstruction *wârkâ [SW 170] only explains the Enets and Nenets and possibly the northern-central Selkup records. A word only attested in northern Samoyed, but reconstructable on the basis of Ng *d'aa* ~ TN *ya* as PS *jaâ (Janhunen 1983 *jäâ*) 'flour', comes from an Indo-Iranian word for 'seed' and represents a parallel borrowing to Finnic *jüvä 'grain' (< Pre-Indo-Iranian). The Enets words TE *d'aud'a* ~ FE *d'aasa* appear to be opaque compounds with TE *ud'a* ~ FE *osa* 'meat' as the latter part. In Selkup, the word for 'flour', Sk (№ 825) **mukā** is a more recent Russian loanword. Finally, presumably reflecting an Iranian word meaning 'sharp', a word attested in Nenets and Selkup alone, FN *tyexa* 'fishing hook' vs. Sk (№ 1200) **tēqā** ~ **tă̄qā** 'bear spear, sword', has been reconstructed as PS *täjkå [SW 154], yet the derivative TN *tyesyeta* ~ FN *tyesyaqta* < PN *tyesyaqta 'iron-head arrow' (a participle of a possessive verb, which is the only attestation in Tundra Nenets) suggests that there was no *j in PS as it would have resulted in PN **tyæsyaqta > TN **tyísyeta* instead, while the Selkup records could well reflect a vowel-glide sequence, but the variation in the first-syllable vowel remains nevertheless unexplained.

A currently widespread word of ultimately Iranian origin, TN *nyany°* vs. FN *nyany°* vs. Sk (№ 1658) **ńāń** 'bread', with parallels in Mansi, Khanty, and the northern Yeniseian languages, must have entered north-western Siberia relatively late through the initial mediation of Komi. Notably, the Tundra Nenets and Forest Nenets words are not cognates reflecting PN **nyanyə as this would have resulted in FN **nyaj°* instead. Furthermore, Nganasan and Enets have an older Russian loanword for 'bread', viz. Ng *kiriba* ~ TE *kiroba* ~ *kiriba* ~ FE *kiriba* (< Russian *xleb*).

Among the Turkic borrowings in Samoyed, the word for 'hundred' has been mentioned several times above, i. e. PS *jür deriving from an *r* Turkic language, currently only represented by Chuvash, but *r* Turkic languages evidently had an extensive Siberian range in the past. Another widespread word of Turkic origin is that for 'horse', yet here the variation typical of loanwords is again evident, cf. Ng *d'indå* ~ *d'ündå* ~ TE *d'uda* ~ Yr ‹júda› ~ TN [Siberian] *yuna* ~ [European] *yuno* ~ Sk (№ 1512) **ćüntə** etc., reconstructed by Janhunen as PS *juntå (~ *juntə̑) [SW 49], where *juntå explains the Enets, Yurats, and Siberian Tundra Nenets reflexes, while **juntə̑ rather represents an independently unattested root created by reconstruction techniques along the lines of SW **kåjə̑- and **kontə̑- mentioned in section 7. The Samoyed word for 'navel', mentioned below among potential Mongolic borrowings but possibly of Turkic origin, could be reconstructed as PS *küən [SW 79 *küŋ] on the basis of northern Samoyed evidence, cf. Ng *kiiŋ* (: pl *kiinəq*) ~ TE *śuuq* ~ TN *syúh* < PS *küən [SW 79 *küŋ], yet Selkup points to a final *ń instead, cf. Sk (№ 2573) **śə̄ń**, with **śə̄l'** as an unexpected dialectal variant. Besides an early overview by Janhunen (1977b), several publications have been dedicated to Samoyed-Turkic contacts, notably Terent'ev (1999).

Joki (1952) discusses, among numerous Turkic loanwords that entered Kamas and Mator in the course of their separate development, a selection of widespread Samoyed words of potential Mongolic origin, as summarized by Janhunen (1977b: 124), viz. PS *kåjmå > TN *xœwa* 'bone marrow' (replacing PU *wiδim), PS *küən > TN *syúh* 'navel' (cf. above), PS *tər > TN *tər* 'body hair', PS *tirämä > TN *tirebya* 'roe', PS *kür > Ng *kir* 'fly', and PS *wekånå > TN *yexana* 'sturgeon', and while body parts and animal names may include borrowings, an alternative scenario might in my view involve a joint substrate as a contact language for both Samoyed and Mongolic. Two lexical items associated with clothing are further discussed by Janhunen, viz. PS *pajmå > TN *pyíwa* ~ Ng *φajmu* 'boot' and PS *såjtə̑- > TN *sœdə-* ~ Ng *sot'ə-* (< *sojtə-*) 'sew', and while early Samoyed-Mongolic contacts remain problematic, some comparisons may prove promising.

The contacts between Tungusic and Samoyed have been discussed in an extensive monograph (Anikin & Xelimskij 2007), yet many questions remain open. For a long time, it was assumed that Proto-Samoyed *jam [SW 40 *jäm] 'sea' (> Km *nam* ~ TN *yam* 'sea, large river') would be a Tungusic borrowing (cf. Evenki *lāmu* 'sea'; Anikin & Xelimskij 2007: 48–49), but this idea has now been largely rejected (Janhunen 2013). Curiously, apparent Tungusic loanwords are often found in northern Samoyed alone, yet they may have been borrowed relatively early, cf. TN *lokï* 'blunt-headed arrow' ~ FN (replacing PN *muŋkə 'arrow') *luqkyi* 'arrow, bullet, grain of shot' (Anikin & Xelimskij 2007: 104), or Ng *satərə* ~ TE *seδoro* 'Arctic fox' (Anikin & Xelimskij 2007: 107).

