Contact in the prehistory of the Sakha (Yakuts): Linguistic and genetic perspectives
Pakendorf, B.

Contact in the prehistory of the Sakha (Yakuts):
Linguistic and genetic perspectives

Contact in the prehistory of the Sakha (Yakuts):
Linguistic and genetic perspectives

PROEFSCHRIFT

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van de Rector Magnificus Prof. Mr. P.F. van der Heijden,
volgens besluit van het College voor Promoties
te verdedigen op woensdag 12 december 2007
klokke 13.45 uur

door

Brigitte Pakendorf

geboren te Johannesburg, South Africa
in 1970

TABLE OF CONTENTS

ii

LIST OF ABBREVIATIONS USED IN THE GLOSSES:

1	1st person	DER	derivational suffix
2	2nd person	DES	designative/destinative
3	3rd person	DETR	detransitivizer
ABL	ablative	DIM	diminutive
ABS	absolutive	DISJ	disjunction
ACC	accusative	DIST	distributive
ADJR	adjectivizer	DP	discourse particle
ADVR	adverbializer	DSTIMP	distant future imperative
AFF	affirmative	DSTPROH	distant future prohibitive
AGNR	agent nominalizer	DU	dual
ALL	allative	DUR	durative
ALN	alienable possession	E	epenthetic vowel
ANR	action nominalizer	EMPH	emphatic
ANT	anterior (converb)	EXCL	exclusive
AOR	aorist	FIN	finite
ASS	assertive	FOC	focus
ATTR	attributive	FUT	future
AUX	auxiliary	FUTPT	future participle
BEN	benefactive	GEN	genitive
BND	benedictive	GL	goal
CA	connective adverbial	HAB	habitual
CAUS	causative	IMM	immediate-precedence (converb)
CLIT	clitic		
COLL	collective	IMP	imperative
COM	comitative	IMPRS	impersonal
COMP	comparative	INCL	inclusive
CON	connective case	INCP	inceptive
COND	conditional	IND	indicative mood
CONNEG	connegative converb	INDF	indefinite
COP	copula	INF	infinitive
CP	connective particle	INFR	inferential
CVB	converb	INS	instrumental
DAT	dative	INTR	intransitive
DEF	definite	INTS	intensive
		IPF	imperfective (converb)

IPFV	imperfective	PRDPOSS	predicate possessive
IRR	irrealis	PREFL	reflexive possessive
ITER	iterative	PRESCR	prescriptive
ITR	interrogative	PRF	perfect
LA	case functioning in privative	PROG	progressive
	constructions in some Tungusic	PROH	prohibitive
	languages	PROL	prolative
LAT	lative case	PROP	proprietive
LIM	limitive case	PRS	present
LOC	locative	PRSPT	present participle
M	masculine	PRV	privative (converb)
MDL	modal suffix	PRXIMP	immediate future imperative
MDS	modalis case	PRXPROH	immediate future prohibitive
MIN	minimal inclusive	PST	past
MOD	modal (converb)	PSTPT	past participle
MON	monitory imperative	PTCP	participle
MULT	multiplicative	PTL	particle
NARR	narrative	PURP	purposive
NEG	negative	REC	reciprocal
NR	nominalizer	REL	relative case
NOM	nominative	REFL	reflexive
NFUT	non-future	RES	resultative
NPST	non-past	Q	interrogative marker
OBJ	objective pers.marking	SBJ	subjective pers. marking
OBL	oblique	SEQ	sequential
OF	object focus	SG	singular
ORD	ordinal numeral	SIM	simultaneous (converb)
PART	partitive	SIML	similative
PASS	passive	SOC	sociative
PF	perfective (converb)	SS	same subject
PFV	perfective	SUB	subjunctive
PL	plural	SUCC	successive (converb)
POLIMP	polite imperative	TEMP	temporal (converb)
POS	posterior (converb)	TERM	terminative
POSS	possessive	TR	transitive
PRD	predicative marker	TRM	terminative case
PREC	precative	TRNS	translative
PRED	predicative person marking	VOC	vocative

VOL	volitional
VPOT	voluntative-potential
VR	verbalizer

1 INTRODUCTION[1]

Sakha (also known as Yakut) is a very divergent Turkic language that has copied a large number of words from Mongolic and is surrounded by Tungusic languages (Evenki and Èven[2]). A number of ethnographers mention the inter-marriage of the Sakha people with indigenous north Siberian groups as well as the linguistic assimilation of the latter in the course of Sakha prehistory (e.g. Seroševskij [1896] 1993: 230f; Dolgix 1960: 461, 486; Tugolukov 1985: 220). Not surprisingly, therefore, a large number of differences that distinguish Sakha from its Turkic relatives are attributed to contact with Evenki and/or Mongolic (Ubrjatova 1960: 78, 1985: 46; Širobokova 1980: 140; Schönig 1990: 95f; Johanson 2001: 1732). This study is an attempt at elucidating the contact influence the Sakha may have undergone in their prehistory, both from a molecular-genetic perspective (i.e. intermarriage/admixture) and from a linguistic point of view.

