

schung zu folgenden Vergleichspaaren, in denen jeweils auf der einen Seite indogermanische, auf der anderen Seite uralische Sprachen vertreten sind: a) germanisch-ostseefinnische Sprachkontakte, b) germanisch-lappische Sprachkontakte, c) baltisch-ostseefinnische Sprachkontakte, d) frühe slavisch-ostseefinnische Sprachkontakte, e) baltisch-lappische bzw. balto-slavisch-lappische Sprachkontakte sowie f) Kontakte des Finnisch-Ugrischen mit dem Indo-Iranischen (Koivulehto 1999, 8-12). Besonders das letztgenannte Thema hat ihn in den letzten Jahren zunehmend fasziniert – wie er auch sonst im Laufe seiner wissenschaftlichen Karriere immer weiter und tiefer in Raum und Zeit eingedrungen ist. Er verwirft die Hypothese einer indogermanisch-uralischen Urverwandtschaft, denn „alle Lautentsprechungen erklären sich als Ergebnisse einfacher Substitutionsvorgänge“ (ebd., S. 13). Hier, wie überhaupt bei der Untersuchung der Reflexe von indogermanischen „Laryngalen“ in Lehnwörtern (Koivulehto 1991), im Bereich der germanischen und finnisch-lappischen Lautgeschichte sowie in der Diskussion über die uralische Urheimat hat der Jubilar Ergebnisse erzielt, die nicht nur für die Sprachwissenschaft, sondern auch allgemein für die Erhellung der prähistorischen Zeit von größter Bedeutung sind.

#### Literatur

- HYP (1997) = Helsingin yliopisto. Professorinmatkella 1918-1996. Hg. von Veli-Matti Autio. Helsinki.
- KKO (2002) = Kuka kukin on: henkiltietoja nykypolven suomalaisista = Who's who in Finland. Helsinki.
- Koivulehto, Jorma (1971): 'Jäten' in deutschen Mundarten. Wortgeographisch-etymologische Untersuchungen. Helsinki (= *Annales Academiae Scientiarum Fennicae* B 170).
- Koivulehto, Jorma (1991): Uralische Evidenz für die Laryngalthese. Wien (= Sitzungsberichte der Österreichischen Akademie der Wissenschaften, Philosophisch-historische Klasse 566/Veröffentlichungen der Kommission für Linguistik und Kommunikationsforschung 24).
- Koivulehto, Jorma (1999): Verba mutata. Quae vestigia antiquissimi cum Germanis alisque Indo-Europaeis contactus in linguis Fennicis reliquerint. Ed. Klaas Ph. Ruppel. Helsinki (= *Mémoires de la Société Finno-Ougrienne* 237).
- Koivulehto, Jorma (2004): Mistä lähtien suomea Suomessa? In: Miia Pesonen/Harri Westermarck (Hg.): *Suomen kansa – mistä ja miksi?* Helsinki (= *Studia generalia*), S. 81-94.
- Nikkilä, Osmo (1999): Jorma Koivulehto, 65 Jahre. In: Koivulehto (1999), S. XVII-XXV.
- SP (2000) = Suomen professorit – Finlands professorer. Hg. von Veli-Matti Autio. Helsinki.

Ante Aikio

## An essay on substrate studies and the origin of Saami<sup>1</sup>

### 0. Introduction

Lapland is known to have been inhabited for over 9,000 years. However, the history of the Saami – or Lapps, as they used to be called – as a distinct ethnic group is much shorter. It is generally agreed that the first settlers of Lapland spread via the Norwegian coast from Western Europe at the end of the last Ice Age when inland Scandinavia was still covered by the retracting ice sheet. These people were not the linguistic ancestors of the Saami. The Uralic original home must have been located far from Western Europe, somewhere in the forest zones of present-day Russia, and the breaking up of established proto-languages happened significantly later than the end of the Ice Age. Thus, the languages spoken in the European north prior to the Uralic and Indo-European expansions will forever remain unknown to science.

Even though the Saami language is intrusive in its present territory, the background of the Saami people is somewhat different. Studies in population genetics have confirmed the old hypothesis of the “racially” distinct character of the Saami; to put this in modern terminology, the Saami people form a genetic outlier in the European context. Significantly, this outlier status also holds in comparison to the geographically adjacent and linguistically related Finnic peoples. Thus, even though the Saami are linguistically Uralic, they have inherited a significant genetic component from the non-Uralic first settlers of Lapland.

The only way to reconcile the linguistic and genetic facts is to assume that language shifts have taken place. The linguistic lineage leading to Proto-Saami must have split off from Proto-Uralic somewhere outside Lapland. At some later stage this language began to spread northward and was ultimately adopted by the subarctic hunter-gatherers in the Fennoscandian north, pushing their original languages to extinction. While these languages did not survive the course of prehistory, they have probably left a trace of themselves behind: general contact linguistic principles predict that Saami must have adopted a substrate component during its expansion.

<sup>1</sup> I wish to thank Pekka Sammallahti and Jussi Ylikoski for comments on an earlier version of this paper.

This paper has two purposes: first, to outline a critical methodological framework for palaeolinguistic substrate studies, and second, to analyze and date the substrate component in Saami. Linguistically Saami provides an ideal case for such a study. The development of the Saami languages has been reconstructed in detail, which allows for a rigorous and exact stratification of linguistic material. The ultimate aim of the analysis is to answer a fundamental question of Saami ethnic history. As it is reasonable to define ethnic groups according to their language, a successful dating of the "Saami language shift" will reveal when Lapland became Lappish.

### 1. Notes on research history

Early research has often hinted at the presence of substrate in Saami. However, there exists no solid linguistic treatment on the subject, and the hypothesis of a language shift was originally motivated by purely extralinguistic circumstances, i.e. the assumption of the "racially" deviant character of the Saami. For instance, the eminent lappologist K. B. Wiklund found it unthinkable that the Saami and Finnish peoples could share a common origin:

es dürfte ganz undenkbar sein, dass zwei von einander anthropologisch so weit verschiedene völker wie die jetzigen finnen und lappen während einer verhältnismässig so kurzen zeit wie etwa zwei tausend jahren ... sich aus einem einheitlichen volke hätten entwickeln können (K. B. Wiklund 1896, 10).

He assumed that the Saami had originally spoken an unknown language, which he confusingly termed "Protolappisch". According to a rival theory the Saami had originally been a Samoyed tribe which had switched its language to a Finno-Ugric one (e.g. K. Nielsen 1913, 204-205). The idea of a Saami-Samoyed connection was also non-linguistic in origin, going back to times which predate the birth of modern science. Both theories were similar in that they interpreted the Saami as "racially" non-Finno-Ugric and assumed that they have abandoned their original language in favor of a Finno-Ugric one.

The hypothesis of an unknown substrate language has later been supported by referring to the large number of unetymologized Saami words, but this line of argumentation has not progressed beyond uncritical compilation of word lists (see e.g. T. I. Itkonen 1948 I, 165-167). On the other hand, there have been attempts to substantiate the Saami-Samoyed connection with etymological compa-

rison, but modern research has shown most of these etymologies as erroneous (Helmski 1996). Against this background it is not altogether surprising that during the latter half of the twentieth century the language shift theory became unpopular among specialists in Saami linguistics. Paavo Ravila initially cautiously supported the Samoyed or "Protolappisch" background of the Saami (1935; 1936), but later concluded that there is no linguistic evidence suggesting that the Saami had earlier spoken another language (1957). Major Lappologists have later either doubted the existence of substrate (E. Itkonen 1961, 99; Sammallahki 1995, 151) or considered the subject impossible to study (Korhonen 1981, 26); recently the possibility of substrate influence has been heatedly denied by Lars-Gunnar Larsson, who states that "not one single linguistic circumstance in the Saami language supports this assumption of a language shift", and that "it is the task of linguistics to deal with linguistic matters" (2001, 241).

It must be noted that the prevailing skeptical view is based on *argumentum ad ignorantiam*: the absence of positive evidence of substrate provides no proof of its nonexistence. While it is now easy to criticize Wiklund's ideas of being based on obsolete racial speculations, the quote above qualifies as an overstatement of the opposite kind. Because critical linguistic methods have not been applied in the search for substrate in Saami, the claim that no substrate exists can be dismissed as axiomatic. Moreover, it must be kept in mind that while Wiklund's theory of the origin of the Saami was clearly wrong in all the details, its basic assumption of a language shift has retained its plausibility in the face of modern genetic research. Thus, there is a sound reason to expect the existence of substrate in Saami even in the absence of any confirming evidence for the time being; this has recently been noted also by Carpelan and Parpola (2001, 78). Larsson is, of course, quite right in his second point: linguistic claims must be based on linguistic evidence. The only way to avoid the pitfalls of earlier research is to set up a linguistic method which can reliably detect the presence of unknown substrate.

### 2. A framework for palaeolinguistic substrate research

The study of loanwords is the basic method for detecting prehistoric language contacts. However, in palaeolinguistic substrate studies this standard method faces a dead end: the extinction of the unknown languages has created an etymological event horizon beyond from which no positive evidence can be retrieved. Because of this problem the possibility of proving contacts with un-

known languages has often been viewed with methodological skepticism. However, this agnosticism is unfounded, as various case studies allow one to draw conclusions on the nature of substrate influence, and these conclusions in turn yield typological predictions of what to expect in the case of an unknown substratum.

A number of well-examined cases, including e.g. the Saami substrate in Bothnian Finnish and the Finnic substrate in the northern dialects of Russian, suggest that the basic component of substrate is lexical. In addition, interference may also occur on the level of phonology and (morpho)syntax, but because it is often difficult to satisfactorily demonstrate influence in these domains even in the case of known substrate languages, they will be left outside consideration in the present paper.<sup>2</sup> Substrate vocabulary tends to concentrate in certain semantic fields, including e.g. topographic nouns, words related to the biological world, and terms for culture-specific items and concepts, but the influence is not necessarily restricted to these domains. In addition to loanwords proper substrate influence is characterized by extensive borrowing of place names, especially topographic names. (See Saarikivi 2000 and in press (a) for more detailed discussion.)

Substrate influence from unknown languages has recently been a subject of discussion in Indo-European linguistics. Salmons (1992, 267) has summarized criteria (originally deriving from Edgar Polomé) for the identification of probable substrate words in the northwest Indo-European languages. Furthermore, he notes that substrate hypotheses are more plausible when they are based on systematic patterns instead of isolated lexical items. Because of this it is actually more fruitful to set criteria that apply to the entire lexicon; this makes it possible to identify typological features in the material without requiring that all indicators of substrate origin coincide in the case of individual lexical items. Modifying *mutatis mutandis* the criteria summarized by Salmons, the following test can be devised. If each of the following four criteria is fulfilled, the case for substrate in the lexicon is strong:

- a) The amount of vocabulary of unknown origin is significant, preferably comprising at least several hundred basic stems.
- b) Words of unknown origin concentrate in semantic fields which are especially liable to substrate influence.

2 According to a widespread notion substrate influence is predominantly phonological and syntactic and involves little lexical loans, but this view has been shown to be mistaken (Saarikivi 2000).

- c) The words in question frequently show non-native structural features, e.g. phonemes or phonotactic combinations which are of secondary origin.
- d) Some of the words show irregular sound correspondences between languages or dialects, indicating separate borrowings which reflect different sound substitution strategies or distinct but related source idioms.<sup>3</sup>

The approach employed by Salmons and many other Indo-Europeanists can also be improved on one point: place names must be analyzed on a par with loanwords proper, as they are especially liable to borrowing during language shifts. Thus, a positive result in the lexical test must further be tested against place names. If there is a significant substratum in the lexicon, it is predictable that toponyms have also been extensively borrowed. This would predictably be manifested as large numbers of topographic place names which are of unknown origin. If the language shift is not recent, major topographical formations are more likely to carry substrate names, because macrotoponyms are more resistant to replacement.

This kind of onomastic analysis requires a good understanding of the stratification of the nomenclature. Luckily, in the case of Saami the task is relatively simple due to the geographically peripheral location. Inland Lapland must have been inhabited by the Saami alone from the time the last "Palaeo-Laplandic" languages disappeared up to the recent colonization of the area by Scandinavians, Finns and Russians. Thus, only two significant strata should occur in this area: autochthonous names and substrate names. This hypothesis indeed seems to be reflected in the Saami nomenclature in a predictable way. A huge majority of Saami place names are completely transparent, but the ones that are not are often completely opaque, unique and denote major topographical formations. On the Arctic Coast the situation is somewhat more complicated. The coastal contact zone between Saami and Scandinavian developed already in prehistoric times predating the Viking Age, and this is reflected as a complex stratification of borrowed place names. However, as far as Scandinavian toponyms can be weeded out of the material, the remaining names can be analyzed from the same premises as in the inland areas.