The putative similarities between Yukagir and Samoyed (or Uralic in general) have been subject to various interpretations. These have been summarized and elaborated by Aikio (2014b), who concludes that, while most of the proposed cognates prove faulty, a rather sizable corpus of potential loanwords from Proto-Samoyed (often Pre-Proto-Samoyed) to Proto-Yukagir may be assumed. My own interpretation of the data would be areal rather than temporal, i. e. the source of the Uralic borrowings in Yukagir could

have been a Para-Samoyed idiom established as a parallel offshoot to Proto-Samoyed proper. Such a scenario obviously remains hypothetical, but operating with pre-pro-to-languages carries a speculative element as well. The Uralo-Yukagir long-range hypothesis, firmly rejected by Aikio, suffers from the same problem as the Indo-Uralic hypothesis mentioned above, namely that the grammatical differences between the Uralic family (or its grammatically conservative Samoyed branch) and the Yukagir family would suggest a considerable time depth, yet the Yukagir words in question are highly similar to their proposed Uralic cognates. Typologically, Uralic and Yukagir are obviously quite similar, exhibiting many features of characteristically agglutinative SOV languages, which is not unusual in Siberian context except for the Yeniseian family (see Vajda, this volume). Yeniseian may be mentioned in this context also because a few Uralic borrowings have been suggested there. Many of them, however, point to a source not immediately identifiable with Samoyed but possibly another hypothetical offshoot, perhaps closer to Khanty.

Inspired by Aikio (2014b), who identifies a group of several words widely known in Uralic and presumably borrowed into Yukagir but not attested in Samoyed, I venture the following case studies of lexical change. (i) PU *sula- 'melt, thaw' has been replaced with PS *ju- 'melt, warm up' (already mentioned in section 7), a verb that looks basic and ancient, yet with no apparent cognates. Nenets has another verb PN *kolka- > TN *xolka-* 'melt, ripen' ~ FN *kołka-* 'melt', which, because of its syllable-final *l, cannot derive from an early stage of the language (cf. the sound change *l > *j discussed in section 2). What is more, 'ripen' in Forest Nenets is *ku-* < PN *kú-, in other words, another basic-looking verb that has no known cognates. (ii) PU (?) *śola 'bowel, intestine, gut' (whatever the PU reconstruction is, it cannot be **śali suggested by Aikio 2022 since that is based on secondary reflexes in Finnic, cf. the discussion on PU *śorwa 'horn' above) has been replaced with PS *wätäw [SW 175 *wetə̂-jə̂j] (see section 7), a word of potential substrate origin found in all Samoyed languages. (iii) PU *śoδ'ka 'goldeneye' (a species of duck, important in ancient mythology) has no fixed term in Samoyed to the extent that Sk (№ 2386) **kulčaŋ** is the only single-word expression recorded for 'goldeneye'. The Nenets languages have similar phrases for 'goldeneye', viz. TN *ŋæwa ŋarka* (literally "big-headed") vs. FN *ŋyiwi ŋałka* (literally "big-brained"), manifesting a syntactic pattern not otherwise discussed in this chapter. The main point to be highlighted here is that the lexis of the Samoyed languages is still known rather unevenly.

Language contacts are obviously crucial for locating the Samoyed *Urheimat* and to some extent also the areas where the contact languages were spoken. Furthermore, several distinct substrate zones can be identified on the basis of lexical differences. For instance, since the names of sea mammals found in Nenets are not cognate with those shared by Enets and Nganasan, it may be concluded that two different aboriginal languages were the donors of these terms. Helimski (2001b) identifies three distinct substrates in north-western Siberia, i. e. (i) the "Proto-Yugran" language, responsible for many if not most similarities between Mansi and Khanty, (ii) the *Syix°rtya* language, referring to the TN name for an aboriginal population, in much of the current Nenets

territory, and (iii) the Taimyr substrate language, influencing Nganasan in the first place but also Enets to some extent. As for the names of sea mammals, Nganasan and Enets notably share a name for 'seal', recorded as Ng *śia* 'seal' ~ TE *sia* 'ringed seal' ~ FE *sie* id., without a recognized cognate in Nenets, where the words TN *nyək°* 'ringed seal' and *ŋərtyih* 'bearded seal' not only have no cognates outside Nenets, but *ŋərtyih* in particular shows a consonant cluster that could not have been inherited from Proto-Samoyed. Similarly, TE *kuaga* 'white whale' is unrelated to TN *webarka* id. (the latter looking like a productive derivative of *weba* 'leaf' which is, needless to say, semantically opaque). The word for 'whale', by contrast, appears similar across northern Samoyed, cf. Ng *koliiŋ* 'whale, white whale' (apparently also covering species of large fish) ~ TN *xaleh* 'whale' (> TE *xale* id.), yet the second-syllable vowel correspondence is irregular, pointing to a secondary contact within northern Samoyed, the Enets word being a more recent Tundra Nenets borrowing. For a fascinating connection with the concept of 'mammoth' see Kaheinen (2022) and the literature therein.

Furthermore, TN *tyiw°tyey°* 'walrus' represents a transparent derivative of *tyibya* 'tooth' via the possessive verb *tyiwəq-* 'have teeth' with an imperfective participle *tyiw°tya* from which *tyiw°tyey°* incorporates an adjectival marker to become lexicalized in the meaning. Since the topic was not discussed before, it must be noted here that the widely celebrated idea first suggested by Abondolo (1996: 57–58) that the Samoyed word form 'tooth' (PS *timä > Ng *t'imi* ~ TN *tyibya*) must have originated from a productive infinitive of PU *sewi- 'eat', i. e. *sewi-mä > *sew-mä seems problematic to say the least: PU *sew-mä would not result in an initial *t'i* in Nganasan but rather in **tį-* in its stead, and, contrary to the assumption of Aikio (2002), based on contemporary but inaccurate data, Proto-Uralic sequences involving a postvocalic glide never yield a Proto-Samoyed single vowel (Salminen 2023: 392–394).