This introductory chapter presents an overview of the Sakha language and prehistory, as well as an overview of the languages and prehistory of the populations they are or were in contact with, i.e. Evenks, Èvens, Yukaghirs, and Mongolic-speaking groups (section 1.1). A discussion of the current theories and approaches to language contact follows in section 1.2, while previous studies of the impact of language contact on Sakha are presented briefly in section 1.3. In section 1.4 I outline the aims of this study and the general methodology followed.

1.1 The Sakha and their Siberian neighbours

1.1.1 The Sakha

The Sakha are one of the northernmost Turkic-speaking peoples in Eurasia. Although in the English-speaking literature they are frequently referred to as Yakuts (e.g. Gordon 2005: 507; Balzer 1994), their own ethnonym is Sakha, and they call their language *saχa tïl–a* [Sakha tongue–POSS.3SG] 'language of the Sakha'. Following the wishes of my consultants in Yakutia, I use the native ethnonym in this thesis[3]. According to the 2002 census, there are currently 443,852 Sakha in the

---

[1] In addition to the countless people mentioned in the acknowledgements, I sincerely thank Frederik Kortlandt and Bernard Comrie for crucial support and very constructive comments.
[2] Given the possibility of confusing the ethnonym Even at the beginning of a sentence with the English word 'even' [i:ven] I use the symbol for transliteration of the Russian letter Э (Ė) in the name of the people as well as their language. Since the name Evenk (Evenki for the language) is unambigous, I write it in its English form.
[3] For practical reasons, the term Yakut was retained as ethnonym in the publications of the genetic data (Pakendorf et al. 2006, Pakendorf et al. 2007).

Russian Federation, the vast majority of which reside within the autonomous Republic Sakha (Yakutia) (cf. Figure 1.1). Language retention among the Sakha is high – according to the 2002 population census, approximately 93% of Sakha know their heritage language, and only approximately 87% know Russian; among the rural population this figure is even lower, with only approximately 83% of the Sakha claiming a knowledge of Russian (Federal'naja služba gosudarstvennoj statistiki 2004: 19, 24, 113, 130)[4]. Amongst urbanized Sakha knowledge of Russian is more widespread, since in towns Russians and Ukrainians dominate numerically, whereas villages are predominantly mono-ethnically Sakha [with the exception of some villages in the north and northeast, where settlements are multiethnic, consisting of Sakha and minority peoples (Maslova 2003a: 2; personal observation)]. In Sakha rural settlements, older people are sometimes still monolingual Sakha speakers, as are children under school age, notwithstanding the fact that often the only television channels that can be received in such settlements are Russian (personal observation). As can be seen from the data of the 2002 census (456,288 speakers of Sakha as opposed to 443,852 people who claimed Sakha ethnicity; Federal'naja služba gosudarstvennoj statistiki 2004: 124), Sakha is endangering minority languages in Yakutia, especially Evenki and Ėven (Pis'mennye jazyki Rossii 2000: 576, 2003: 641, 668; Federal'naja služba gosudarstvennoj statistiki 2004: 151). Thus, in the Ėveno-Bytantaj district Sakha has nearly completely replaced Ėven, with only a few older Ėven speakers remaining (Raisa Starostina, pers. comm.; own observation).

The Republic Sakha (Yakutia) covers an enormous territory of more than 3,000,000 km^2 – roughly six times the area of France, and about one sixth of the area of the Russian Federation (Safronov 2000:11; Microsoft Encarta Reference Library Premium 2005). Although nowadays Sakha are settled over most of this territory, at the time of first Russian contact in the 17th century (the Yakutsk fort was founded in 1632) the Sakha were concentrated mainly in a fairly small area of central Yakutia, between the Lena, Amga and Aldan rivers (Dolgix 1960: 377, cf. Figure 1.2). Thus, their expansion over the large area they inhabit today occurred quite recently, in the 17th and 18th centuries (Dolgix 1960: 360ff; Forsyth 1992: 63; Wurm 1996a: 971f).