Because the method outlined above operates with the exclusion of alternative etymological explanations, it is certain that some non-substrate items will be-

3 Naturally, the test can be applied only when the phonological and lexical stratification of the language(s) in question is well understood. If there are serious gaps in reconstruction or basic etymological research, claims of secondary phonological features or of unknown origin are on a shaky basis.

come erroneously included in the research material. Nevertheless, this should not have a significant impact on the results. If probable substrate words and place names can be shown to occur in masses, there is no serious alternative to the assumption of a language shift.

### 3. Substrate vocabulary and nomenclature in Saami

The method outlined in the previous chapter can now be applied to the Saami lexicon and nomenclature. For the purpose of this paper it is sufficient to conduct a limited study based primarily on North Saami materials; an in-depth analysis of all the relevant data must be left for future research. There is no reason to assume that concentration on North Saami would somehow distort the overall picture. First, North Saami is located in the geographical core of Lapland, and if the existence of a non-Saami substrate can be demonstrated in this area, the same most probably applies to the rest of Lapland as well. Second, even a most superficial inspection of the existing materials reveals that the characteristics discussed below are typical of the lexica and nomenclature of the other Saami languages, too.

The following list contains North Saami words which have no known etymology and no acceptable cognates in the other Uralic branches. The words have been chosen from three semantic fields which can a priori be judged as likely substrate domains. It should be noted that the distribution of the words shows great variation, ranging from a pan-Saami one to attestation in only one dialect area of North Saami. This kind of distributional variation is typical of substrate vocabulary, e.g. of the Saami loanwords in the dialects of Finnish and Karelian (Aikio 2002) and of the Finnic loanwords in north Russian dialects (Saarikivi 2000).

#### Animals and animal body parts, including:

- birds: *állat* 'snow bunting', *biehkan* 'rough-legged buzzard', *bovttás* 'puffin', *bupmálas* 'northern fulmar', *cagan* 'Eurasian oystercatcher', *čiekkás* 'black guillemot', *čoočču* 'bovto- 'snipe', *šiehta* 'fiehttag- 'scaup', *širon* 'rock ptarmigan' (< PS \**kierun*), *goašst* 'merganser', *gnoxxat* 'Lapland jay' (Finn. dial. *kuusanka* etc., which has erroneously been considered cognate, is obviously a borrowing from Saami), *hávda* 'eider' (? << \**ávda*, cf. SaaSk *avdá*, note also Old Norse *aér* 'eider' – due to the sound correspondence between Saami and Norse the word cannot be a direct loan in either direction), *šiesmi* 'young swan', *liidnu* 'eagle owl', *loađgu* 'short-eared owl', *skuoŋŋ* 'owl, esp. snowy owl', *soymir* 'gull (Larus fuscus?)', *viros* 'sandpiper'

- fish: *beahcet* 'fish tail', *cuolppa* 'fish meat', *čuovža* 'powan', *dápmot* 'brown trout', *diktus* 'haddock', *ásovvi* ~ *ásovvi* 'female salmon with roe', *golis* 'giant pike', *gudjur* 'sea trout which has wintered in fresh water', *grovžur* 'trout', *gunža* 'sea trout', *loahkka* 'torak', *soarvvi* 'grayling', *šákka* 'capelin' (< PS \*šápšš), *šuoŋja* 'giant shark (basking shark?)', *valas* 'red char (living in the ocean)', *veakst* 'fin' (< PS \*veapsš), *vuorru* 'salmon which winters in fresh water'
- marine mammals: *áičnu* 'grey seal', *áičču* 'bearded seal', *áičču* 'common seal', *buovjja* ~ *buovjag-* 'beluga', *dealljá* 'harp seal', *dearvu* 'grey seal', *gáhtir* 'seal's flipper', *jeagis* ~ *jeagis* 'bearded seal', *šiepmá* 'young seal', *morša* 'walrus', *noarvi* 'seal', *njuorju* ~ *njuorjju* 'seal' (assim. << \**nuorjjo*, cf. Saal *nuárju*), *oaiču* ~ *oaičču* 'ringed seal', *riehkka* 'a middle-sized mottled seal(?)', *roahkka* 'common seal', *rohka* 'full-grown male seal', *skuoŋga* 'baleen', *skuoŋgir* ~ *skuoŋgun* : *skuoŋgun-* 'ethmoid bone', *viekst* 'young common seal'
- land mammals: *báikta* 'castor gland (of a beaver)', *scent gland (of an ermine)*, *čearvris* 'otter', *fuočču* 'game animal', *gákkur* 'she-wolf, she-bear' (< PS \*kávšš), *geahtá* 'wolverine', *gumpe* 'wolf', *guohsi* 'one-year old beaver', *njállá* 'arctic fox', *ođgi* 'young fox', *ruomas* 'wolf', *sáhpán* 'mouse', *váisi* 'wild animal' (~ SaaS *væijsije* 'beast of prey, elk')

**Topography:** *ája* : *áđjag-* 'spring', *áđju* 'heath, moor', *áđga* 'grassy terrain along a river', *balasa* 'frost peat mound', *bákti* 'cliff, rock', *bársi* 'isolated mountain', *beasstá* 'pass (in mountains); strait which dries up during summer' (< PS \**peackš*), *bovccis* 'side channel of a river', *bovdna* 'large tussock, hillock in a bog', *cahca* ~ *čahca* 'narrow pass (in mountains, between bogs, etc.)', *ceavnnil* (plurale tantum) 'impassable terrain', *čavil* : *čavilg-* 'wild tract', *čievra* 'gravel', *čierri* 'gravelly ground', *čoarvá* 'ford; place where a river runs underground', *čunu* : *čuadno-* 'fine sand', *čuodjá* 'long and narrow bay of a lake', *dievvá* 'hill', *fielbmá* ~ *vielmis* 'small but deep river, deep place in a small river', *geađgi* 'stone, rock', *geavvips* 'big, foaming rapids, waterfall', *gieva* : *gievng-* 'boghole', *giezzi* 'short river between two lakes', *iŋku* 'shady place, place where the sun does not shine', *šalvi* 'short stretch of smooth water between two rapids', *šargga* 'open water in a lake', *šiertá* ~ *šiertá* 'large, round mountain', *šuggi* 'depression (in terrain)', *šuoŋva* 'scree', *šuoŋvji* 'peat', *šuoŋvji* ~ *šuoŋvji* 'marshland in a low valley in the uplands', *šuoŋva* 'scree', *šuoŋvji* 'hillside', *šuoŋvi* ~ *šuoŋvi* 'river head', *šuoŋvi* 'birch forest surrounded by bogs' (< PS \**mēršššš*, cf. Saal *moorššš*), *šuoŋvji* 'small rapids', *šuoŋvji* 'shoal, shallows', *šuoŋvji* 'treeline', *šuoŋvji* 'flat stone, stone slab' (< PS \**rāpššš*), *šuoŋvji* (plurale tantum) 'rough, rocky terrain which hinders travel', *šuoŋvji* 'high, barren mountain', *šuoŋvji* ~ *šuoŋvji* 'flat seashore with a clay bottom', *šuoŋvji* 'valley up in the mountains which is difficult to access', (\**šuoŋvji* 'rocky mountain' (in North Saami preserved only in place names), *šuoŋvji* 'place where there has been a forest fire', *šuoŋvji* 'hole, pit; depression, hollow, (river) valley', *šuoŋvji* 'narrowing (e.g. in a river)', *šuoŋvji* 'narrowing in a gorge', *šuoŋvji* ~ *šuoŋvji* 'heath', *šuoŋvji* 'landing place, boat-shore', *šuoŋvji* 'narrowing in a gorge', *šuoŋvji* 'dense thicket', *šuoŋvji* 'heap of rocks or rocky slope near the seashore', *šuoŋvji* 'flat, wet bog', *šuoŋvji* 'dense compound *šuoŋvji* 'willow thicket', cf. SaaSk *váissš* 'thicket with young pine trees'), *šuoŋvji* 'hillside', *šuoŋvji* 'fjord, long and narrow bay of a lake'

Snow, ice and weather: *adajjo-* 'to get covered by snow', *asikkas* 'sheet ice' (< PS \**ētkēs*), *áidnet* 'to snow thinly' (a derivative, cf. SaaSk *aim* 'thin, freshly fallen snow' < PS \**ājine*), *biegga* 'wind', *bihci* 'rime', *bulži* 'coating of ice', *ceavvi* 'hard, compact snow', *chiehka* ~ *chiehtá* 'layer of clouds, storm-cloud; mist or haze over the sea', *coakci* 'foothold (esp. in snow)' (< PS \**coapce*), *cuokca* 'naturally formed bridge of ice or snow; patch of perpetual snow up in the mountains in summer' (< PS \**cuopce*), *cuoyju*: *cuggo-* 'strong snow-crust', *čahki* 'hard lump of snow (which e.g. sticks on an animal's foot)', *čavttas* 'small snowdrift', *čearga* 'snowdrift', *čovvga* 'light' (< PS \**čuokvė* or \**čuovpė*), *dáiki* 'weather', *dierpmis* 'thunder; thunder god', *dohpa-* 'to stick (of snow)', *duolli* 'frost on the ground', *gocáhtá* 'calm weather', *guovla* 'overhanging snowdrift' (< PS \**kuoplė*), *jassa* 'patch of perpetual snow up in the mountains in summer', *liehmu* 'mild weather in winter', *muovla* ~ *muovhla* 'deep snow; weather when one sinks into snow', *njeaŋga-* 'to storm with snow (so that tracks, way, etc., get covered)', *oakti* 'rain shower', *rákšu* 'heavy rain in cold weather', *roavru* 'ice with a hollow space under it', *ruokpa* 'lack of snow', *rusta* 'frost mist' (< PS \**ruste*), *seavjás* 'granular snow', *seavdnjaj* 'darkness', *sievla-* 'to sink and get stuck in soft snow', der. *sievlla* 'soft snow' (< PS \**sieplė*), *soarvi* 'slush', (\**svobhat* 'bridge of snow (e.g. over a gap)' (in North Saami preserved only in place names), *svonjar* 'beam of light', *suovvi* 'wet and sticky snow', *šuhci* ~ *šuhci* 'hard-frozen rime on trees', *veššv* 'frosty wind', *váhččv* 'calm after a storm', *vuožži* 'water on ice'

Saami easily fulfills the first two criteria for substrate influence, i.e. quantity and semantic distribution: there are very many words of unknown origin in semantic fields which are susceptible to substrate influence. It must be stressed that the list above is far from exhaustive. One could easily multiply the number of items by adding etymologically opaque plant names, words for primitive hunting and fishing equipment and methods, and butchering terminology (including body parts). Also the abundant etymologically opaque vocabulary related to reindeer has been left out; while reindeer herding on the large scale is a relatively late innovation among the Saami, many of the lexical roots in this field derive from earlier terminology related to the hunt of wild reindeer. Furthermore, the number of potential substrate items would substantially increase if one also were to take words which are not attested in North Saami into account.

As to the third criterion, many of the words in the list show phonotactic structures which suggest non-native origin. These features can be divided into four groups.<sup>4</sup>

4 The treatment is based on established reconstructions and sound laws, which are not separately referred to. Detailed presentations of Saami historical phonology are provided by Korhonen (1981, 76-200) and Sammallahti (1998, 181-202).