Paleolinguistic methods can obviously be applied using both inherited and borrowed vocabulary. From a cultural point of view, lexicon related to shamanism is of major interest, and the words for 'shaman', different in Nganasan versus the other Samoyed languages, was briefly discussed above. The word for 'shaman's drum' would seem to show regular correspondences reflecting PS *päŋkir [SW 119 *peŋkär] > Ng *xeńd'ir* ~ TE *ped'iq* ~ TN *pyency°r* but Sk (№ 588) **pi̯ŋkər** 'musical instrument' is irregular for both the first and second syllable vowels. A morphological analysis of PS *päŋkir would rather be *pä(-)ŋk-ir, i. e. a derivative of a verbal stem which possibly contains an iterative marker *-ŋkə̂ (cf. Janhunen 1998: 476), with parallel instrumental derivative in Nenets, cf. TN *pyeŋkəbcy°h* 'shaman's drumstick' < Pre-Nenets *pä(-)ŋkə-psin, whose connection to Sk (№ 1792) **qapśəń** id. seems spurious, and only a Kamas cognate is accepted by Alatalo. Another otherwise widespread stem *såmpə̂- 'shamanize' again lacks a cognate in Nganasan.

When it comes to material culture, a word for 'domestic reindeer', reconstructed here as PS *tẹa, and related terms dating back to Proto-Samoyed obviously suggest that elements of reindeer husbandry had already started to develop and spread by that time. Other paleolinguistic cues indicate, not surprisingly, that Proto-Samoyed speakers made much of their living by fishing and hunting in the deep taiga.

Ethnohistorical conclusions can naturally be drawn from place names as well. The only place name in SW would also be the most discussed case, viz. the name of Yenisei, for which *Jentəsiń [SW 43 *jentâsä] may be reconstructed on the basis of Enets and Nenets cognates, TE *D'edośiq* ~ TN *Yenəsy°h*, taking into account Ng *D'entəd'iə* which would be interpreted as a generalization of the original vowel stem (in Nenets this word almost always occurs in the genitive, notably TN *Yen°syə°h yam* 'Yenisei' and *Yen°syə°h xəbyi* 'Ket'). The initial CV sequence of the Selkup variant, Sk (№ 1652) **ńantəsi**, is clearly secondary, but otherwise it corresponds to the northern Samoyed cognates. Anikin and Helimski consider it a Tungusic loanword (Anikin & Xelimskij 2007: 47–48), yet its complex phonological shape and the distribution of the lexeme has led most specialists to the conclusion that the Tungusic word is borrowed from Samoyed. Place names can, of course, give indications about the subsequent homelands, such as that of Proto-Nenets, which, besides the name of Yenisei, also had opaque names for the rivers Taz and Pur, i. e. TN *Tasu°* < PN *Tasəmə and *Pyur°h* < *Pyurəŋ, respectively. Tundra Nenets place names further to the west, however, are often transparent, cf. TN *Səlyah yam* 'Ob' ("Cape River"), *Ya-mal°* 'Yamal' ("Land's End"), or *Ŋarkaq pœq* 'the Urals' ("Big Rocks").

Lexical features of Samoyed are obviously beyond the scope of this essay, but I offer a few remarks, based on Nenets but apparently reflecting Samoyed as a whole.

Furthermore, including Samoyed data into cross-linguistic lists of basic vocabulary proves challenging with regard to the concept of 'bird' in particular, as the Samoyed languages by and large lack a single-word generic term. In Tundra Nenets, 'bird' would be expressed by collocations such as *tyírtya sarmyik°* "flying animal" or simply *tyírtya* "flying" in the European dialects and *nuw°h sarmyik°* "sky's animal" in the Siberian dialects, while in Forest Nenets, 'bird' would be referred to a single word but with a wider meaning, viz. *xalaqku* 'animal, bird'. By contrast, both of the Nenets languages have generic words for 'small bird', TN *lœqmor* vs. FN *ătïtyew°*, which are obviously unconnected. There is also no word that would cover 'tail', but three words cover the semantic space, cf. TN *tœwa* 'tail of a mammal' vs. *mortiq* 'tail of a bird' vs. *yabco* 'tail of a fish'.

A curious difference among the Samoyed languages involves 'finger' and 'toe', where Selkup, for instance, has a generic word covering both concepts, i. e. Sk (№ 782) **mūnə** 'finger, toe', while Nenets employs the collocations (TN) *ŋudah tarka* ("hand's branch") 'finger' vs. *ŋœh tarka* ("foot's branch") 'toe'. At the same time, Nenets has single words for four of the five fingers, viz. TN *pyík°cya* 'thumb', *ŋump°ya* 'index finger', *nyumcy°da* ("nameless") 'ringfinger', and *war°ti°* 'little finger' (from *war°* 'edge'), only 'middle finger' having a multi-word expression *yerki° ŋudah tarka* "central hand's branch".

As for 'child', there are two distinct words, cf. TN *nyú* 'own child' vs. *ŋəcyeki°* 'any child'. Furthermore, Nenets in particular does not have separate words for 'daughter' and 'son' (both TN *nyú*) or 'girl' and 'boy' (both TN *ŋəcyeki°*), attributes *nye* 'female' and *xasawa* 'male' being used for ambiguation when necessary, as in TN *nye nyú* 'daughter' or *xasawa ŋəcyeki°* 'boy'.