---

[4] Of course, it is not quite clear what the label владеющий русским языком ('knowing Russian') really entails; whether this indicates just a basic knowledge of Russian or whether some degree of fluency is required. Judging from my own field observations, the percentage of fluent Russian speakers in rural areas is certainly lower than 80% when children are included in the count.
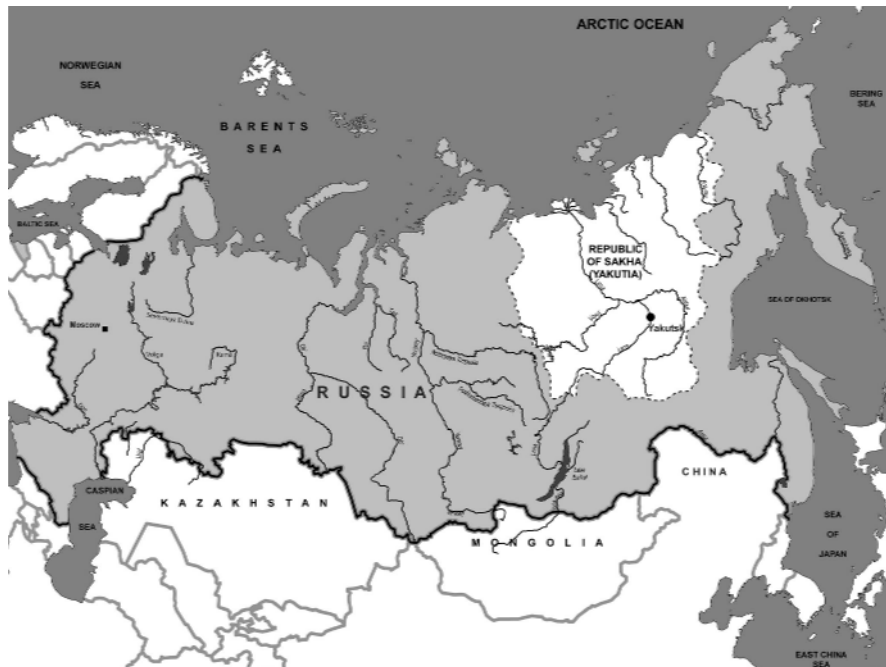
Figure 1.1: The location of the Republic of Sakha (Yakutia) within the Russian Federation. © MPI for Evolutionary Anthropology.

The main mode of subsistence among the Sakha is cattle- and horse-breeding; since the collapse of the Soviet Union this is practised on the level of basic subsistence economy. Both cattle and horses are kept for meat, cows in addition providing milk, which is the basis of many Sakha food products, especially in late spring and early summer. In addition, hunting of game and fowl as well as fishing supplement the economy. Cattle are kept in barns during the winter and throughout that time (often seven to eight months) need to be fed with hay; therefore, hay-making is the most important event in the Sakha calendar. The Sakha horses, however, are able to fend for themselves even in winter, when they dig in the snow for fodder (in temperatures reaching −50° C and below). They are half-wild and roam free practically all year; only in early spring are mares brought to enclosures to ensure their safety at the time of foaling (personal observation).

Figure 1.2: The approximate distribution of the language families of Siberia at the time of first Russian contact. Map adapted from Dolgix (1960) and Wurm et al. (1996: map 106). © MPI for Evolutionary Anthropology.

1.1.1.1 The Sakha language

The Sakha language clearly belongs to the Turkic language family, with a large number of basic words (numerals, words for body parts, kinship terms, and some livestock terminology) and the nominal inflection being retained to a large degree. However, there exist several differences between Sakha and Common Turkic[5] as well, such as a number of sound changes, a large amount of Mongolic lexical copies, and differences in the verbal TAM system, so that mutual comprehension between speakers of other Turkic languages and Sakha is low

---

[5] Common Turkic designates the Turkic languages with the exception of Chuvash and possibly Khalaj (Johanson 1998b: 81; Lars Johanson, pers. comm.).

(Stachowski & Menz 1997). These differences, and especially the large number of copied Mongolic words, led Radloff (1908) to suggest that Sakha was initially a language of unknown affiliation that was mongolicized and only later became turkicized – a view that cannot, however, be supported nowadays.

Turkic languages are spoken over a very large area of Eurasia, from Manchuria and northeastern Siberia in the east (Fuyü and Sakha, respectively) to Anatolia, Moldavia and Lithuania in the west (Turkish, Gagauz and Karaim, respectively), and from the Taimyr Peninsula and the coast of the Arctic Sea in the north (Dolgan and Sakha) to Iran in the south (Khalaj and Qashqa'i). The Turkic language family is sometimes classified as one of the branches of the disputed Altaic language family, together with Mongolic and Tungusic, and, even more controversially, Korean and Japanese (Comrie 1981: 39ff; Ruhlen 1991: 328f; Janhunen 1996: 237ff; Kortlandt [2004] 2006; Robbeets 2005: 423). Due to large-scale population movements in the history of the Turkic peoples, the genealogical classification of the individual languages is not straightforward, since areal influence cuts across genealogical relationships. Thus, the currently accepted classification of the Turkic languages comprises three branches that are defined through genealogical relatedness as well as one branch that is defined mainly by the geographic proximity of the languages involved; in addition, two further branches are represented by individual languages (Chuvash and Khalaj). The three branches defined primarily on genealogical grounds (Schönig 1997: 123; Johanson 1998b: 82f) are: southwestern Oghuzic (with Anatolian Turkish, Azerbaijanian, Turkmen and Gagauz as the main representatives), northwestern Kypchakic (including, amongst others, Kazakh, Kirghiz, and Tatar), and southeastern Uighuric (Uzbek, Uyghur, and Yellow Uyghur, to name a few). The Siberian Turkic languages (Altai-Sayan Turkic in the south and Lena Turkic – Sakha and Dolgan – in the north) are genealogically heterogenous and are grouped together mainly on geographical grounds. Chuvash and the very archaic Khalaj are the sole representatives of the Oghuric and the Arghu branch, respectively[6] (Johanson 2001: 1720). Chuvash is the only living descendant of the language of the Turkic Bolgars, a group that split off from the remainder of Turkic peoples in the first half of the first millennium AD (Golden 1998: 18; Johanson 1998b: 81). Four languages, Sakha and Dolgan, Chuvash, and Khalaj are very divergent, indicative of an early separation from the remainder of the Turkic languages (Schönig 1997: 120). Sakha has only one close relative, namely Dolgan, a language spoken by a group of mixed ethnic origins on the Taimyr Peninsula (Ubrjatova 1966). Dolgan is structurally close enough to Sakha that it is