- 1) Features which could in theory be projected back to Uralic reconstructions, but which do not appear in shared Uralic vocabulary and are thus likely to be of non-native origin. Such features which look most unusual from a Uralic perspective include three-consonant clusters (e.g. *gocáhtá* 'calm weather' < PS \**koalktē*, *suovka* 'dense thicket' < PS \**suovkkē* or \**suomkkē*, *jiervá* 'large, round mountain' < PS \**jiervtē*, *gumpe* 'wolf' < PS \**kumppti*), geminate sonorants (*ličnu* 'eagle owl' < PS \**limnō*, *dáipmot* 'brown trout' < PS \**táimmōtktē*), two *r*'s or dental affricates in the same word (*roavru* 'ice with a hollow space under it', *cahca* 'narrow pass', *coakci* 'foothold', *cuokca* 'natural bridge of ice or snow'), and various consonant clusters which are unattested in shared Uralic vocabulary, including \**ls* (*balsa* 'frost peat mound', *gocáhtá* 'merganser'), \**lč* (*bulži* 'coating of ice'), \**sm* (*husmi* ~ *huspi* 'river head'), \**sn* (*rusta* 'frost mist' < \**ruste*), \**pl* (*guovla* 'overhanging snowdrift' < \**kuoplė*, *sievla-* 'to sink in soft snow' < \**cuopce*), \**čk* (*asikkas* 'sheet ice' < \**coapce*, *cuokca* 'natural bridge of ice or snow' < \**čuokvė* or \**čuovpė*), *čk* (*asikkas* 'sheet ice' < \**ētkēs*), \**kv* or \**pv* (*čovvga* 'light' < \**čuokvė* or \**čuovpė*). Consonant clusters which are anomalous in suffixal position in nouns can also be included in this group, i.e. \**pp* (*dállat* 'snow hunting' < \**áilēppē*), \**st* (*maras* 'birch forest surrounded by bogs' < \**mērāstē*), \**lk* (*čavil* 'wild tract' < \**čēvēlčkē*) and \**ŋk* (*guovssar* 'Lapland jay' < \**kuokskŋkē*). Also certain vocalic combinations, in particular structures that would go back to a Pre-Saami sequence of a long vowel + a single consonant + a non-high vowel (e.g. *bupmálas* 'northern fulmar' and *bihci* 'rime' which would require the reconstruction of earlier \**pimálas* and \**pīčá*, respectively) belong to this category of unlikely Uralic words. Most probably, the majority of these phonotactic combinations do not go back to old strata of lexical innovations but instead became established during a relatively late phase of borrowing.
- 2) Features which cannot be projected back to Uralic reconstructions but which may derive from a phase of Pre-Saami anteceding the reorganization of the Proto-Saami vowel system. The only feature that can be unambiguously counted in this group is the secondary sibilant *š* (*gákká* 'she-wolf, she-bear', *morša* 'walrus', *šákká* 'capelin', *šuoŋja* 'giant shark', etc.).
- 3) Features which can only derive from the phase of Proto-Saami after the rather radical reorganization of the Uralic vowel system. New structural types which became possible as a result of these vowel shifts include the following present North Saami vowel combinations: a) *á-á*, *ea-á* and *oa-á*, b) bisyllabic stems ending in *-á* (a couple of Uralic words have irregularly developed into Saami *á*-stems, but these cases are very rare), c) the vowel combination *á-u* (except in Uralic \**u*-stems) and d) the first syllable vowels *i* and *u* and the vowel combinations *ie-i*, *ie-u* in stems with Proto-Saami consonant clusters in medial position. (Note that all the secondary vowel combinations can appear in derivatives based on old Uralic stems, though.) As for the words in group c) and d), one could in theory project them back to earlier levels of reconstruction, but this would force one to postulate forms which violently break the rules of Uralic phonotaxis: e.g. the words *ákkčv* 'common seal', *gázzv* 'short river between two lakes', *ličnu* 'eagle owl', *husmi* ~ *huspi* 'river head', and *rusta* 'frost mist' would have to be derived from the Pseudo-Uralic reconstructions \**ákkō*, \**kēncčá*, \**lŋmo*, \**lŋsma* and \**rŋsni*, which appear absolutely impossible in the light

of Uralic phonotaxis. As for words in groups a) and b), one could reconstruct nothing, because they show innovative vowel combinations which cannot even in principle be traced back to Pre-Saami.

4) Features which cannot even derive from common Proto-Saami. These include the initial cluster *sk-* (*stuogga* 'baleen', *stuoggir* 'ethmoid bone', *skuolffi* 'owl'), the phoneme *f* in non-initial position (*skuolffi* 'owl', *uffir* 'heap of rocks or rocky slope near the seashore'), and the unvoiced nasal /*M*/ (*tietnu* 'mild weather in winter', *-hm-* stands for /*M*/ in the North Saami orthography). None of these features is attested in Kola Saami. On the basis of Scandinavian loanwords it can be inferred that both *sk-* and *-f-* were adopted in the west during the early separate development of the Saami languages, but never spread to Kola Saami. These areal features thus emerged in a phase when Proto-Saami began to diverge into dialects anticipating the modern Saami languages.

About 40 % of the vocabulary listed above contains features of phonology in categories 2)–4), and even many of the remaining 60 % show unusual phonotactic combinations which suggest late origin (i.e., category 1). Also semantics can imply loan origin. For instance, the word *buovjija* 'beluga' (attested solely in North Saami) could in principle be projected back to Proto-Uralic \**powjik* or \**parjik*, but the referent of this word must have been unknown to the linguistic ancestors of the Saami before they reached the coast of the Arctic Ocean. Evidently, the "Uralic" phonology of this word is coincidental and it must have been adopted from an extinct substrate language at roughly the same time as the phonologically clearly non-Uralic words *morša* 'walrus' and *šuoŕja* 'giant shark'.

Finally, a part of the listed vocabulary shows irregular sound correspondences within Saami, which fulfills the fourth criterion for substrate influence. A couple of words show inexplicable irregularities between dialects, such as *állat* 'snow bunting' ~ dial. *hálpri* 'id.', *reaz'si* ~ *reakčá* 'flat seashore with a clay bottom'. One also finds probably related words in the other Saami languages which nevertheless cannot be traced back to a common Proto-Saami form or explained as later borrowings between the differentiated idioms. Examples include:

Saam *állat* 'snow bunting' (< PS \**állēppē*) ~ Saal. *albba* (< PS \**ēlpe*) || Saam *bársi* 'isolated mountain' (< PS \**pāršē*) ~ Saas *bársje* 'steep edge of a mountain top' (< PS \**pāršē*) || Saam *bovccis* 'side channel of a river' (< PS \**pokeēs*) ~ Saas *páđu'tes* 'id.' (< PS \**pokeēs*) (cf. also Finn. *puolás* 'id.', which matches neither of these forms) || Saam *častias* 'small, hard snowdrift' (< PS \**česēs*) ~ Saas *fihtese* 'snowdrift (in summer)' (< PS \**čotēs* or *čutitēs*) || Saam *duollu* 'frost on the ground' (< PS \**tuolō*) ~ Saas *taal* 'id.' (< PS \**tälēk*) ~ Saas *doelmes* 'id.' (< PS \**tuolmēs*) (note also Old Norse *þeli* 'frozen ground', which matches none of these forms) || Saam *fuodđu* 'game animal' (< PS

\**vuodđjē*) ~ Saas *vđđrenasse* ~ *vđđrenasse* 'beast of prey, esp. wolf' (< PS \**vuorum*/l<sup>essē</sup>) || Saam *ilku* 'shady place' (< PS \**irkō*) ~ Saas *ierhtie* 'id.' (< PS \**eaktkē*) ~ Saal *ipio* 'id.' (< PS \**iptō*) ~ Saas *ōhkas* 'id.' (< PS \**ētkōs*) || Saam *jassa* 'patch of snow in summer' (< PS \**jēsē*) ~ Saas *jēcc* 'id.' (< PS \**jēccē*) || Saam *jiesmi* 'young swan' (< PS \**jiesmē*) ~ Saas *jie* 'desm' 'id.' (< PS \**jiešmē* or \**jiešmē*) ~ obsolete Swedish Saami \**jeudahr* 'id.' (? < \**jiev-* or \**jiev-*) || Saam *locadgu* 'short-eared owl' (< PS \**locādkō*) ~ Saal *luđđu*, Saal. *loarhtko* 'id.' (< PS \**loarhtkō*) ~ Saal *loarvggo* 'id.' (< PS \**loarvkō*) ~ *loarvkō* 'id.' (< PS \**loarvkō*) ~ Saap *lav'htkuo* 'id.' (< PS \**loarvkō*) || Saam *muovla* ~ *muovhla* 'deep snow' (< PS \**muovlē*) ~ Saas *marhile* 'id.' (< PS \**mēvlē* or \**muovlē*) || Saam *sáhpoh* 'mouse' (< PS \**sáhpoh*) ~ Saap *sachpiek* 'id.' (< PS \**sáhpēkkē*) ~ Saal *sáhpig*, Saak *sáppli* 'egg' 'id.' (< \**sáppli*) || Saam *skuoŕla* 'owl, especially snowy owl' ~ Saap *skuok'la* 'eagle owl' || Saam *snovhka* 'dense thicket' (< PS \**snovkkē* ? < PS \**suomkkē*) ~ Saas *semhtie* 'dense broadleaf forest' (< PS \**semhtē* or \**somhtē*) || Saam *sákká* 'capelin', Saal *sápsá* 'powan' (< PS \**sápsē*) ~ Saak *šább* 'id.' (< PS \**šáppō*) (cf. also Saam *čuovža* 'powan', which would regularly go back to the strikingly similar Pre-Saami form \**čawšil*) || Saam *šuhci* ~ *šuhci* 'hard-frozen rime' (< PS \**šuhci* / \**šuhci*) ~ Saas *šijhtie* 'rime' (< PS \**šicē*) || Saam *valas* 'red char living in the ocean' (< PS \**vēlēs*) ~ Saak *villc* 'lake trout' (< PS \**villcē*) || Saam *váđu* 'thicket' (in *šiedgaváđu* 'willow thicket') ~ Saas *váđušš* 'thicket with young pine trees' (< PS \**vēššō*) ~ Saas *vešjtjie* 'brushwood, thicket' (< PS \**vēšjtjē*)

It is important to note that the words which are inherited from Proto-Uralic do not generally show such unpredictable variation. Quite the contrary, the development of inherited items shows a striking degree of regularity, even though the series of innovations leading from Proto-Uralic to the present-day Saami idioms is complex. This suggests that the majority of these kinds of phonologically varying words belong to late vocabulary strata, and the irregularities are due to parallel borrowings.

Parallel borrowings also provide a natural explanation to the irregular lexical doublets and even triplets in one language:

*čievva* 'gravel' vs. *čier'ri* 'gravelly ground' || *guvzá* 'sea trout' vs. *guovžur* 'trout' vs. *guđjur* 'sea trout which has wintered in fresh water' (cf. also northern Karelian *kumsa* ~ *kumssa* 'lake trout', which does not regularly compare to any of these) || *noarvi* 'seal' vs. *rijuorju* (< \**muorjō*) 'id.' (cf. also Finn. *norppa* 'ringed seal', which does not regularly compare to either of these) || *riehkku* vs. *roahkka* vs. *rohkka* (all species or kinds of seal) || *stuogga* 'baleen' vs. *stuoggir* ~ *skuoggum* 'ethmoid bone'

Such doublets also occur in the other Saami languages:

SaaN *éahki* 'lump of frozen snow (which sticks e.g. on an animal's foot)' (< PS \**ékkē*) ~ SaaS *fjihke* 'id.' (< PS \**ékkē*) vs. SaaS *sihkie* 'weather when ice forms and sticks on the bottom of skis' (< PS \**ékkē*) || SaaN *jargja* 'open water in a lake' ~ SaaS *jarge* 'centre, middle (of a lake, river, coral, floor, hide, etc.)' (< PS \**jérgjē*) vs. SaaS *voernge* 'middle, deepest place of a lake' (< PS \**vuornjē*) || SaaN *guovla* 'overhanging snowdrift' ~ SaaS *goeble* 'depression in a mountain side, where snow remains long during the spring' (< PS \**kuopliē*) vs. SaaS *gdebite* 'precipice; overhanging snowdrift' (< PS \**koapliē*) || SaaN *oakti* 'rain shower' ~ SaaSk *đ'ht* 'id.' (< PS \**o(a)kie*) vs. SaaSk *aanitár* 'storm' (< PS \**áktēr*) || SaaN *orda* 'treeline' ~ SaaK *orrd* 'id.' (< PS \**orie*) vs. SaaK *u'rrd* 'wide and level mountain tract, tundra' (< PS \**urrdē*) || SaaN *ráksu* 'rain in cold weather' ~ SaaS *raassjuo* 'torrential rain' (< PS \**rás(š)ō*) vs. SaaS *raavije* ~ *raavje* 'rain in cold weather' (< PS \**rāvžē* / *rēvžē*).