A feature apparently shared with the other Siberian Uralic branches, Samoyed kinship terms are based on age rather than generation. Thus, the focal points in the

system are the age of the speaker and the age of their parents. For instance, younger siblings are referred to by a generic term, e. g. TN *papa* (or *pyebya*) and FN *kăqka*, the gender disambiguated in the same way as in the case of 'child' above. The irrelevance of generation becomes evident from the facts that a person's elder brother and their father's younger brother are expressed by a single word, e. g. TN *nyinyeka ~ nye°ka* and FN *ĭja*, and that the older siblings of one's parents are covered with words also denoting and usually glossed as 'grandmother' and 'grandfather', e. g. TN *xada* and *yiryi* or FN *kata* and *nyĭłyi*, respectively. Samoyed-speaking communities have traditionally been divided into patrilineal moieties, which is reflected in kinship terminology, so that separate words must be used for relatives belonging to the other moiety, cf., for example, TN *tyidya* 'mother's younger brother' (a word of Proto-Uralic origin often referring to 'paternal uncle' elsewhere) and FN *xojuqma* id. As can be seen from many semantic equivalents between Tundra Nenets and Forest Nenets above, at least some of the kinship terms appear to be subject to rapid lexical change.

Colour terms provide another example of the many lexical differences among the Samoyed languages, not untypical of adjectives in general (cf. Szeverényi 2005). There are four basic colour terms in both Nenets languages, in the sense that the words in question cannot refer to anything other than colour. The Nenets languages share the words TN *serako* ~ FN *xelaqku* 'white' (originally diminutives of PS *sẹr > PN *ser 'ice, salt, white') and TN *pəryidye-* ~ FN *pĭłyitye-* 'be black', while Tundra Nenets has the verb *nyar°ya-* 'be red' in contrast to the Forest Nenets nominal form *kemtya*(*q*)*kyï* 'red' (an opaque derivative of PS *kẹm > PN *kem 'blood'). At the same time, TN *tasyexey°* 'yellow' constitutes a basic colour term (while connected with 'egg yolk'), but FN *pătyałăxa* id. may also have the productive meaning 'bile-like' (its Tundra Nenets cognate *pədyarəxa* rather referring to 'yellow-green'). Conversely, FN *tănsya*(*q*)*kyï* 'blue', although related to words such as *tănsyanye-* 'be hazy', only refers to colour, unlike TN *syun°rəxa* 'blue' (productively 'steam-like'). Not surprisingly, comparisons with the other Samoyed languages produce a checkered picture. While PS *sẹr is widely used as the basis for 'white', Ng *heŋkə* 'black' and Sk (№ 2618) **sə̄qə** id., although morphologically simple, are not related to each other. What is more, the common word for 'white' in Selkup would be Sk (№ 1555) **ćə̄qə**, with uncertain cognates elsewhere. Sk (№ 1695) **ńārqə** 'red' is clearly a correlative derivative of TN *nyar°ya-* 'be red', but again, this stem is not found in all Samoyed languages.

Many similar cases could obviously be discussed in the context, but I conclude this essay hoping that despite its uneven coverage of certain topics and particular bias towards Nenets examples, it may nevertheless contribute to future studies aiming toward a more balanced coverage of the Samoyed languages.

Acknowledgements: In the same manner as Samoyed specialists are able to employ one another's data collections, a lot of information is distributed through personal communication, of which this essay has benefitted to a remarkable degree. Jarmo Alatalo must be mentioned in this context again. He has shared with me not only his extensive materials but also his valuable insights on numerous crucial issues, and there is no

doubt that without his contribution this chapter could not have been written. I wish to mention here my closest teachers in the field of Samoyedology stemming from years back, Ago Künnap, Eugene Helimski, and Juha Janhunen, and I am fortunate to have been able to cooperate with Juha up till today. Other close colleagues who have assisted me during the writing of this chapter include Maria Barmich, Olga Kazakevich, Larisa Leisiö, Gerson Klumpp, Anna Urmanchieva, Marina Lublinskaya, Valentin Gusev, Olesya Khanina, Andrey Shluinsky, and Kaisla Kaheinen. I have also received direct help on current problems from Arja Ahlqvist, Riku Erkkilä, Jaakko Helke, Sampsa Holopainen, and Alexandre Ryzhkov-Shukumine, as well as many others who shall remain nameless here, but they know who they are. I am indebted to Kone Foundation and the University of Helsinki for sustaining my work. Finally, I express my deep gratitude to Edward Vajda, the volume editor, for his guidance and patience.

Abbreviations

FE	Forest Enets
FN	Forest Nenets
Km	Kamas
Mt	Mator
Ng	Nganasan
PN	Proto-Nenets
PS	Proto-Samoyed
PU	Proto-Uralic
Sk	Selkup
TE	Tundra Enets
TN	Tundra Nenets

abl.	ablative
acc.	accusative
dat.	dative
du	dual
intr.	intransitive
gen.	genitive
loc.	locative
nom.	nominative
pl	plural
pros.	prosecutive
refl.	reflexive
sg	singular
tr.	intransitive

References

Abondolo, Daniel. 1996. *Vowel rotation in Uralic: Obug[r]ocentric evidence.* (School of Slavonic and East European Studies Occasional Papers 31.) London: School of Slavonic and East European Studies, University of London.

Aikio, Ante. 2002. New and old Samoyed etymologies. *Finnisch-Ugrische Forschungen* 57. 9–57.

Aikio, Ante. 2006. New and old Samoyed etymologies (Part 2). *Finnisch-Ugrische Forschungen* 59. 9–34.

Aikio, Ante. 2012. On Finnic long vowels, Samoyed vowel sequences, and Proto-Uralic *x. In Tiina Hyytiäinen, Lotta Jalava, Janne Saarikivi & Erika Sandman (eds.), *Per Urales ad Orientem: Iter polyphonicum multilingue*, 227–250. (Mémoires de la Société Finno-Ougrienne 264.) Helsinki: Société Finno-Ougrienne.

Aikio, Ante = Luobbal Sámmol Sámmol Ánte (Ante Aikio). 2014a. Studies in Uralic etymology II: Finnic etymologies. *Linguistica Uralica* 50. 1–19.