---

[6] However, Ščerbak (1994: 29ff) includes Khalaj in the Oghuzic group.

sometimes classified as a dialect of the latter (Voronkin 1999: 154); however, due to a large number of lexical differences (changes in the semantics of shared lexical items, innovations, Evenki lexical copies) and phonetic changes there is only a low degree of mutual intelligibility. Its classification as a separate language has therefore both linguistic (Ubrjatova 1966) and sociopolitical grounds (Artem'ev 1999a: 45).

It seems that at least two different Turkic languages have contributed to the Sakha language. One might have been related to the language of the Orkhon inscriptions, as can be seen from many retentions of Old Turkic features; the other may have been a Kypchak language, as seen by some shared features between Kypchak (especially Kirghiz) and Sakha (Širobokova 1977; Ubrjatova 1985: 24; Schönig 1990; Stachowski & Menz 1997; Gogolev 1993: 44f). Although the language is quite homogenous – a further confirmation of the relatively recent spread over the vast area of current settlement – there are some dialectal differences, which are grouped into four major dialectal groups: the central group, the Vilyuy group, the northwestern group, and the northeastern group (Voronkin 1999: 154f). The dialectal differences are assumed to be due to different substrate influences (especially Evenki influence in the northwest), and also to isolation of the inhabitants of individual regions from one another (Voronkin 1999: 30f). The most salient feature of the dialectal system is a phonetic difference in approximately 200 words which in some dialects are pronounced with unrounded vowels (*akan'e*[7] in the Sakha linguistic literature), while in others they are pronounced with rounded vowels (*okan'e*), e.g. *χatïn/χotun* 'housewife', *a:ɣïy/o:ɣuy* 'spider', *seri:n/sörü:n* 'cool' (Voronkin 1999: 57). These are words which in Common Turkic or Mongolic (in the case of copying) contained labially unmatched vowels, i.e. the first syllable was unrounded, while the vowel of the second syllable was rounded, such as *qatun* 'housewife'. Such words go against the Sakha system of labial vowel harmony, in which all vowels must be either rounded or unrounded. In order to resolve this discrepancy, in some areas the second vowel assimilated to the quality of the first vowel (*akan'e*), while in others the first vowel assimilated to the second vowel (*okan'e*). This development is presumably a fairly recent event: in Dolgan, which follows the same labial harmony as Sakha, some of these words have retained their ancient pronounciation, e.g. *katun* (Sakha *χatïn/χotun* 'housewife'). Since the ancestors of the Dolgans still lived in contact with Sakha in the beginning of the 17th century, the retention of labially unmatched words in Dolgan indicates that *akan'e* and *okan'e* in Sakha must have developed later than that (Ubrjatova 1960: 40f). Central Yakutia (i.e. the area of initial settlement by the Sakha) is split among

---

[7] I adopt the Russian-Sakha linguistic terms as they offer a useful way of briefly designating the chief difference in the pronounciation of these words.

dialects showing *akan'e* in the north and those with *okan'e* in the south (Voronkin 1999: 20f), a split that some researchers attribute to Mongolic substrate in the dialects with *akan'e* (Ubrjatova 1960: 42; Širobokova 1980; Voronkin 1999: 57ff; Gogolev 1993: 58, 61f). In Yakutia as a whole, the northeastern region belongs to the dialects with *akan'e*, while the Vilyuy and northwestern areas belong to the *okan'e* dialects (Voronkin 1999: 57f).

The majority of the Mongolic lexical copies in Sakha cannot be assigned to one specific modern Mongolic language; rather, they show similarities to Middle Mongolian/Written Mongolian of the 13th and 14th century (Popov 1986: 46ff; Kałużyński 1962: 39f). Mongolic lexical copies are widespread in all semantic domains, being found amongst designations of social relations, e.g. *jon* 'people, relatives, family' (Pekarskij 1958 [1912]: 840), *eme:χsin* 'old woman, wife', *kergen* 'family, spouse' (Kałużyński 1962: 26, 28); body parts, e.g. *bīlčarχay* 'gland', *berbe:key* 'ankle bone', *čančĭk* 'temple' (Kałużyński 1962: 19, 25, 135); or livestock terminology, e.g. *süöhü* 'livestock', *meččiy* 'graze', *dal* 'corral' (Kałużyński 1962: 35, 40, 44); furthermore, a number of descriptive verbs are copied from Mongolic languages as well, such as *jirbey* 'be tall and slim, appear excessively tall' and *sïntay* 'having a turned-up nose' (Kałużyński 1962: 139, 149).