The word *liehmu* 'mild weather' shows a particularly wide range of irregular correspondents in Saami:

- 1) *liehmu* ~ SaaS *leanmoe* 'id.' (< \**liehMō*); 2) SaaS *leanhmoē* 'id.' (?) (< \**lievMō*); 3) SaaS *iemkies* 'mild (weather)' (< \**iemkēs*); 4) SaaS *ieggēs* 'id.' ~ SaaN *iekkas* 'warm' (< PS \**iepkēs*). Note especially the unvoiced nasal \**M* and the non-Uralic cluster \**mk*.

Due to the secondary phonological features both Uralic inheritance and borrowing from early Indo-European languages can be rejected as a general explanation for the vocabulary discussed above (but, naturally, *some* of the phonologically least marked cases can no doubt be explained in this way). Disregarding *Urschöpfung*, two alternatives to the substrate hypothesis remain: the words with new phonemes or phonotactic combinations are either relatively late borrowings from known languages or sound symbolic and onomatopoeic coinages. Neither of these alternatives is feasible, however. Because both Finnic and Scandinavian loanwords in Saami have for long been the object of detailed study it would be unnatural to assume that hundreds of late loanwords have managed to escape the eyes of etymologists. Onomatopoeia, on the other hand, can perhaps be surmised as an explanation for a couple of bird names, but the fact remains that the vast majority of the words discussed above are certainly not sound symbolic or onomatopoeic. (To be sure, since the 1990s there has been an increasing tendency in Finnish etymological research to explain phonologically marked or varying words as "descriptive" and the variation due to "affect", but the circularity of this line of argumentation is evident.)

It can be concluded that the Saami lexicon shows consistent signs of heavy substrate influence. The remaining task is to test this result against place names. As expected, traces of the lost languages of Lapland can also be uncovered in

the Saami place name system. An inspection of the existing collections of North Saami toponyms reveals an abundance of names of unknown origin. A proportionally large number of such names belong to major topographical formations, which in itself suggests that the names may predate Saami.

A list of monolexical North Saami names of mountains and highlands which have no appellative counterparts in Saami and no known etymology has been compiled below as an example of such material. Name groups where individual names are separated by identifiers (e.g. *Suorra Jeahkkás* 'big J.', *Umma Jeahkkás* 'small J.') have been counted as monolexical. The municipality in which the name occurs is given in parentheses. The Saami identifiers and the abbreviations for the municipalities are explained at the end of the paper. The names occurring in northern Swedish Lapland have been given in the modified standard orthography (preserving certain dialectal features) which is employed in the map series *Fjällkartan* published by Lantmäteriet, Sweden. The list also contains a couple of names which formally compare to some Saami lexemes, but on semantic grounds the connection seems to be coincidental. These kinds of formally comparable stems are given in parentheses ("cf."):

*Aivu* (Rái) || *Suor*, *Umma-Avrrik* (Váh) || *Adnji* (Čoh) || *Suor-Álta*, *Gaska-Álta*, *Umma-Álta* (Bdu) || *Suora Anai*, *Umma Anahás* (Rái) || *Avnui* (Čoh) || *Ávrrui* (Bdu) || *Barbbostat* (Áka) || *Suora Bálak*, *Umma Bálakas* (Nar) || *Bárit* ~ *Bárik* (Ohe) (cf. the next entry) || *Bárrás* ~ *Bárrás* ~ *Bárrák* ~ *Bárrák* (Mál) (cf. the previous entry) || *Beahná* (Gár) || *Beahná* (Gár) || *Beahná* (Muo) || *Beahná* (Ála) || *Buocagas* (Bdu/Mál) || *Buohkágas* (Čoh) || *Boahči* (MVá) || *Boarrá* ~ *Boahrifjálá* (Ála) || *Bierká* (Váh) || *Bihppás* (Bdu) || *Boahči* (MVá) || *Boarrá* ~ *Boahči* (Váh) || *Bierká* (Váh) || *Buohkágas* (Čoh) || *Buošštir* (Mál) (cf. *buošši* 'bad-tempered (of a woman)') || *Cokcu* (Gár/Vaz) || *Covra* (Vaz) || *Covra* (Čoh) || *Častejas* (Guo/Rái) (cf. *častá-* 'to hit, strike') || *Davít*, *Máittit Čavgyá* ~ *Čovgá* (I) (Aná) || *Čáibma* (Ean) || *Čeallu* (Gái) || *Suora-Čierie*, *Ságge-Čierie* (Ean/Rái) || *Čiesti* (Unj) || *Čiesti* ~ *Sesti* (Dav) || *Čillit* (Unj) || *Suor*, *Gaska*, *Umma-Coavrrik* (Váh) || *Cohmmii* (MVá) || *Čohhtit* ~ *Čohhtár* (I) (Fál/Lea) || *Čuosmmir* (Mál) || *Čupmás* (Váh) || *Davju* (Ean) || *Deaigu* (Gáp) || *Dievjá* (Ohe/Aná) || *Diibán* (Gár) || *Dolli* (Váh) || *Dolpi* (Rái) || *Davju* (Ean) (Gái) || *Duorda* (Kart) || *Duorsi* (Rái) || *Altip*, *Fuolip Duorras* (Čoh) || *Gábii* (Lea) || *Altip*, *Lulip Gábhtár* (Čoh) || *Gáiku* (Gár) || *Gálbar* (Rom) (cf. *gáibi* 'caif < Scand.') || *Gárgás* (Báheca) || *Gárgás* (MVá) || *Gásta* ~ (pl.) *Gástáti* (Gái) || *Gealbir* (Mál) || *Gearbbehat* (pl.) (Náv) || *Suor*, *Jorba-Gearbii* (Čoh) || *Geassas* (Gái/Oma) (cf. *geassi* 'to pull') || *Suorra*, *Uhma Giddá* (Dav/Nor) (cf. *giddá* 'spring (Frühling)') || *Giemas* ~ *Giempas-* (Dea) || *Goaddá* (Láh) || *Goahnu* (n) || *Goahntom-* (BeV) || *Goahvit* (Gár) || *Nuorttar*, *Oarjitt Goavvu* (Čoh) || *Gorve* (Gár) || *Govččá* (Fál) (cf. *govčča-* 'to cover') || *Guođggut* ~ *Guođgguk-* (BeV) || *Guohhtár* ~ *Guohhtit* (Dav) (cf. *guohhtá* 'hoe' < Finn.) || *Guoivil* (Čoh) || *Davip*, *Lulip Guohhtit* (Čoh) || *Gurbbeš* (Kár) || *Hoalgir* (Ála) || *Iltis* ~ *Iltas-* (BeV) || *Suorra*, *Uhma Iltis* (Dea/Dav) || *Iškoras* (Kár) || *Suorra Jeahki*, *Umma Jeahkkás* (Lea) ||

*Suorra Jeahkár, Unna Jeahkírás* (Fál) || *Suorra, Unna Jeahkás* (Ean) || *Jerkkon : Jerkkom-* (Dav) || *Jerremás* (Rom) || *Jieiggus : Jigos-* (Guo) || *Jieskátá :* *Jieskátám-* (Ohe) || *Jievdu* (Gár) || *Jolligas* (Čoh) || *Jovri* (Guo) || *Lájká* (Ohe) || *Leantam* (Oma) || *Liesuk* (Čoh) || *Davvi-, Luili-Liiká* (Guo) || *Davit, Luit Máná* (Čoh) || *Máisi* (~ *Máissavárr*) (Bdu) || *Mát'la* (Mál) || *Márkos-, Rás'sa-, Golláta-, Silis-Mát'la* (Ean/Oma) || *Mának* (Váh) (cf. *máná* 'child') || *Márkos* (Oma) (cf. Saal *Márkko*, a highland in the Lule Saami area) || *Suorra-, Duottar-, Jorba-Máváña* (~ *Mávanna*) (Ohe) || *Mollejus* (Guo) (cf. *mol'le* 'to crush, crumble' < *moallu* 'crumb' < Scand.) || *Muhtkomas* (Náv) || *Muhtkírás* (Váh) (cf. *muogir* : *muohkár* 'blackfly') || *Naiggá* (Náv/Rái) || *Nalja* (Čoh) || *Nálgamaš-* (*Nálgamaš*) (Mál) || *Suorra Námá* (~ *Návna* ~ *Nán'na*), *Uha Námimás* (~ *Návimás* ~ *Nánimás*) (Mál) || *Nás'sa* (Guo/Ala) || *Suorra-, Unna-, Ságg-Navgástar* (Ala) || *Njeaiva* (Tra) || *Njullá* (*Nju'lá*) ~ *Njulljá* (Čoh) || *Noalé* (Nar) || *Nohli* (Mál) || *Nuhpi* (Guo/Náv/Ala) || *Nuhpir* ~ *Nuhppir* (Ohe) || *Nuovas, Uha Nuovasas* (Náv/Rái) || *Nussir* (Ala) || *Nussir* (~ *Nussirvárr*) (Ala) || *Nussir* (Fál) || *Suorra Ráppis, Unna Ráppásas* (Ala) || *Raká* (Hám) || *Nuorttat, Oarjitt Rienuk* (Váz) || *Rieváña* (Fál) || *Rivkás* (Guo/Rái) || *Suorra-, Uha-, Bodus-, Ull'o-Roadjá* (Ohe) || *Rosu* (Gár) || *Rovru* (Gán) || *Ruohir* (Ohe) || *Suora Ruvji, Unna Ruvjás* (Rái) || *Salču* (Guo/Kár) || *Salggástak* (Čoh) || *Suorra-Searba, Uha-Searbas* (Dea) || *Silká* (Gár) (cf. *silká* 'silk') || *Sinehat* (pl.) (Ean) || *Sinot* (Gár) || *Skáppas* (Gár) || *Bagit, Vuolit Skázá* (Gár) || *Skiehč(č)aras* (Ohe) || *Skrelá* (Váh) || *Skiełgan* : *Skiełgam-* (Kár) (cf. *skiełgas* 'squint-eyed') || *Suora Skoapmit, Unna Skoapmihás* (Bdu) || *Skooválat* (pl.) (Báhca) || *Stuohki* (Dav) || *Stuorá* (Báhca) || *Snjierriet* (Gár) || *Soančtir* (~ *Soamčtervárr*) (Báhca) || *Solcar* ~ *Soleccar* (Ohe) || *Davit, Lulit Surgu* (Gár) (cf. Saal. *surgu* 'sorrow' < Scand.) || *Šišká* (Váh) || *Suošmir* (Gár) || *Váhčir* (Aná/MVá) (exact parallel in Saal. *Váhijer* ~ *Vearháns* ~ *Veardnjás*) (Guo) || *Yearral* ~ *Virral* (I) (Rom) || *Vidjitt* : *Vidjig-* ~ *Vimit* : *Vimik-* (I) (Báhca) || *Vieksá* (Lea) (cf. *vievssis* : *vieksá-* 'wasp' < Proto-Scand.) || *Virdni* (Ean) || *Virru* (Hám) (cf. Saal. *virru* 'great storm, hurricane') || *Viroš* (Fál) (cf. *viros* 'sandpiper' and also Saal. *virru vuoida-* 'to grease') || *Vuolpoš* (Lea) || *Vuoidul* ~ *Buoidul* (I) (Gát) (cf. *buoidi* 'fat', 'to drive', *vuodja-* 'to swim') || *Bajit, Vuolit Vuokkít* (Gár) (cf. *vuokti-* 'to groan' < \**vuokti-*) || *Vuordnás* (Guo) (cf. *vuortnis* : *vuordnás-* 'oath') || *Vuorji* (Kár/Lea) (cf. *vuorji* 'sparse, thin')

The names in the list above are only meant to serve as examples. Similar opaque names can be found e.g. among names of rivers (cf. *Cuoggá* (Ohe), *Geatnja* (BeV), *Njidgu* ~ *Njedgu* (Unj), *Uvdu* (MVá/Aná)), valleys (cf. *Báváña* (Ala), *Čuita* (Láh), *Dávát* : *Dávag-* (Báhca), *Libmi* (Dav)), coastal islands (cf. *Ávačuohhtu* (Nor), *Bánjal* ~ *Bánjal* (Ski), *Gierbá* (Dav), *Välččit* (Den)) and headlands (cf. *Čorgas* (Dav/Gán), *Dealpu* (MVá), *Iddu* (Ivg), *Vinni* (MVá)). Moreover, loan names are often adopted as hybrids, i.e. by adding a native topographic noun as a generic to a borrowed basis. In fact, the exclusion of biliteral items

evidently leaves out a notable number of likely substrate names, because the Saami name system shows a tendency to avoid monolexical names for certain types of geographical objects. For example, monolexical names of fjords are very rare (but even the rare cases that occur usually seem to be etymologically totally obscure – cf. *Huossu* (Báhca), *Livcu* (Láh), *Stárv* (Ála), *Ušmá* (Nar)). Nevertheless, a number of names with the generic *vuotna* 'fjord' show etymologically opaque qualifiers – cf. e.g. *Duikkávuotna* (Náv), *Itarvuotna* (Ivg), *Iyvuotna* (Gát/Ivg/Oma/Rái/Ski), *Lágesvuotna* (Dav), *Leavdijavuotna* (Lea), *Muorralvuotna* (Muo) (the similarity to *muorra* 'tree' is clearly coincidental), *Skiešvuotna* (Dav), and *Vuorkelvuotna* (MVá). Thus, if one also were to count compound names with an opaque specific, the material would increase dramatically.