Aikio, Ante = Luobbal Sámmol Sámmol Ánte (Ante Aikio). 2014b. The Uralic-Yukaghir lexical correspondences: genetic inheritance, language contact or chance resemblance? *Finnisch-Ugrische Forschungen* 62. 7–76.

Aikio, Ante = Luobbal Sámmol Sámmol Ánte (Ante Aikio). 2022. Proto-Uralic. In Marianne Bakró-Nagy, Johanna Laakso & Elena Skribnik (eds.), *The Oxford guide to the Uralic languages*, 3–27. Oxford: Oxford University Press.

Alatalo, Jarmo. 1994. De sölkupska dialekterna på övre Ob. *Journal de la Société Finno-Ougrienne* 85. 7–20.

Alatalo, Jarmo. 1998. *Сӯ̄ссыӷӯй э̄җипсан: Қэ̄тқый қӯланни* [Selkup dictionary according to the Ket dialect] = *Slovar’ sel’kupsko-russkij i russko-sel’kupskij na ketskom dialekte sel’kupskogo jazyka* [Selkup-Russian and Russian-Selkup dictionary according to the Ket dialect]. Maksimkin Jar/Xel’sinki.

Alatalo, Jarmo. 1999. Zur samojedischen Flexion. In Cornelius Hasselblatt & Paula Jääsalmi-Krüger (eds.), *Europa et Sibiria*, 13–33. (Veröffentlichungen der Societas Uralo-Altaica 51.) Wiesbaden: Harrassowitz.

Anikin, A[leksandr] E. & E[vgenij] A. Xelimskij. 2007. *Samodijsko–tunguso-man’čžurskie leksičeskie svjazi* [Samoyed-Tungusic lexical connections]. Moskva: Jazyki slavjanskoj kul’tury.

Barmič, M[arija] Ja. & I[nna] A. Vèllo. 1994. *Slovar’ nenecko-russkij i russko-neneckij (lesnoj dialekt)* [Nenets-Russian and Russian-Nenets dictionary (Forest dialect)]. Sankt-Peterburg: Prosveščenie.

Burkova, S[vetlana] I. 2010. Kratkij očerk grammatiki tundrovogo dialekta neneckogo jazyka. [A Short study of Tundra Nenets grammar]. In *Dialektologičeskij slovar’ neneckogo jazyka*, 179–349. Ekaterinburg: Basko.

Castrén, M. Alexander. 1854. *Grammatik der samojedischen Sprachen. Herausgegeben von* Anton Schiefner. St. Petersburg: Kaiserliche Akademie der Wissenschaften.

Castrén, M. Alexander. 1855. *Wörterverzeichnisse aus den samojedischen Sprachen. Bearbeitet von* Anton Schiefner. St. Petersburg: Kaiserliche Akademie der Wissenschaften.

Collinder, Björn. 1955. *Fenno-Ugric vocabulary.* Stockholm: Almqvist & Wiksell.

Comrie, Bernard. 1981. *The languages of the Soviet Union.* Cambridge: Cambridge University Press.

Donner, Kai. 1944. *Kamassisches Wörterbuch.* Bearbeitet und herausgegeben von A[ulis] J. Joki. (Lexica Societatis Fenno-Ugricae 8.) Helsinki: Société Finno-Ougrienne.

Donner, Kai, U[uno] T. Sirelius & Jarmo Alatalo. 2004. *Sölkupisches Wörterbuch.* Zusammengestellt und herausgegeben von Jarmo Alatalo. (Lexica Societatis Fenno-Ugricae 30.) Helsinki: Société Finno-Ougrienne.

Gusev, V[alentin] Ju. 2008. Auslautnoe *-w* v samodijskix jazykax: nganasansko-matorskaja izoglossa [The final *-w* in the Samoyed languages: a Nganasan-Mator isogloss]. In *Materialy 2-j meždunarodnoj konferencii po samodistike*, 114–128. Sankt-Peterburg: Nestor-Istorija.

Gusev, V[alentin] Ju. 2022. Refleksy prasamodijskix **a* i **ä* i nganasanskij singarmonizm [The reflexes of Proto-Samoyed **a* и **ä* and the Nganasan synharmonism]. In Valentin Gusev, Anna Urmanchieva & Aleksandr Anikin (eds.), *Siberica et Uralica: In memoriam Eugen Helimski*, 39–62. (Studia Uralo-Altaica 56.) Szeged: University of Szeged.

Hajdú, Péter. 1963. *The Samoyed peoples and languages.* Bloomington: Indiana University.

Hajdú, Péter. 1968. *Chrestomathia Samoiedica.* Budapest: Tankönyvkiadó.

Hajdú, Péter. 1988. Die samojedischen Sprachen. In Denis Sinor (ed.), *The Uralic languages: description, history and foreign influences*, 3–40. Leiden: Brill.

Helimski = Xelimskij, Evgenij. 1986. Etymologica 1–48. *Nyelvtudományi Közlemények* 88. 119–143.

Helimski = Xelimskij, Evgenij. 1992–1993. Etymologica 49–79. *Nyelvtudományi Közlemények* 93. 101–123.

Helimski = Xelimskij, E[vgenij] A. 1993. Prasamodijskie **ə̂* i **ə̈*: praural’skie istočniki i nganasanskie refleksy. [The Proto-Samoyed **ə* and **ə̈*: their Proto-Uralic sources and Nganasan reflexes]. In Marianne Sz. Bakró-Nagy & Enikő Szíj (eds.), *Hajdú Péter 70 éves*, 125–133. Budapest: MTA Nyelvtudományi Intézet.

Helimski = Xelimskij, Evgenij. 1996. Selkup as lingua franca. In Stephen A. Wurm, Peter Mühlhäusler & Darrell T. Tryon (eds.), *Atlas of languages of intercultural communication in the Pacific, Asia, and the Americas.* I: *Maps*, 1035. Berlin/New York: Mouton de Gruyter.