Sakha does not have a long literary tradition: the first text books in Sakha were published based on a writing system devised by S.A. Novgorodov in the 1920s; this writing system was exchanged for a unified Turkic alphabet in 1929, which in 1939 was replaced by the Russian-based Cyrillic alphabet still in use today (Voronkin 1999: 35). In the early 1930s the Sakha standard language was officially based on the dialects of the districts around Yakutsk: Kangalas, Namcy and Megin, with *okan'e* and word-initial [s] as its most salient features (Voronkin 1999: 39f).

1.1.1.2 Origins of the Sakha

There is a general consensus that the Sakha are not indigenous to Yakutia, but immigrated from an area further to the south. This can be seen both from their Turkic language and their subsistence pattern of cattle and horse pastoralism. Their ancestors are identified as the Kurykans known from Chinese chronicles and archaeological finds on the shores of Lake Baykal in South Siberia, whose culture is dated to the 6th to 10th century AD. Judging from runic inscriptions found in conjunction with these archaeological sites, the Kurykans are presumed to have been a Turkic-speaking population (Okladnikov 1955; Konstantinov [1975] 2003; Širobokova 1977; Gogolev 1993; Alekseev 1996). The main mass of Turkic-speaking Sakha ancestors is taken to have immigrated to the middle reaches of the

Lena river in the 13th or 14th century (Gogolev 1993: 61, 88f; Alekseev 1996: 46), although, as shown by a runic inscription on the Lena dated to the 9th or 10th century AD, some small scattered groups reached this area already at the end of the first millennium (Okladnikov 1955: 326ff; Konstantinov [1975] 2003: 18f; Alekseev 1996: 28, 45f). Okladnikov (1955: 332, 365) and Alekseev (1996: 45f) propose that cultural and ethnic contacts between the indigenous inhabitants of Yakutia (in their view, mainly Yukaghirs) and the Turkic-speaking immigrants started at that time; while Konstantinov ([1975] 2003: 19) rather assumes that these initial Turkic-speaking groups were very small and had no influence on the local populations.

Okladnikov (1955: 289), Gogolev (1993: 94, 96) and Alekseev (1996: 35, 45) assume that the immigrating Turkic-speaking groups interacted with the indigenous inhabitants of Yakutia, while Konstantinov ([1975] 2003: 68f) claims that the immigrating group of Turkic-speakers did not admix with local populations. However, the degree of substrate influence postulated by Gogolev and Alekseev is quite different: the former sees the south Siberian cultural elements as clearly predominant (Gogolev 1993: 122), while the latter claims that indigenous groups played a major role in the formation of the Sakha culture and ethnic identity (Alekseev 1996: 45); furthermore, while Gogolev (1993: 126) sees admixture predominantly with Tungusic groups, Alekseev (1996: 48) denies any notable contact with Tungusic-speakers, claiming a predominant role for 'Paleoasiatic' groups (mostly Yukaghirs) in Sakha prehistory[8].

Given the large number of Mongolic substance copies in the Sakha language (Kałużyński 1962, *passim*; Pakendorf & Novgorodov, in preparation), it is obvious that the Sakha ancestors were in close contact with Mongolic-speaking groups. Most of the Mongolic copies cannot be traced to any specific Mongolic language, which may be an indication that they were in contact with several dialects over a long period of time, from approximately the 12th/13th century up to the 15th or even 16th century (Kałużyński 1962: 122, 126); however, Širobokova sees close ties with Buryats (Širobokova 1980: 143, 146). Some Mongolic-speaking tribes are presumed to have been assimilated by the Turkic-speaking Kurykans in the 6th-10th centuries AD (Gogolev 1993: 44), but the main contacts must have taken place later than that. Mongolic-speaking tribes are believed to have migrated to Lake Baykal in the 11th century under pressure of the expanding Khitans in Mongolia, leading to an

---

[8] It should be noted that for most of the time period and geographical area under consideration there exist only archaeological data. In the absence of inscriptions (which are, however, found only in southern Siberia), these data do not contain any indication of the language spoken by the producers of the cultural artefacts. Therefore, a lot of the work on Sakha prehistory remains quite speculative.