It is revealing that many of the etymologically obscure Saami mountain names show similar phonological anomalies as the group of potential substrate loanwords discussed earlier. These include unetymological vowel combinations (e.g. *Čáibma*, *Čiesti*, *Iskoras*, *Jieiggus*, *Rivkkoš*, *Searba*, *Surgu*), final *-á* (this is especially common, cf. e.g. *Beachná*, *Dievjá*, *Lájká*, *Likčá*, *Máisiá*, *Roadjá*, *Vieksá*), initial *sk-* (e.g. *Skázá*, *Skiehč(č)aras*, *Skiełgan*), initial *h-* (*Hoalgir*), the secondary sibilant *š* (*Buoššir*, *Šišká*, *Suošmir*), the unvoiced sonorants */R/*, */M/* and */L/* (*Boahrrá*, *Čohmmi*, *Nohli*), three-consonant clusters (e.g. *Salču* < \**sélččá*, *Solcar* < \**solceččē*, *Avku* < \**ävkkō*, *Lájká* < \**lájkkō*), and other consonant clusters which are not native to Uralic, such as \**js* (*Máisiá*), \**sm* (*Čuosmmir*), \**mn* (*Námndá*), \**ʃ* (*Nálfa*). About 50 % of the names show definitely non-Uralic phonological features, i.e. features in the categories 2)–4) in the phonological stratification discussed above. Even many of the remaining cases contain sound combinations which suggest late origin: e.g., the names *Ráppis*, *Salču*, *Solcar* and *Avku* could in principle be projected back to Pre-Saami, but this would result in the forms \**ráppás*, \**silččo*, \**sulččira* and \**ivkko*, which are entirely improbable from the perspective of Uralic phonotaxis.

Before drawing any conclusions, however, other potential explanations must be considered. The possibility of autochthonous formation provides the most serious alternative. The Saami name system contains also native-based monolexical names, such as e.g. the mountain names *Ailegas* ('sacred'), *Akšu* ('ax'), *Caggi* ('prop'), *Fállát* (pl. 'whales'), *Noaidi* ('shaman'), *Oahpis* ('friend; guide'), *Ruutu* ('cauldron'), *Ruokvadas* ('red thing'), *Vierca* ('ram'), *Adjit* (an obsolete word which probably meant 'thunder god', cf. Saal *Ajjih* 'id.'). Thus, some of the opaque names can simply be based on lost appellatives. However, the idea that all or even most of the names could be accounted for in this way



can be falsified by *reductio ad absurdum*. The assumption that all the opaque names are native formations would force one to reckon with hundreds of lost Saami words most of which were only used once in place name formation before their disappearance. The idea of the former existence of such words is rendered even more unrealistic by the non-native phonological structures exhibited by many of the names in question.

The remaining alternatives, including borrowing from known contact languages, have even less explanatory power. It is impossible to find any plausible points of comparison in the neighboring languages for at least the majority of the names. Moreover, in the light of known history it is impossible to assume that the Saami communities in inland Lapland have heavily borrowed toponyms from Scandinavian or Finnish. *Urschöpfung* can be categorically dismissed as an explanation, and in the case of place names the same applies to onomatopoeia and sound symbolism, too. Thus, there remains no real alternative to assuming that a large part of the names derive from the extinct languages of Lapland. This solution also provides an explanation to the discrepancy between typology and historical phonology that is manifest in the material: it is likely that the majority of the names are old because they denote major topographical formations, but due to their phonological structure they cannot be old in Saami.

The substrate hypothesis can now be considered proved, and we can proceed by examining if some of the substrate toponyms could actually be analyzed. In the case of unique names this is naturally impossible, but substrate influence also produces recurring name types which reflect the "toponymic basic vocabulary" of the source language. Even if the source language is unknown it ought to be possible to decrypt the original naming motive expressed by some recurring elements. This can be illustrated with a thought experiment. Let us assume that Saami had completely disappeared and we knew nothing of the language of the medieval 'Lapps' who inhabited most parts of Finland. A careful analysis of the Finnish nomenclature would nevertheless reveal recurring name types, for example the numerous lake and pond names of the shape *Kuukasjärvi*, *Kukkajärvi*, *Kuukka*, *Kukkamo*, etc. Because these names consistently denote lakes of oblong form, one could correctly identify a substrate language lexical item \**Ku(u)Kka(s)* 'long' or the like (cf. PS \**kukkä*, attrib. form \**kukkäs* 'long' > Saan *guhkkä*, *guhkes*). Exact phonological and semantic reconstruction would, of course, remain impossible.

As expected, some of the mountain names in the list above are not entirely unique, but have more or less exact parallels either within the list or outside it; examples of the latter include e.g. *Läpkkä* (Ohc) ~ *Läpkkavárri*, another mountain

(Náv); *Viräni* (Ean) ~ *Viränejávri*, a lake, and *Viränehohkka*, a mountain top (Guo); *Likkä* (Guo) ~ *Likkajohka*, *jávri* and *-čearru*, a river, a lake, and a mountain ridge (Kár); *Mollejus* (Guo/Rái) ~ ? *Mollesjávri*, a lake, and *Mollesjohka*, a river (Rái); *Vuorji* (Kár/Lea) ~ *Vuorján* : *Vuorjám-*, a river (MVÁ); *Skuohtki* (Dav) ~ Saal *Skuogetjikkä*, a very high mountain top in the Lule Saami area (the similarity to Scandinavian *skog* 'forest' is certainly coincidental, because the mountains in question are far above the treeline). In these cases the examination of maps reveals no shared feature between the places in question, but a couple of such recurring elements appear to yield to this kind of analysis. A thorough study in this direction would be very time-consuming, not to mention difficult due to the lack of appropriate place name collections from most Saami languages. Nevertheless, even a superficial examination of the existing collections of North Saami place names reveals a couple of structurally and semantically identifiable elements which are probably non-Saami in origin.<sup>5</sup>

1. *-ir* ~ *-Yr* 'mountain, uplands' (< PS \**-ērē* ~ \**-Yr(ē)*) || *Báhhtervári* (MVÁ) || *Buočšir* (Mái) || *Čáddervári* ~ *Sáddervári* (and *Čádder-* ~ *Sádderjávri*, an adjacent lake) (MVÁ/Aná) || *Čohhil* ~ *Cohkkir* (Fál/Lea); if the former form is primary, *-hkk-* must be due to folk-etymological contamination with *čohkka* 'steep or pointed mountain top' || *Čuošmmir* (Mái) || *Alip*, *Lulip Čáhtitár* (Čoh) || *Gáibbar* (Rom) || *Gealbir* (Mái) || *Guohkkar* ~ *Guohkkel* (Dav) || *Suorra Háhtir*, *Unna Háhtitaraš* (Fál/Lea) (perhaps < *háhtta* 'hat') || *Hoalgir* (Ála) || *Huhtitčohkat* (Vár) || *Iškoras* (Kár) || *Jeahkár* (2 instances, see entry 3 below) || *Miohktiris* (Váh) || *Nuhp(p)ir* (Ohc) (cf. *Nuhpi* (Guo/Náv/Ála)) || *Mussir* (as many as five instances, see entry 4 below) || *Ruohtir* (Ohc) || *Stieht(ē)aras* (Ohc) (see entry 2 below) || *Soanččir* : *Soanččir-* (~ *Soanččervári*) (Báhcca) || *Solc(c)ar* (Ohc) || *Suošmir* (Gár) || *Váhčir* ~ Saal *Váhčir* (Aná/MVÁ) (exact parallel in Saal: *Váhčjer*).

There are relatively many North Saami names of mountains and uplands which show the toponym formant *-ir* ~ *-Yr*, in a couple of cases followed by another formant *-ks*. In contrast, the same formant is rare in other types of names (but cf. *Aigir*, a heath (Lea) – the similarity to *áigi* 'time' is obviously coincidental). Because in almost every case the base of the name is etymologically obscure, it is likely that *-ir* reflects a substrate language

5. Regrettably, it is not possible to provide here maps illustrating the relevant features of the places discussed. However, all the names in entries 2–6 can be located with the help of the online map services "Norgesglasset" maintained by Statens Kartverk, Norway (<http://ngis2.statkart.no/norgesglasset/default.html>), and "Mapsite" maintained by The National Land Survey of Finland ([http://www.kartta.fi/index\\_e.html](http://www.kartta.fi/index_e.html)). Note that the special characters used in Saami orthography are not adequately handled by the place name search engines on these sites, and thus have to be replaced. The replacements necessary for finding the names discussed here are: *č*, *Č* > *ts*; *-čč-* > *ts*; *đ* > *α* (for *Mapsite*); *ē* > *è*; *Č* > *É* (sic) (for *Norgesglasset*).

morpheme or lexeme with the meaning 'mountain', 'uplands', or the like. These kinds of toponymic formants which reflect source language generics or derivative suffixes are well attested in various cases of substrate influence; cf. e.g. such frequent endings as *-nem* (< Finnic *\*nēmi* 'cape') and *-la* (< Finnic *oikonym* suffix *\*-la*) in the Finnic substrate toponyms in Northern Russia (Saarikivi 2000). The assumption of substrate origin of the formant *-ir* is further supported by the fact that it attaches to the recurring substrate bases *\*skiehč(č)-, jeahk(č)-* and *nuss-* which appear to be semantically analyzable (see below).

It is conceivable that the name *Háhhtir* may derive from the Scandinavian loanword *háhtita* 'hat'. The shape associative mountain name type *Hatten* ('the hat') is relatively common in Norwegian (Sandnes/Stemshaug 1997 s.v. *Hatten*, cf. s.v. *Háhhtir*), and even Saami names such as *Háhhtanvárrri* ('hat-mountain') are occasionally attested. On the other hand, a formation of the type *háhtir* < *háhtita* cannot be explained as a regular derivative in Saami. There are two possibilities: either the form *Háhhtir* has been formed from *háhtita* 'hat' on the analogy of other mountain names ending in *-ir*, or the correspondence is coincidental. The same anomalous morphological relationship occurs between the mountain name *Čohkkár* (~ *Čohhtit*) and the noun *čohkka* '(steep or pointed) mountain top'. Even if this correspondence is not accidental, it does not form a counterargument to a substrate etymology. The word *čohkka* has no satisfactory etymology (the suggested comparison to Komi *čuk* 'hill, hillock', and Nenes *soxa* 'id.' is not phonologically regular), and the word may thus derive from the substrate languages of Lapland; the irregular variation of the stem vowel (cf. *čohkka* < *\*čohkkē* vs. Saai *čohkke* < *\*čohkkē*) could also be seen as supporting this idea. If this is the case, a name of the shape *Čohkkár* may be a direct borrowing from a substrate language formation consisting of *\*čohkk-* 'steep, pointed top' + *\*ir* 'mountain'.