Helimski, Eugen. 1997. *Die matorische Sprache.* (Studia Uralo-Altaica 41.) Szeged: JATE Finnugor Tanszék.
Helimski, Eugene. 1998a. Nganasan. In Daniel Abondolo (ed.), *The Uralic languages*, 480–515. London/New York: Routledge.
Helimski, Eugene. 1998b. Selkup. In Daniel Abondolo (ed.), *The Uralic languages*, 548–579. London/New York: Routledge.
Helimski = Xelimskij, E[vgenij] A. 2000 [1976]. Ob odnom perexodnom severnosamodijskom dialekte [On one transitional northern Samoyed dialect]. In E[vgenij] A. Xelimskij, *Komparativistika, uralistika*, 50–55. Moskva: Jazyki russkoj kul'tury. [Originally published (without appendix) in *Proisxoždenie aborigenov Sibiri i ix jazykov* [The origin of the aborigines of Siberia and their languages] 3. 89–93. Tomsk 1976.]
Helimski, Eugene. 2001a. Samoyedic studies: a state-of-the-art report. *Finnisch-Ugrische Forschungen* 56. 175–216.
Helimski = Xelimskij, Evgenij. 2001b. Ural'cy i ix predšestvenniki: belye pjatna na ėtnoistoričeskoj karte Severnoj Evrazii i ural'skie jazyki. [The Uralic speakers and their predecessors: blank spots on the ethnohistorical map of Northern Eurasia and the Uralic languages]. In *Congressus nonus internationalis fenno-ugristarum.* IV: *Dissertationes sectionum, Linguistica I*, 332–336. Tartu.
Helimski, Eugen. 2005. The 13th Proto-Samoyedic vowel. In Beáta Wagner-Nagy (ed.), *Mikola-konferencia 2004*, 27–39. Szeged: SzTE Finnugor Tanszék.
Holopainen, Sampsa. 2019. *Indo-Iranian borrowings in Uralic.* Helsinki: Faculty of Arts, University of Helsinki. <https://helda.helsinki.fi/handle/10138/307582>
Irikov, S. V. 1988. *Slovar' sel'kupsko-russkij i russko-sel'kupskij* [Selkup-Russian and Russian-Selkup dictionary]. Leningrad: Prosveščenie.
Jalava, Lotta. 2015. *Borders of (in)finiteness: Tundra Nenets predication in change.* Helsinki: Department of Finnish, Finno-Ugrian and Scandinavian Studies, University of Helsinki. <https://helda.helsinki.fi/handle/10138/154753>
Jalava, Lotta. 2022. *Tundraa myöten, joen mukana.* Prosekutiivisijan alkuperästä samojedilais-itämerensuomalaisin vertailuin. In Kaisla Kaheinen, Larisa Leisiö, Riku Erkkilä & Toivo E. H. Qiu (eds.), *Hämeenmaalta Jamalille*, 145–158. Helsinki: University of Helsinki.
Janhunen, Juha. 1977a. *Samojedischer Wortschatz: gemeinsamojedische Etymologien.* (Castrenianumin toimitteita 17.) Helsinki: [Helsingin yliopisto].
Janhunen, Juha. 1977b. Samoyed-Altaic contacts: present state of research. In Juha Janhunen (ed.), *Altaica*, 123–129. (Mémoires de la Société Finno-Ougrienne 158.) Helsinki: Société Finno-Ougrienne.
Janhunen, Juha. 1981. Uralilaisen kantakielen sanastosta. (Über den Wortschatz des Protouralischen). *Journal de la Société Finno-Ougrienne* 77. 219–274.
Janhunen, Juha. 1982. On the structure of Proto-Uralic. *Finnisch-Ugrische Forschungen* 44. 23–42.
Janhunen, Juha. 1983. On early Indo-European-Samoyed contacts. In Juha Janhunen, Anneli Peräniitty & Seppo Suhonen (eds.), *Symposium saeculare Societatis Fenno-Ugricae*, 115–127. (Mémoires de la Société Finno-Ougrienne 185.) Helsinki: Société Finno-Ougrienne.
Janhunen, Juha. 1998. Samoyedic. In Daniel Abondolo (ed.), *The Uralic languages*, 457–479. London/New York: Routledge.
Janhunen, Juha. 2013. The Tungusic languages: a history of contacts. In Kim Juwon & Ko Dongno (eds.), *Current trends in Altaic linguistics*, 17–60. Seoul: Altaic Society of Korea.
Janurik, Tamás. 2021. *Kamassz szótár.* Székesfehérvár: Budenz Alkotóház.
Joki, Aulis J. 1952. *Die Lehnwörter des Sajansamojedischen.* (Mémoires de la Société Finno-Ougrienne 103.) Helsinki: Société Finno-Ougrienne.
Kaheinen, Kaisla = Kaxejnen, Kajsla. 2022. K voprosu o rannix kontaktax meždu samodijskimi jazykami: zametki ob istorii ngan. *колиий* [On the early contacts between Samoyed languages: notes on the history of Nganasan *колиий*]. In Valentin Gusev, Anna Urmanchieva & Aleksandr Anikin (eds.), *Siberica et Uralica: In memoriam Eugen Helimski*, 63–75. (Studia Uralo-Altaica 56.) Szeged: University of Szeged.
Katzschmann, Michael. 2008. *Chrestomathia Nganasanica.* Norderstedt: Books on Demand.

Khanina, Olesya & Andrey Shluinsky. 2022. How different are they and why? In Kaisla Kaheinen, Larisa Leisiö, Riku Erkkilä & Toivo E. H. Qiu (eds.), *Hämeenmaalta Jamalille*, 181–200. Helsinki: University of Helsinki.