extended period of joint settlement and cultural contact between the Turkic-speaking ancestors of the Sakha and the Mongolic immigrants (possibly the current-day Buryats). Based on archaeological data as well as epic tales and legends, the Sakha ancestors are assumed to have left the Baykal area only in the 13th century to avoid Mongol military campaigns against the Yenissey Kirghiz and others (Konstantinov [1975] 2003: 70) or as a result of ethnic clashes with Mongolic-speaking tribes (Gogolev 1993: 61). However, the period between the 6th and 13th centuries AD was one of continuous tribal conflict and upheaval involving large-scale population movements in South Siberia. Thus, from the middle of the 6th century a series of Turkic Empires existed in modern-day Mongolia that were engaged in continuous warfare with their neighbours, leading to a number of population displacements in South Siberia (Spuler 1966: 132, 138, 159). From the 10th century onwards, Mongolia was conquered by the Khitans (an ethnic group of as yet unknown linguistic identity – Janhunen 1996: 139ff), who themselves were displaced by the Tungus-Manchu-speaking Jurchen in 1125 (Spuler 1966: 188). The Jurchen were displaced less than a century later by the rising Mongol Empire. It is therefore quite possible that the Turkic-speaking ancestors of the Sakha migrated north at any time during this period in order to evade the warfare and political domination imposed by the successive tribal dynasties in Mongolia/South Siberia.

A further possible source of the Mongolic copies could be a Mongolic-speaking group settled on the Lena before the arrival of the Turkic-speaking Sakha ancestors (Dolgix 1960: 498; Janhunen 1996: 162). Thus, Ubrjatova (1960: 42) claims that there must have been Mongolic-speaking groups in the northern areas of Central Yakutia contemporary with the Sakha, whose later shift from Mongolic to the Turkic language explains the development of *akan'e* (cf. section 1.1.1.1).

Sakha epic tales agree with the archaeological, linguistic, and ethnographic data in depicting the Sakha ancestors as having immigrated from the south. They mention three legendary heroes as the ancestors of the Sakha: the first, Omogoj, is viewed as personifying the Turkic-speaking Kurykans; he is depicted as arriving on the Middle Lena before the others. The second legendary hero is Ėllej who is often depicted as being of Tatar or Kirghiz origin; he is shown as arriving on the Lena later, and as being the 'Kulturträger' of the Sakha and the founding father of nearly all Sakha clans. Only two of the Sakha clans (the Namcy and Bajagantaj *ulus*[9]) are claimed to have descended from Omogoj (Konstantinov [1975] 2003: 44f; Gogolev

---

[9] A continuation of the original clan system is retained in the administrative division of the Republic, which is divided into 33 districts, or *ulus*, which is the Sakha word for 'clan'. Thus, it is possible that in Central Yakutia descendants of individual clans are settled predominantly in the corresponding districts.

1993: 117f). The third hero, who does not feature in the legends as much as the other two, is Uluu-Xoro who is identified with a Mongolic tribe, the Xoro; he appears in Yakutia later than Omogoj and Ėllej and may represent a third immigration into Yakutia by Mongolic-speakers who further influenced the Sakha language; this could explain the relatively young age of Mongolic copies into Sakha (Gogolev 1993: 119).

A previous molecular-genetic study of the Sakha (Pakendorf et al. 2002, Pakendorf et al. 2003) indicated female Tungusic and Mongolic admixture in the Sakha and a strong bottleneck undergone by the men. Unfortunately, due to lack of comparative data, the origins of the Sakha men (who appear quite divergent from Finno-Ugric speaking groups, Buryats, and Russians) couldn't be elucidated. These genetic results are indicative of either a small group of Turkic-speaking men intermarrying preferentially with Tungusic-speaking women (if the Sakha men should be shown to be of Turkic origin), or of a case of language shift of an originally Tungusic-speaking population after a severe reduction of the male population – in the case that the Sakha men should be of Tungusic origin (Pakendorf 2001). One of the most interesting genetic features of the Sakha is the very high frequency of men carrying the Y-chromosomal single nucleotide polymorphism (abbreviated as SNP) Tat C (Pakendorf et al. 2002, 2006). Tat C belongs to the group of slowly evolving markers (also called 'unique event polymorphisms') of which it is assumed that they arose only once in human prehistory; therefore, sharing of the derived state at such a polymorphic site (such as Tat C) indicates shared ancestry (or admixture). Tat C is found predominantly in northern Eurasia, with a distribution from Finns and Saami in the west to Eskimos in the east (Lahermo et al. 1999; Karafet et al. 2002). In South Siberian Turkic groups it is present in approximately 10%, with a range of 2% in Shors to 25% in Tofa (Derenko et al. 2006). In Mongols it is found in low frequencies of 2-6% (Karafet et al. 2002; Derenko et al. 2006), while in Buryats the frequency is much higher: between 19% and 58% (Zerjal et al. 1997; Karafet et al. 2002; Derenko et al. 2006). This could be indicative of a shared substrate in Tofa, Buryats and Sakha. However, comparison of short tandem repeats (STRs) on Sakha Tat-C-carrying Y-chromosomes with those from other populations (mainly Finno-Ugric groups and Buryats) showed a striking divergence between Sakha and others (Pakendorf et al. 2002, 2006). Although the frequency of Tat C is quite high in Finno-Ugric populations (Lahermo et al. 1999), among Samoyedic-speaking groups the distribution is uneven, with a range of 0% in Selkups to 51.7% in Forest Nenets (Karafet et al. 2002). Since the easternmost Samoyedic groups, the Selkups and Nganasans, practically lack Tat C (it is present in Nganasans with a frequency of only 2.6%), a Samoyedic origin of the Sakha men is rather unlikely. Thus, the origins of Sakha men still remain a mystery.