2. *skiehč(č)-* 'watershed' (< PS *\*skiehč(č)-*) || *Skiehččájávri* (Guo), a lake. || *Skiehččán* (*Skiehččam-*), a river originating from lake *Skiehččánjávri*; *Skiehččamvárrit*, fells. (Aná/Kár) || *Skiehč(č)aras* (Ohc), a mountain.<sup>6</sup>

The element *skiehč(č)-* appears in three name families in the water system of the river Deatnu (Norw. Tana). All the names are connected with watersheds. The river *Skiehččán* originates from lake *Skiehččánjávri* and flows north into the river Anárjohka; on its eastern side is a mountain ridge including the fells *Skiehččamvárrit*, which forms the watershed between the rivers *Skiehččán* and *Fáška*, which belong to different water systems (note that *Fáška* is also an opaque and phonotactically unusual name). Lake *Skiehččájávri* in the upper course of the river Áhkkanas is situated next to the watershed between the water systems of Deatnu and *Guovdageaineatnu/Álahaieatnu*. The fell *Skiehč(č)aras* is situated in the watershed between the rivers *Buolbmátjohka* and *Čárajohka/Sillisjohka*, which belong to different water systems and flow in opposite directions (note that also *Čára-* seems to be a substrate element; see entry 5 below). In this name *-r-* seems to be the substrate language formant meaning 'mountain' (see above).

<sup>6</sup> This substrate etymology was jointly discovered by myself, Samuli Aikio and Janne Saarikivi.

3. *jeahk(č)-* 'isolated (mountain)' (< PS *\*jeak(č)-*) || *Suorra Jeahkár*, *Umma Jeahkiras* (Lea) || *Suorra Jeahkár*, *Umma Jeahkiras* (Fál) || *Suorra*, *Umma Jeahkkás* (Ean)

The element *jeahk(č)-* appears in three mountain names, twice suffixed with the formant *-ir* 'mountain' and once with the deminutive suffix *-š* (*jeahkkás* < PS *\*jeak(č)äššē-ščē* or *\*jeak(č)šjē-ščē*). The element resembles the Scandinavian loanword *jeahkki* 'glacier' which also occurs in toponyms, but the three names seem to have nothing to do with glaciers. Instead, the feature shared by the mountains in question is that they are isolated, i.e. surrounded by low-lying terrain and river valleys.

4. *nuss-* 'mountain top on the edge of a highland area, projection of a mountain' (< PS *\*nuss-*) || *Nussir* (Ála) || *Nussir* ~ *Nussirvárrri* (Ála) || *Nussir* (Fál) || *Galmmat-Nussir* (Kár), an extension of the fell *Galmmat* (? cf. *galmmas* 'cold') || *Gurbbeš-Nussir* (Kár), an extension of the fell *Gurbbeš* (? cf. *guorbbas* 'barren (soil, land), hairless').

The element *nuss-* appears in as many as five names, always suffixed with the substrate formant *-ir* 'mountain'. All the names denote mountain tops which are situated on the edge of a highland area, next to a river valley or, in one case, near the seashore. They are often connected on one side to a higher mountain.

5. *čára-* 'uppermost (lake)' (< PS *\*čäreč(č)im*) || *Čárajávri*, two lakes (Ohc) || *Čárajávri*, a lake (Guo) || *Čárajávri*, a lake (Rái).

The element *čára-* appears in the names of three lakes, all of which are the uppermost ones in their water systems.

6. *sáll-* 'major island in the sea' (< PS *\*säl-*) || *Sállan* : *Sállam-* (Ákny/Hám) || *Sállan* : *Sállam-* (MVÁ) || *Sállir* (Norw. Kvalsøy, earlier *Salarøy* ? < Saami) || *Dávnesálla* ? < *távme-sälē*, cf. Norwegian *Tamsøy* (< Saami) (Lea).

The element *sáll(a)-* appears in four island names. The first three of these denote major islands in the sea, and even the fourth, *Dávnesálla*, belongs to the largest island in the fjord *Porsángu*. The name *Dávnesálla* looks formally like a compound, but both of its parts are opaque. Leem (1756, 579) cites the earlier form as "*Daume-sálo*" (? = *\*Dávmesálla*), and connects this with "*daume*" 'twig, spray'. Frette (1984) hesitatingly suggested the reconstruction *\*Dávmes-álla*, containing the attributive form *\*dávmes* of a hypothetical adjective *\*dávmei* 'full of twigs, sprays' and the generic *\*álla* 'island'. This etymology is unlikely for a variety of reasons (most of these problems were acknowledged by Frette, too). First, the word "*daume*" mentioned by Leem is a hapax legomenon; it is not even included in Leem's later Saami dictionary (1768-1781), even though the place name "*Daume-sálo*" itself is listed. Second, the word contains the consonant cluster *\*vm*, which is not attested in shared Saami vocabulary and which is phonotactically illegal even in present-day North Saami; thus, even if the original form of the name was *\*távme-sälē*, the non-native cluster suggests foreign origin. Third, the element *\*álla* has never been attested as a lexeme and it only occurs as a component in one other island name, the name of the islands *Čáramasálar* (pl.) (*čáramas* 'crow'). Furthermore, a homonymous element is attested in a couple of mountain names (*Állavárrri*, and *Álla* in the mountain name list), which suggests that if *álla* ever occurred as an appellative its original meaning was not

'island'. It seems much more likely that the name contains the element \**sálta* which is connected with the concept of 'large island'.

It is puzzling that the substrate element \**säl-* strikingly resembles the shared Finnic-Saami word \**salaw* > Finn. *salu* '(archaic) large island', Saami *suolu* 'island'. However, even if the two are connected, this does not suggest that the substrate languages in Lapland were Uralic. The word \**salaw* has no cognates in the more eastern branches of Uralic, and it is usually taken to be a Baltic loanword, cf. Lithuanian *salá* 'island' (< Proto-Baltic \**salā*). But also the Baltic word is isolated and has no Indo-European etymology. Thus, Saarikivi's (in press, a) suggestion that the word may originally derive from an extinct substrate language seems well possible. If the correspondence between Finnic-Saami(-Baltic) \**salaw* and the Laplandic substrate item \**säl-* is not due to coincidence, the word must have been shared between languages spoken in the coast of the Arctic Ocean and further south in the eastern Baltic Sea region – either as a loanword or due to genetic relationship.

It is naturally legitimate to ask how one does know that the six name elements discussed above derive from substrate languages instead of merely representing lost Saami words. The answer is that in the case of individual items one cannot know this for certain. Nevertheless, certain factors strongly speak against native origin. The most evident case is the formant *-ir* 'mountain', which almost exclusively combines with roots of unknown origin and meaning – the assumption that *-ir* was originally a native suffix would thus only lead to circular argumentation. From this it follows that the name elements \**jeak(k)-* and \**nus-* can also be deduced as substrate items because they occur combined with the substrate formant *-ir*. Also \**skteč(č)-* 'watershed' is on phonotactic grounds probably a substrate element. If one assumed that there once was a North Saami appellative \**skrehčča* 'watershed' which disappeared after the formation of the three names, this would only raise the question of the origin of this word – due to the initial cluster *sk-* it would have to be a late borrowing. The same naturally applies also to \**sálta* 'large island' and \**čára* 'uppermost': if there have ever been such lexemes in Saami, they are in any case of unclear origin.

If one accepts only two instances as a recurring name type, at least two more cases can be added to the ones discussed above. The names *Čiesti* (Unj) and *Čiesti* ~ *Šesti* (Dav) (< PS \**čiestē* ~ ? \**šiestē*) both denote small fells with a steep cliff on the seashore. An element \**iñč-* 'outermost island' seems to occur in two island names. *Ižtot* (< PS \**iñčujē-k*) (Norw. Lille Ekkerøy) (Čáh) is a small outer island on the border of open sea at the mouth of the fjord Várjavuotna (Norw. Varangerfjorden), near the northern shore. *Fávle-Ižžát* and

*Gádde-Ižžát* (< PS \**iñčā-k*) (Muo) is a group of two relatively large outer islands (cf. *fávli* 'open sea', *gáddi* 'shore').

Careful studies based on more extensive materials would probably reveal more recurring and semantically analyzable name elements. Nevertheless, this line of study would probably also quickly meet its limits, as only those recurring name elements can be analyzed which have expressed a clearly distinctive topographic feature in the substrate language. Needless to say, only a minority of all place names in any language are of this kind. In fact, critical etymological study of toponyms involves so many methodological limitations that even in the case of known substrate languages only a minority of borrowed place names can be reliably etymologized (for more detailed discussion see Aikio, in press). In the case of unknown languages the prospects of onomastic analysis are still much bleaker.

#### 4. The dating of the Saami expansion

The analysis conducted in the previous chapter is admittedly sketchy, and a more thorough examination of the existing materials would no doubt yield a clearer picture of the substrate elements in Saami. Nevertheless, the brief examination presented above suffices to verify the substrate hypothesis. The North Saami lexicon and nomenclature exactly fulfill the criteria devised to test the presence of unknown substrate.

The next step is to date the language shift. Due to the very narrow distribution of many substrate items, the numerous irregular correspondences between the Saami languages, and late phonological foreignisms such as the initial consonant cluster *sk-*, the last phase of substrate borrowing appears to have occurred at the time when Proto-Saami had already dialectally disintegrated. Because the formation of new dialect isoglosses is an inevitable result of a large-scale linguistic expansion, it is natural to assume that this disintegration took place contemporaneously as Proto-Saami spread over the present Saami area and pushed the native languages of Lapland to extinction on its way. Under such circumstances also heavy borrowing of place names must have taken place.

The Scandinavian loanwords in Saami provide a useful fixed point for the dating of these early dialect boundaries. Some Proto-Scandinavian loans retain the foreign initial cluster *sk-* in western and central Saami (e.g. SaaN *skáži* 'seashell' ~ SaaK *kálf* 'id.' < Proto-Scand. \**skaljō-*), whereas in older Indo-European borrowings foreign *sk-* is always reflected as plain *k-* in Saami. Because

the initial cluster *sk-* also occurs in probable substrate items, the last phase of substrate influence appears to have been contemporaneous with the adoption of Proto-Scandinavian loanwords. Wälund (1947) suggested that Proto-Saami disintegrated as late as 700-800 A.D. because there are relatively many Scandinavian loanwords showing a wide or uniform distribution in Saami. However, the argument is based on an unrealistically late dating of the break-up of Proto-Scandinavian. Moreover, Sammallahti (1984, 145) points out that some of these loanwords show phonological irregularities in Kola Saami which suggests that they have spread between already diverged dialects. On the other hand, even phonologically regular correspondences do not exclude the possibility of dialectal diffusion because borrowings between dialects and closely related languages easily become conformed to patterns of regular sound correspondence (see Trask 2000, s.v. *loan nativization*).

It can be inferred that Proto-Saami disintegrated slightly earlier than Proto-Scandinavian, and a part of the Proto-Scandinavian loanwords became adjusted to regular sound correspondences when they were mediated to Kola Saami. Moreover, due to such borrowed place names as e.g. *Mueffie* (the South Saami name of *Mo, Vefsén*) < Proto-Scand. \**mōha-*, which preserves a reflex of intervocalic \**-h-*, Proto-Saami appears to have reached central Scandinavia already in the Proto-Scandinavian phase (Bergsland 1996). As Proto-Scandinavian can be equated with Early Runic in the period 200-500 A.D. (H. F. Nielsen 2000), we can take 500 A.D. as the latest possible dating for the break-up of Proto-Saami. It is more difficult to set the *terminus post quem*, but considering typical rates of linguistic change it is hardly possible to date the emergence of early Saami dialect boundaries much before the Common Era. Hence, the language shift from the unknown "Palaeo-Laplandic" languages to Saami was probably completed between approximately 0-500 A.D. The dating might appear surprisingly late, but the linguistic data leaves little possibility for dramatically stretching the timetable backwards. Even if one allows considerable latitude in the dating, any date preceding the early Iron Age is out of the question. Needless to say, this result implies the rejection of all suggestions of Saami ethnic continuity in Lapland since the Stone Age or even the Bronze Age.

In this connection it is useful to examine the results achieved in the research on Indo-European loanwords in some more detail. The studies conducted by Jorma Koivulehto during the recent decades have greatly elaborated and changed the picture of Indo-European loans in the Uralic languages, including Saami. In earlier research it was commonly maintained that Saami had adopted few (if any) independent Indo-European loanwords before the Proto-Scandinavian

phase, and the older strata of Baltic and Germanic borrowings that are present in Saami were explained as mediated by Finnic. However, Koivulehto has demonstrated that there are several consecutive strata of independent borrowings in Saami which precede both the Proto-Scandinavian and Proto-Saami levels of reconstruction (the literature is too extensive to be thoroughly discussed here; Koivulehto 2002 provides an up-to-date summary with further references).