Klumpp, Gerson. 2002. *Konverbkonstruktionen im Kamassischen.* (Veröffentlichungen der Societas Uralo-Altaica 58.) Wiesbaden: Harrassowitz.

Klumpp, Gerson. 2005. Die '9' im Kamassischen und im Nganasanischen. *Studia Etymologica Cracoviensia* 10. 65–67.

Klumpp, Gerson. 2019. The numerals of Abakanic. *Ural-Altaische Jahrbücher*. Neue Folge 27. 71–92.

Klumpp, Gerson. 2022. Kamas. In Marianne Bakró-Nagy, Johanna Laakso & Elena Skribnik (eds.), *The Oxford guide to the Uralic languages*, 817–844. Oxford: Oxford University Press.

Körtvély, Erika. 2005. *Verb conjugation in Tundra Nenets.* (Studia Uralo-Altaica 46.) Szeged: University of Szeged.

Koškareva, N[atal'ja] B. 2005. *Očerki po sintaksisu lesnogo dialekta neneckogo jazyka* [Studies in Forest Nenets syntax]. Novosibirsk: Ljubava.

Kosterkina, N[adežda] T., A[leksandr] Č. Momde & T[at'jana] Ju. Ždanova. 2001. *Slovar' nganasansko-russkij i russko-nganasanskij* [Nganasan-Russian and Russian-Nganasan dictionary]. Sankt-Peterburg: Prosveščenie.

Künnap, Ago. 1971. *System und Ursprung der kamassischen Flexionssuffixe.* I: *Numeruszeichen und Nominalflexion.* (Mémoires de la Société Finno-Ougrienne 147.) Helsinki: Société Finno-Ougrienne.

Künnap, Ago. 1978. *System und Ursprung der kamassischen Flexionssuffixe.* II: *Verbalflexion und Verbalnomina.* (Mémoires de la Société Finno-Ougrienne 164.) Helsinki: Société Finno-Ougrienne.

Kuznecova, A[riadna] I., O[l'ga] A. Kazakevič, E[lena] V. Gruškina & E[vgenij] A. Xelimskij. 2002. *Sel'kupskij jazyk* [The Selkup language]. Sankt-Peterburg: Prosveščenie.

Kuznecova, A[riadna] I., E[vgenij] A. Xelimskij & E[lena] V. Gruškina. 1980. *Očerki po sel'kupskomu jazyku: tazovskij dialekt* [Studies in the Selkup language: the Taz dialect] I. Moskva: Izdatel'stvo Moskovskogo universiteta.

Labanauskas, K[azimir] I. 2002. *Rodnoe slovo* [The native word]. Sankt-Peterburg: Prosveščenie.

Lehtisalo, T[oivo]. 1927. *Über den [V]okalismus der ersten [S]ilbe im [J]uraksamojedischen.* (Mémoires de la Société Finno-Ougrienne 56.) Helsinki: Société Finno-Ougrienne.

Lehtisalo, T[oivo]. 1936. *Über die primären ururalischen Ableitungssuffixe.* (Mémoires de la Société Finno-Ougrienne 72.) Helsinki: Société Finno-Ougrienne.

Lehtisalo, T[oivo]. 1956. *Juraksamojedisches Wörterbuch.* (Lexica Societatis Fenno-Ugricae 13.) Helsinki: Société Finno-Ougrienne.

Leisiö, Larisa. 2006. Passive in Nganasan. In Werner Abraham & Larisa Leisiö (eds.), *Passivization and typology*, 213–230. (Typological Studies in Language 68.) Amsterdam: Benjamins.

Leisiö = Lejsië, L[arisa]. 2014. Imennye kategorii vremeni v nganasanskom i drugix severnosamodijskix jazykax [Nominal categories of tense in Nganasan and the other northern Samoyed languages]. *Voprosy jazykoznanija* 2014(1). 39–59.

Mikola, Tibor. 1988. Geschichte der samojedischen Sprachen. In Denis Sinor (ed.), *The Uralic languages: description, history and foreign influences*, 219–263. Leiden: Brill.

Mikola, Tibor. 2004. *Studien zur Geschichte der samojedischen Sprachen.* Aus dem Nachlass herausgegeben von Beáta Wagner-Nagy. (Studia Uralo-Altaica 45.) Szeged: SzTE Finnisch-Ugrisches Institut.

MS = Helimski 1997.

Nikolaeva, Irina. 2014. *A grammar of Tundra Nenets.* Berlin: De Gruyter Mouton.

Rédei, Károly. 1986–1991. *Uralisches etymologisches Wörterbuch* 1–3. Budapest: Akadémiai Kiadó/Wiesbaden: Harrassowitz.

Salminen, Tapani. 1993. Word classes in Nenets (and a few words about their Uralic parallels). In Ulla-Maija Kulonen (ed.), *Festschrift für Raija Bartens zum 25.10.1993*, 257–264. (Mémoires de la Société Finno-Ougrienne 215.) Helsinki: Société Finno-Ougrienne.