1.1.2 Evenks and Ėvens

The Evenks and Ėvens, who speak closely related Tungusic languages, are spread over a large area of Central and Eastern Siberia, notwithstanding their relatively small number. Thus, according to the census of 2002, there are 35,527 Evenks and 19,071 Ėvens in the Russian Federation. The total number of speakers of Evenki is given as 7,584, and the total number of speakers of Ėven is given as 7,168, suggesting that a maximum of 21.3% of Evenks and 37.6% of Ėvens still speak their heritage language[10] (Federal'naja služba gosudarstvennoj statistiki 2004: 19, 124). The main areas of settlement of Evenks are between the Nižnjaja and Podkamennaja Tunguska in the west, the upper reaches of the Lena, Barguzin, Vitim, and Olëkma rivers with the northern tributaries of the Amur in the southwest, and the Lower Amur, the Oxotsk Sea coast as well as some areas of Sakhalin in the southeast (Atknine 1997: 110, cf. Figure 1.3). Ėvens are settled in several areas of northeastern Yakutia, predominantly between the Yana and Kolyma rivers, along the Oxotsk Sea coast, and on Kamchatka (Novikova 1960: 9); however, the latter represent a very recent immigration (Severnaja Ėnciklopedija 2004: 1114; Wurm 1996a: 972f; cf Figure 1.2). Evenks and Ėvens are traditionally fully nomadic reindeer-herders and hunters; until sovietization, the domesticated reindeer were kept predominantly for transport, while subsistence was based on fishing and hunting wild reindeer. Reindeers are mainly ridden and used as pack-animals, which distinguishes the Evenks and Ėvens from Samoyedic reindeer herders in Western Siberia, such as the Nenets, although sleds are used by Ėvens living in the forest-tundra and on Kamchatka as well (Novikova 1960: 13; Severnaja Ėnciklopedija 2004: 1106, 1114, 635).

---

[10] These figures are lower than those given by the sociolinguistic encyclopedia Pis'mennye jazyki mira (2003: 640, 642, 667, 668); here, of 29,901 Evenks in the Russian Federation (data from the 1989 census), 9891 (i.e. 33%) are said to speak their heritage language, while of 17,055 Ėvens 7850 (i.e. 46%) are claimed to have retained their heritage language.

Figure 1.3: The approximate current-day distribution of the languages of Siberia. Map adapted from Wurm et al. (1996: map 109). © MPI for Evolutionary Anthropology.

1.1.2.1 Tungusic languages

Evenki and Èven belong to the Northern Tungusic branch of the Tungusic language family. Although the relationship of the languages belonging to this family is widely accepted, the internal classification of the Tungusic language family as a whole has not yet been unanimously resolved. One reason for the difficulties besetting the classification of the Tungusic languages is their shallow time depth and, similar to the Turkic languages, the nomadic lifestyle of some of the groups. This brought groups speaking different dialects and different languages into contact with each other, and also into contact with speakers of different languages (Whaley et al. 1999: 289, 313). Thus, Sunik (1968: 54) postulates two main branches: Manchu (consisting of the extinct Jurchen language on the one hand, and Manchu with its dialect Sibo on the other) and Tungusic. The latter he splits into two

branches, Northern Tungusic (also called the Siberian, or Evonki, group) with the languages Evenki, Solon, Negidal, and Ėven; and Southern Tungusic (also called the Amur, or Nanay, group) with the languages Nanay, Ulča, Orok, Oroč, and Udihe (Sunik 1968: 54). Comrie (1981: 58) also postulates two main branches; however, instead of grouping the Siberian Tungusic with the Amur Tungusic languages, he postulates a primary split between Northern (Siberian, Evenki) Tungusic and the other languages (the Southern Tungusic branch), with the latter comprising a southwestern branch (Manchu and Sibo, as well as Jurchen), and a southeastern branch consisting of the Amur Tungusic languages. Janhunen (1996: 78) prefers to "[…] recognize four main branches, corresponding to the four languages of Manchu, Nanai, Udeghe and Ewenki (with Ewen)", a classification also followed by Tsumagari (1997: 175; see also Kortlandt [1998] 2006). A further classification postulates three main branches, Northern Tungusic, Amur Tungusic, and Manchu (Atknine 1997: 111). However, according to Janhunen (1996: 78) the genealogical validity of Amur Tungusic is not clear, especially the position of Udihe relative to Evenki and Nanay. Another classification is that of Doerfer (1978), which is accepted to some degree by Whaley et al. (1999). This classification also argues for three primary branches, here called Northern, Central, and Southern Tungusic, with the Northern branch split into a Northeastern (Ėven and Arman) and a Northwestern group (the latter consisting of Evenki, Solon and Negidal). The Central branch is split into a Central-Eastern group containing Oroč and Udihe, and a Central-Western group consisting of Kili, Nanay, Ulča and Orok, while the Southern branch contains Jurchen and Manchu. However, what distinguishes Doerfer's classification from those of others is that he doesn't postulate a binary family tree model, but rather proposes a network, with some languages or dialects being in transition to others, e.g. the Western dialect of Ėven is depicted as being in transition to Evenki (though still closer to Ėven) (Doerfer 1978: 4, 5). One of the conclusions Whaley et al. (1999: 313) come to in their paper is that the Northwestern Tungusic languages, and possibly the entire Tungusic language family, cannot be classified using the traditional family tree model, since on the one hand contact influence has led to diffusion of features between different dialects and families, and on the other hand the shallow time depth of the language family means that the languages are too similar, so that sound correspondences do not define clear groups. Throughout the following, I will for practical purposes refer to Evenki, Ėven and Negidal as the Northern Tungusic languages, and to Nanay, Ulča, Orok, Udihe and Oroč as the Amur Tungusic languages, without the intention of making any genealogical claims.