It is illuminating to compare the substrate vocabulary in Saami against the stratification of Germanic (and Baltic) loanwords. Koivulehto (op. cit.) divides the Germanic loans in Saami into three main strata. For the purposes of this presentation these consecutive layers of borrowings can be illustrated with the three distinct Saami reflexes of Germanic \**ha-*; the examples have been taken from Koivulehto (op. cit.).

1. Borrowings which show identical patterns of sound substitution in Saami and Finnic and which have participated in all the Saami and Finnic sound changes; these loanwords are often shared between the two language branches. In this stratum Uralic \**k-* has been substituted for Germanic \**h-*, cf. SaaN *gnoos'ri* 'guest' (~ Finn. *kansa* 'people') < Finnic-Saami \**kansa* < Germanic \**hansaz-* 'crowd, host, etc.'
2. Pre-Saami borrowings which were adopted after phase 1 but prior to the Saami vowel rotations, including the shift \**a* > \**ö* > \**uo*. These loans show Ø as the substitute for Germanic \**h-*. Cf. SaaN *vuoopman* ~ *vuoma* 'a kind of hunting fence' < Pre-Saami \**amin* < Germanic \**hamen-* 'hunting net, etc.'
3. Borrowings adopted into Proto-Saami after the vowel rotation, e.g. SaaN *áyyi* 'hay' < Proto-Scand. \**hazja-* 'id.'

Soon after the beginning of phase 3 Proto-Saami disintegrated dialectally and new phonotactic features such as initial *h-* and medial *-f-* became established in western and central Saami through borrowing, cf. e.g. SaaN *háittis* 'very hot (of stove etc.)' < Proto-Scand. \**haitiaz* 'hot' and SaaN *márfi* 'sausage' < Proto-Scand. \**marhu-* 'fat in the intestines'. In many individual words it is difficult to distinguish between subsequent phases, as an indisputable phonological criterion cannot always be found, and some Germanic loans have probably been mediated to Pre-Saami by Pre-Finnic (and even vice versa). As for the Baltic loans in Saami, the majority of them apparently belong to the phase 1. A part of them were probably adopted during phase 2, but it is difficult to show this due to the lack of appropriate phonological criteria. By phase 3 the adoption of Baltic loans had probably ceased.

The fact that the last period of substrate influence coincides with the adoption of Proto-Scandinavian loanwords (phase 3) reveals that Saami has been in con-

tact with unknown languages long after the beginning of intensive interaction with Germanic- and Baltic-speaking populations. The Indo-European loanwords and substrate elements together yield a consistent picture of the earlier history of Saami. It is clear that the phases 1 and 2 in the development of Saami must have taken place at a relatively southern latitude because otherwise e.g. the adoption of Baltic loanwords would not have been possible. The location of this Saami original home cannot be exactly pinpointed in the light of present research, but it must have been situated somewhere in the vicinity of the Gulf of Finland, Lake Ladoga and Lake Onega; this rough placement provides the possibility of contact both with Germanic via the Baltic Sea and with Baltic to the southeast of the Gulf of Finland. Interestingly, it seems that even some Scandinavian loanwords belonging to the phase 3 have been adopted along these latitudes. This is shown by the fact that some of them have been further transferred as substrate items to Finnish and Karelian from the extinct Saami languages in these areas. A couple of them are attested as dialectal loanwords (e.g. Finn. *ume* 'mist' (southeastern Finland) < PS *\*omV-* 'id.' < Proto-Scand. *\*hūm-* 'dusky, half-dark'), whereas still others are reflected in Saami substrate place names (e.g. *Raasa-* < PS *\*rāšē* 'grass' < Proto-Scand. *\*grasa-* 'id.' and *Mella-* < PS *\*miellē* 'steep sandbank' < Proto-Scand. *\*melha-* 'id.') (see Aikio, in press, for more details).

In addition to Indo-European loanwords there is also another, more general reason for assuming that Proto-Saami spread from a restricted homeland to its present territory in the Iron Age. As noted above, linguistic divergence is the natural result of language spread. It is known that now extinct Saami languages were once widely spoken in southern Finland and Karelia, and the Saami substrate toponyms in this area demonstrate that these languages had undergone the same set of Proto-Saami sound changes as their surviving sister languages in the north (see e.g. Aikio, in press; Saarikivi, in press, b). An assumption that Saami reached its maximal distribution *before* the Proto-Saami sound changes would imply that Lapland, Finland and Karelia were covered by one huge speech community, as otherwise the sound changes could not have been uniformly completed. Due to the geographical extent and the topographical and ecological diversity of the area such a situation can hardly have been possible. Instead, prior to the expansion of Saami the linguistic map of Fennoscandia has probably been a mosaic of languages much like the regions later inhabited by hunter-gatherers, such as North America (see e.g. Mithun 2001).

Thus, it can be inferred that Saami originated in a relatively southern homeland in the vicinity of the Gulf of Finland. By the early Iron Age the phonological innovations leading to Proto-Saami had been accomplished in this homeland,

and only after this – for reasons unknown at present – the language expanded northward, soon covering most of Finland, Karelia and Lapland. During this process Saami must have swallowed up a number of distinct and possibly even unrelated substrate languages. Evidently, not all the substrate elements in Saami have been adopted at the same time or in the same area. Indeed, some of the probable substrate words discussed in this paper show varying areas of distribution in the Saami substrate toponyms of southern Finland and Karelia. For instance, the words *leakšá* 'wide marshland' and *luspi* ~ *lusmi* 'river head' are found in Karelia in the forms *Lieksa(-)* and *Lusma(-)*, whereas *vuotna* 'fjord, long and narrow bay' is attested in both Karelia and inland Finland in the shape *Vuon(n)-*. Some items show a more northern distribution, e.g. *roarvi* 'place where there has been a forest fire' (> *Rova-*) and *bákti* 'cliff, rock' (> *Pahta-*) which are almost exclusively attested in the Finnish Lapland. The southern boundaries of these distributions presumably roughly reflect the latitudes where Saami adopted these substrate lexemes on its way north. A more detailed discussion of these kinds of lexical horizons in substrate nomenclature and their implications to the origin of Saami is provided by Saarikivi (in press, b).

## 5. Conclusion

In conclusion, it can be said that K. B. Wiklund was principally on the right track with his "Protolappisch" theory, even though he did not manage to provide linguistic substantiation to his claims. Nevertheless, one qualification is in order. There is no reason to think that the Saami have "abandoned" their native language and "borrowed" a new one from their Finno-Ugric neighbors. This idea merely resulted from the mistake of viewing human populations as unchanging monoliths which could be "racially" classified e.g. as Finno-Ugric and non-Finno-Ugric. Considering basic sociolinguistic regularities, the idea that the native inhabitants of Lapland would have switched their languages to Proto-Saami without any population mixture is clearly unrealistic.

This assumption of population mixture is in no contradiction with the observed genetic distance between the present-day Finns and Saami. The spread of Saami from the south to the north must have occurred step by step, and during each of these steps the intruding population must have become mixed with the shifting group. A sufficient number of such steps would result in the genetic component of the first expanding group becoming diluted almost to null. It is true that such a process would have left a genetic continuum behind, but even

this would have been later wiped out by the Finnic expansion. The medieval hunter-gatherer Saami of southern Finland and Karelia were by far outnumbered by the spreading Finnic farmers, and thus even a process of total assimilation must have contributed relatively little to the Finnic genetic stock.

The model outlined above is naturally quite sketchy, and the reconstruction of the details of the Saami ethnogenesis remains a challenge for future research. It remains to be seen, for instance, whether the results of this paper can be correlated with the archaeological record. In particular, Christian Carpelan's (2003, 87) recent suggestion that the expansion of Saami would be reflected in the quick spread of Kjelmoey Ware in Lapland in the beginning of the Iron Age (650 B.C. >) appears promising, even though the dating is somewhat early compared to the linguistic data. But in any case, a more detailed analysis of both linguistic and archaeological data is necessary before exact conclusions can be drawn.

Finally, an interesting implication of the above model can be mentioned. Because the last phase of substrate influence from the "Palaeo-Laplandic" languages seems to have coincided with the adoption of Proto-Scandinavian borrowings into Saami, it seems evident that these lost languages have come to contact with Scandinavian as well. This being the case, it is likely that they have contributed loanwords to Scandinavian, too – even though the interaction was probably not as intensive as that with Saami, considering the extralinguistic aspects of the situation. However, whether the linguistic traces of this contact can be uncovered must be left to Scandinavianists to decide.

## 6. Appendix: A list of identifiers in the North Saami names discussed in this paper

*bođus* 'loose; isolated (of mountain)' || *davit* (~ *davip*) 'more northern' || *davvi* 'north' || *duottar* 'uplands, tundra' || *šivle*-compound form of *šivli* 'deep water, open water, open sea' || *gaska* 'middle' || *gáddá*-compound form of *gáddi* 'shore' || *goldáa* 'drift-net' (in *Goldáa-Málla*, taken over from the name of the adjacent lake *Goldáajávri*) || *gorba* 'round (attrib. form)' || *lullit* (~ *lullip*) 'more southern' || *lullit* 'south' || *Márkos*- (in *Márkos-Málla*) based on the fell name *Márkos*, opaque || *ráš'sa* 'barren mountain' || *Ságge*-formally identical with *sággi* 'stick, pin' (compound form *ságge-*), but the connection is not semantically clear. The identifier also appears in a couple of autochthonous names, e.g. *Ságge-Muorjearri* (Muorjearri 'berry mountain'). || *sillis* 'thin (of fish)' (in *Sillis-Málla*, taken over from the name of the adjacent lake *Sillisjávri*) || *stuoara* ~ *stuoara* ~ *stuoar*- 'big, large (attrib. form)' || *ul'lo-* (in *Ul'lo-Roaxjja*) compound form of *ul'la* 'wool' (sic) || *unna* 'small (attrib. form)'; in

connection with this adjective the head noun usually occurs in the diminutive (with the suffix -š).

## 7. Abbreviations

### 7.1. Municipalities

Aná = Anár (Inari), Finland || Ákrj = Ákrováhki (Hasvik), Norway || Ála = Álaheadju (Alta), Norway || Báhca = Báhcavuotna (Bátsfjord), Norway || Báhcca = Báhccavuotna (Bátsfjord), Norway || Bdu = Beardu (Bardu), Norway || BeV = Bearalváhki (Berlevág), Norway || Čáh = Čáhceuoalu (Vadsø), Norway || Čoh = Čohkkiras (Jukkasjärvi), Sweden || Dav = Davvesiida (Lebesby), Norway || Dea = Deatnu (Tana), Norway || Ean = Eanodat (Enontekiö), Finland || Fál = Fálesnuorri (Kvalsund), Norway || Gáj = Gággaviilka (Garnvik), Norway || Gái = Gáivuotna (Kåfjord), Norway || Gár = Gárasavvon (Karesuando), Sweden || Guo = Guovdageaidnu (Kautokeino), Norway || Hám = Hámmerfeasta (Hammerfest), Norway || Ivg = Ivgu (Lyngen), Norway || Karl = Karlisey, Norway || Kár = Kárásjohka (Karasjok), Norway || Láh = Láhppi (Loppa), Norway || Lea = Leavdnja (Lakselv), Norway || Muo = Muosát (Måsøy), Norway || MVá = Máitta-Várjjat (Sør-Varanger), Norway || Mál = Málšelv, Norway || Nar = Narvik, Norway || Náv = Návuoatna (Kvænangen), Norway || Nor = Nordkapp, Norway || Ohc = Ohcejohka (Utsjoki), Finland || Oma = Omasvuotna (Storfjord), Norway || Rái = Ráina (Nordreisa), Norway || Rom = Romssa (Tromsø), Norway || Ski = Skiervá (Skjerveøy), Norway || Tra = Trausøy, Norway || Unj = Unjárga (Nesseby), Norway || Vaz = Vazáš (Vittangi), Sweden || Váh = Váhšir (Saal. Váhšier) (Gällivare), Sweden || Vár = Várggát (Vardø), Norway.

### 7.2. Languages

PS = Proto-Saami || Saal = Inari Saami || SaaK = Kildin Saami || Saal = Lule Saami || Saan = North Saami || Saap = Pite Saami || SaaS = South Saami || SaaSk = Skolt Saami || SaalT = Ter Saami.