Salminen, Tapani. 1997. *Tundra Nenets inflection.* (Mémoires de la Société Finno-Ougrienne 227.) Helsinki: Société Finno-Ougrienne.
Salminen, Tapani. 1998a. *A morphological dictionary of Tundra Nenets.* (Lexica Societatis Fenno-Ugricae 26.) Helsinki: Société Finno-Ougrienne.
Salminen, Tapani. 1998b. Nenets. In Daniel Abondolo (ed.), *The Uralic languages*, 516–547. London/New York: Routledge.
Salminen, Tapani. 2007. Notes on Forest Nenets phonology. In Jussi Ylikoski & Ante Aikio (eds.), *Sámit, sánit, sátnehámit*, 349–372. (Mémoires de la Société Finno-Ougrienne 253.) Helsinki: Société Finno-Ougrienne.
Salminen, Tapani. 2012. Traces of Proto-Samoyed vowel contrasts in Nenets. In Tiina Hyytiäinen, Lotta Jalava, Janne Saarikivi & Erika Sandman (eds.), *Per Urales ad Orientem: Iter polyphonicum multilingue*, 339–358. (Mémoires de la Société Finno-Ougrienne 264.) Helsinki: Société Finno-Ougrienne.
Salminen, Tapani. 2014. Suomalais-samojedilaisia muotovertailuja [Finnish-Samoyed form comparisons]. In Nobufumi Inaba, Jorma Luutonen, Arja Hamari & Elina Ahola (eds.), *Juuret marin murteissa, latvus yltää Uraliin [The roots in the Mari dialects, the canopy reaches to the Urals]*, 289–300. (Mémoires de la Société Finno-Ougrienne 270.) Helsinki: Société Finno-Ougrienne.
Salminen, Tapani. 2023. Suomalais-samojedilaisia sanavertailuja [Finnish-Samoyed word comparisons]. In Sampsa Holopainen, Jeongdo Kim & Niklas Metsäranta (eds.), *Elämä ja etymologia [Life and Etymology]*, 377–399. Helsinki: Books on Demand.
Sammallahti, Pekka. 1974. *Material from Forest Nenets.* (Castrenianumin toimitteita 2.) Helsinki.
Sammallahti, Pekka. 1988. Historical phonology of the Uralic languages. In Denis Sinor (ed.), *The Uralic languages: description, history and foreign influences*, 478–554. Leiden: Brill.
Sammallahti, Pekka. 1998. *The Saami Languages.* Kárášjohka: Davvi Girji.
Siegl, Florian. 2008. A note on personal pronouns in Enets and Northern Samoyedic. *Linguistica Uralica* 44. 119–130.
Siegl, Florian. 2012. Yes/no questions and the interrogative mood in Forest Enets. In Tiina Hyytiäinen, Lotta Jalava, Janne Saarikivi & Erika Sandman (eds.), *Per Urales ad Orientem: Iter polyphonicum multilingue*, 399–408. (Mémoires de la Société Finno-Ougrienne 264.) Helsinki: Société Finno-Ougrienne.
Siegl, Florian. 2013. *Materials on Forest Enets, an indigenous language of Northern Siberia.* (Mémoires de la Société Finno-Ougrienne 267.) Helsinki: Société Finno-Ougrienne.
Simoncsics, Péter. 1998. Kamassian. In Daniel Abondolo (ed.), *The Uralic languages*, 580–601. London/New York: Routledge.
SkWb = Donner, Sirelius & Alatalo 2004.
Sorokina, I[rina] P. 2010. *Ėneckij jazyk* [The Enets language]. Sankt-Peterburg: Nauka.
Sorokina, I[rina] P. & D[arija] S. Bolina. 2001. *Slovar' ėnecko-russkij i russko-ėneckij* [Enets-Russian and Russian-Enets dictionary]. Sankt-Peterburg: Prosveščenie.
Sorokina, I[rina] P. & D[arija] S. Bolina. 2005. *Ėneckie teksty* [Enets texts]. Sankt-Peterburg: Nauka.
Sorokina, I[rina] P. & D[arija] S. Bolina. 2009. *Ėneckij slovar'* [Enets dictionary]. Sankt-Peterburg: Nauka.
SW = Janhunen 1977a.
Szeverényi, Sándor. 2005. Lexical categorization of property concept words in the Samoyedic languages. In Beáta Wagner-Nagy (ed.), *Mikola-konferencia 2004*, 75–93. Szeged: SzTE Finnugor Tanszék.
Terent'ev, V[ladimir] A. 1999. Drevnejšie tjurksko-samodijskie kontakty [Ancient Turkic-Samoyed contacts]. *Journal de la Société Finno-Ougrienne* 88. 173–200.
Tereščenko, N[atal'ja] M. 1965. *Nenecko-russkij slovar'* [Nenets-Russian dictionary]. Moskva: Sovetskaja Ėnciklopedija.
Tereščenko, N[atal'ja] M. 1973. *Sintaksis samodijskix jazykov* [The syntax of the Samoyed languages]. Leningrad: Nauka.
Tereščenko, N[atal'ja] M. 1979. *Nganasanskij jazyk* [Nganasan language]. Leningrad: Nauka.
Trosterud, Trond. 2006. *Homonymy in the Uralic two-argument agreement paradigms.* (Mémoires de la Société Finno-Ougrienne 251.) Helsinki: Société Finno-Ougrienne.

UEW = Rédei 1986–1991.
Urmanchieva = Urmančieva, Anna. 2016. Analitičeskie formy «modal'nogo inferentiva» v lesnom ėneckom [Analytical forms of "modal inferential" in Forest Enets]. *Linguistica Uralica* 52. 122–140.
Verbov, G[rigorij] D. 1973. Dialekt lesnyx nencev [The dialect of the Forest Nenets]. In *Samodijskij sbornik*, 3–190. Novosibirsk: Akademija Nauk SSSR, Sibirskoe otdelenie.
Volkov, O[leg] S. & I[van] A. Stenin. 2019. Andativ i ventiv v jazykax Sibiri: k tipologii glagol'noj orientacii [Andative and ventive in the languages of Siberia: towards a typology of verbal orientation]. *Linguistica Petropolitana* 15. 289–319.
Wagner-Nagy, Beáta. 2001. *Die Wortbildung im Nganasanischen.* (Studia Uralo-Altaica 43.) Szeged: SzTE Finnugor Tanszék.
Wagner-Nagy, Beáta. 2019. *A grammar of Nganasan.* Leiden/Boston: Brill.
Ylikoski, Jussi. 2017. On the tracks of the Proto-Uralic suffix *-*ksi* – a new but old perspective on the origin of the Mari lative. *Journal de la Société Finno-Ougrienne* 96. 369–419.