## 8. References

### 8.1. Archives

The Archive of Names. Research Center for the Languages of Finland, Helsinki. (Materials used: The collections of North Saami place names in the municipalities of Inari, Utsjoki and Enontekiö, a copy of the collection of North Saami place names in Norway by Thor Frette.)

## 8.2. Maps

- Áltá (1991). (A single map sheet.) Designed by Hans Ragnar Mathisen. Keviselie, Bodø. Deavuotna ja Vuolle-Deatnu. (A single map sheet.) Designed by Hans-Ragnar Mathisen (uncredited). Time and place of publication unknown.
- Fjällkartan. Topografisk karta skala 1 : 100 000. Lantmäteriet, Sweden.
- Návuoana. (A single map sheet.) Designed by Hans-Ragnar Mathisen (uncredited). Time and place of publication unknown.
- Njárggat, vuonat ja sullot. (A single map sheet.) Designed by Hans-Ragnar Mathisen (uncredited). Time and place of publication unknown.
- Sállan. (A single map sheet.) Designed by Hans-Ragnar Mathisen (uncredited). Time and place of publication unknown.

## 8.3. Literature

- Aikio, Ante (2002): Sápmelaš loanasánit suomelaš gielaide davvejoavkkus. Unpublished Master's thesis, University of Oulu; available on the internet at <http://cc.oulu.fi/~anaikio/> hp.html.
- Aikio, Ante (in press): The Study of Saami Substrate Toponyms in Finland. In: Ritva Liisa Pitkänen/Janne Saarikivi (eds.): The Borrowing of Place Names in the Finno-Ugric Languages. Debrecen/Helsinki (= Onomastica Uralica 4).
- Bergsland, Knut (1996): Bidrag til sydsamenes historie. Tromsø (= Sámi dulkamiid guovddáš/Senter for samiske studier, Skriftserie nr. 1).
- Bergsland, Knut/Mattsson Magga, Lajla (1993): Áarjelsaemien-daaeroen baakoerjia – Syd-samisk-norsk ordbok. [Lakseliv].
- Carpelan, Christian (2003): Inarilaisten arkeologiset vaiheet. In: Veli-Pekka Lehtola (ed.): Inari – Aanaar. Inarin historia jääkaudesta nykypäivään. Oulu, pp. 28-95.
- Carpelan, Christian/Parpola, Asko (2001): Emergence, contacts and dispersal of Proto-Indo-European, Proto-Uralic and Proto-Aryan in archaeological perspective. In: Christian Carpelan/Asko Parpola/Petteri Koskikallio (eds.): Early Contacts between Uralic and Indo-European: Linguistic and Archaeological Considerations. Papers presented at an international symposium held at the Tvärminne Research Station of the University of Helsinki, 8-10 January 1999. Helsinki (= Mémoires de la Société Finno-Ougrienne 242), pp. 55-150.
- Collinder, Björn (1964): Ordbok till Sveriges lappska ortnamn. Uppsala.
- Frette, Thor (1984): On some "apparently Norwegian" place names in Finnmark, Norway. In: Bernt Brendemoen/Eved Hovdhaugen/Ole Henrik Magga (eds.): Riepmočála. Essays in honour of Knut Bergsland presented on the occasion of his seventieth birthday. Oslo, pp. 66-79.
- Frette, Thor (unpublished). Norsk-samiske stedsnavnpar. Del 1. (Printout from a computer database, dated 21.9.1987. 584 pp.)
- Friis, Jens Andreas (1887): Lexicon Lapponicum cum Interpretatione Latina et Norvegica – Ordbok over det lappiske sprog med latinsk og norsk forklaring. Christiania.

- Grundström, Harald (1946-1954): Lulelappak ordbok. Uppsala (= Skrifter utgivna genom dialekt- och folkminnesarkivet i Uppsala. Ser. C:1).
- Helimäki, Eugen (1996): Protolappisch und Samojedisch: die vorgeschlagenen Wortgleichungen im Lichte der heutigen etymologischen Forschung. In: Lars-Gunnar Larsson (ed.): Laponica et Uralica. 100 Jahre finnisch-ugrischer Unterricht an der Universität Uppsala. Uppsala (= Studia Uralica Uppsaliensia 26), pp. 51-69.
- Itkonen, Erkki (1961): Suomalais-ugrilaisen kielen- ja historiantutkimuksen alalta. Helsinki.
- Itkonen, Erkki (1986-1991). Inarilappisches Wörterbuch. Helsinki (= Lexica Societatis Finno-Ugricae XX).
- Itkonen, T. I. (1948): Suomen lappalaiset vuoteen 1945. I-II. Porvoo/Helsinki.
- Itkonen, T. I. (1958): Koltan- ja kuolanlappin sanakirja. Helsinki (= Lexica Societatis Finno-Ugricae XV).
- Koivulehto, Jorma (2002): Contact with non-Germanic languages II: Relations to the East. In: Oskar Bandle/Kurt Braunmüller/Ernst Håkon Jahr/Allan Karer/Hans-Peter Naumann/Ulf Teleman (eds.): The Nordic Languages. An International Handbook of the History of the North Germanic Languages. Volume 1. Berlin/New York, pp. 583-594.
- Korhonen, Mikko (1981): Johdatus lappin kielen historiaan. Helsinki.
- Lagercrantz, Eliel (1939): Lappischer Wortschatz. Helsinki (= Lexica Societatis Finno-Ugricae VI).
- Larsson, Lars-Gunnar (2001): Baltic influence on Finnic languages. In: Östen Dahl/Maria Koptjevskaja-Tamm (eds.): The Circum-Baltic Languages, Vol. 1: Past and Present. Amsterdam/Philadelphia (= Studies in Language Companion Series 54), pp. 237-253.
- Leem, Knud (1756): En Lappesk Normen Clator efter Den Dialect, som bruges af Fjeld-Lapperne i Porsanger-Fjorden. Tronhiem.
- Leem, Knud (1768-1781): Lexicon lapponicum bipartitum. Niddrosia.
- Lehtiranta, Juhani (1989): Yhteissamalainen sanasto. Helsinki (= Mémoires de la Société Finno-Ougrienne 200).
- Lindahl, Erik/Öhrling, Johan (1780): Lexicon Lapponicum. Holmiæ.
- Mithun, Marianne (2001): The Languages of Native North America. Cambridge.
- Nielsen, Hans Frede (2000): The Early Runic Language of Scandinavia: Studies in Germanic Dialect Geography. Heidelberg.
- Nielsen, Konrad (1913): Die wissenschaftliche Bedeutung des Lappischen. In: Finnisch-Ugrische Forschungen 13, pp. 186-206.
- Nielsen, Konrad (1979): Lappiaki (samiskt) ordbok. 2. opplag. Oslo.
- Qvigstad, J. (1904): Lappiske Navne paa Pattedyr, Krybdyr og Padder, Fiske, Leddyr og lavere Dyr. In: Nyt Magazin for Naturvidenskab, B. 42, H. IV, pp. 339-387.
- Qvigstad, J. (1935): De lappiske stedsnavn i Troms fylke. Oslo (= Instituttet for sammenlignende kulturforskning, Serie B. Skrifter, XXVIII).
- Qvigstad, J. (1938): De lappiske stedsnavn i Finnmark og Nordland fylker. Oslo (= Instituttet for sammenlignende kulturforskning, Serie B. Skrifter, XXXIII).
- Qvigstad, J. (1944): De lappiske appellative stedsnavn. Oslo (= Instituttet for sammenlignende kulturforskning, Serie B. Skrifter, XLII).

- Ravila, Paavo (1935): Die stellung des lappischen innerhalb der finnisch-ugrischen sprachfamilie. In: *Finnisch-Ugrische Forschungen* 23, pp. 20-65.
- Ravila, Paavo (1936): Die Lappen und Fennoskandien. In: *Germanen und Indogermanen. Festschrift für Herman Hirt. II. Heidelberg*, pp. 97-107.
- Ravila, Paavo (1957): Lappalaisten ja lapin kielen ongelmia. In: *Suomalainen Suomi*, pp. 465-469.
- Saarikivi, Janne (2000): Kontaktikäroinen kielenmuutos, substraatti ja substraattiniimitistö. In: *Virtittäjä* 104, pp. 393-415.
- Saarikivi, Janne (in press, a): Paleo-European substratum interference in western branches of Uralic? In: *Journal de la Société Finno-Ougrienne* 90.
- Saarikivi, Janne (in press, b): Über die samischen Substratennamen in Nordrussland und Finnland. In: *Finnisch-Ugrische Forschungen* 58.
- Salmons, Joe (1992): Northwest Indo-European Vocabulary and Substrate Phonology. In: Roger Pearson (ed.): *Perspectives on Indo-European Language, Culture and Religion: Studies in Honor of Edgar C. Polomé*. Vol. 2. Washington, D.C. (= *Journal of Indo-European Studies*, Monograph Series 9), pp. 265-279.
- Sammallahti, Pekka (1984): Saamelaiten esihistoriallinen tausta kielitieteen valossa. In: *Jari Gallén* (ed.): *Suomen väestön esihistorialliset juuret. Tvärminnen symposiumi 17.-19.1.1980*. Helsinki (= *Bidrag till Kännedom av Finlands Natur och Folk* 131), pp. 137-156.
- Sammallahti, Pekka (1989): *Sámi-suoma sátnegirji - Saamelais-suomalainen sanakirja*. Ohcejohka.
- Sammallahti, Pekka (1995): Language and Roots. In: *Congressus Octavus Internationalis Flenno-Ugristarum, Jyväskylä 10.-15.8.1995. Pars I. Jyväskylä*, pp. 143-153.
- Sammallahti, Pekka (1998): The Saami languages: an introduction. *Käräjöjohka*.
- Sandnes, Jern/Stemshaug, Ola (eds.) (1997): *Norsk stadnamleksikon*. 4. utgåva. Oslo.
- Trask, R. L. (2000): *The dictionary of historical and comparative linguistics*. Edinburgh.
- Wiklund, K. B. (1896): *Entwurf einer urfäppischen lautlehre*. Helsingfors.
- Wiklund, K. B. (1910-1911): *De lapska och finska ortnam vid Kiruna och Torneråsk*. In: *Le Monde Oriental* IV, pp. 185-228, V, pp. 1-36.
- Wiklund, K. B. (1947): *Lapparna*. Stockholm (= *Nordisk Kultur* X).

Raimo Anttila

## Sap from the birch grove – *sulaa* mahlaa koivulehdosta

The current handbooks and other “obvious” sources take Finnish *sula* ‘melted, unfrozen’ and its related forms as a Uralic inheritance, as is indeed likely – without any further speculation about its existence or source there. I will propose here that an Indo-European loan source cannot be ruled out.

### A. General [Baltic] Finnic

The [Baltic] Finnic semantic gamut is extremely rich, to pick out some highlights: Finnish ‘molten, not frozen, sheer, pure, genuine, flexible, able, good’; Estonian ‘molten, not frozen, soft (bog), mild (winter), pure, liquid’. Karelian shows about the same, but adds ‘friendly, lovely, happy, dear’. This adjective is curious in that it is indeclinable in Estonian and also in Finnish it leans slightly that way. As it has also spawned a *-va* adjective (like *iha* ‘flesh/meat’ ~ *lihava* ‘fat’), *sulava* (*suleva*) with about the same meanings as *sula* itself, we have an original noun here, although that might not be so important in Uralic where this boundary is rather fluid. Note also the Estonian adjective *sulakas* ‘rather melted or soft’. *Sula* is indeed all around in Baltic Finnic a noun for ‘thawing weather’, and its nominal character is further enhanced by its frequency in compounds. Many derivatives designating wet or snowless spots exist, matched by a richness of verbs meaning melting and (in more modern times) smelting (note, e.g., a “technicality” like Finnish *sulake* ‘fuse’).<sup>1</sup>

### B. Karelian – add food and drink

1. Particularly in Karelian and other related idioms we have a derivative *sulo(i)* ‘pleasant, pleasing, etc.’ (cf. *korva* ‘ear’, \**korvoi* > *korvo* ‘a wooden tub with

1 I will leave out here all speculation about Finnish *sulhainen* ‘bridegroom’ and Estonian *sulane* ‘farm hand’ (cf. *best man*, [condescending] *my good man* etc.). Karelian is rich in epithets of *sula*, etc., for the bridegroom, which could of course be natural folk etymology. Note further Karelian *sulaja* ‘yes-answer to marriage proposal’.