Among the Northern Tungusic languages, Evenki, Solon and Negidal are very closely related (to the extent that Solon and Negidal can be classified as Evenki

dialects), even though Solon and Negidal are spoken in Manchuria and on the Lower Amur, respectively (Janhunen 1996: 72f; cf. Figure 1.3). It is sometimes claimed that the Negidals are the descendants of the Evenks (Black 1988: 25; Forsyth 1992: 207; Janhunen 1996: 67, 72f, 79 *inter alia*); Xasanova & Pevnov (2003: 285) however suggest that Evenki and Negidal are descendants of a common ancestor, rather than Negidal being a descendant of Evenki. Furthermore, the Evenki dialects spoken on the Chinese side of the Amur river are often classified as a separate language, Oroqen (Atknine 1997: 114). Among the Amur Tungusic languages, Nanay, Ulča and Orok can be grouped together as forming a dialectal continuum, while Oroč can be classified as a dialect of Udihe (Janhunen 1996: 62f, 65). Ethnic Manchu are confined to China, while the Amur Tungusic peoples live in the Russian Far East on the Lower Amur and the Japanese Sea Coast. As mentioned above (section 1.1.2), the Northern (Siberian) Tungusic Evenks and Ėvens are spread over a huge territory from the Yenissey river to the Oxotsk Sea.

All the Tungusic languages consist of several dialects, some of which are different enough to be classified as distinct, though closely related languages (Sunik 1962: 21f). Evenki is grouped into three dialectal groups, each of which consists of several dialects; 51 dialects are recognized in total. The three dialect groups are distinguished mainly by their phonetic realization of the phoneme /s/: in the northern dialect group (spoken in the north of the Evenk National District) [h] is spoken in word-initial and in intervocalic position, e.g. *hulaki:* 'fox', *ahi* 'woman', while in the eastern dialect group (spoken in the Far East as well as in the south of Yakutia), [s] is spoken word-initially, while in intervocalic position [h] is spoken, e.g. *sulaki:* 'fox', *ahi* 'woman'. The southern dialect group (spoken in the southern areas of the Evenk National District and north of Lake Baykal) comprises two subgroups, the 'hissing' subgroup in which [s] is spoken both in word-initial and in intervocalic position, e.g. *sulaki:* 'fox', *asi* 'woman' and the 'hushing' subgroup where /s/ is pronounced [š] word-initially and intervocalically (Sunik 1962: 22; Nedjalkov 1997: xixf; Bulatova & Grenoble: 1999: 3; Atknine 1997: 117). The Evenki standard language is based on the Podkamenno-Tunguska dialect of the southern dialect group (Nedjalkov 1997: xx; Atknine 1997: 117).

Ėven, too, is classified into three major dialect groups, eastern, central, and western. The eastern dialect group, which has [s] in intervocalic position and word-finally as well as [ə] in non-first syllables, is spoken from the Kolyma river to the Oxotsk Sea coast and on Kamchatka. The central dialect group, characterized by [h] in intervocalic and word-final position and [ə] in non-first syllables, is spoken predominantly along the Indigirka river. The western dialect group, which is characterized by [h] both intervocalically and word-finally as well as [o] in non-first

syllables, is spoken in northern Yakutia from the Lena to the western half of the Yana-Indigirka watershed. The standard language is based on the eastern dialect group, predominantly on the Ola dialect (Novikova 1960: 17ff).

1.1.2.2 The origins of the Evenks and Ėvens

"In view of the amazing linguistic unity of the whole Ewenki-Ewen complex over the vast extenses of Siberian taiga between the Lower Yenisei in the northwest and the Amur in the southeast, it is clear that the modern Northern Tungusic ethnic groups were formed relatively recently by diffusion of population and language from a single limited source." (Janhunen 1996: 167f)[11]

There exist two divergent hypotheses concerning the origins of the Evenks and Ėvens. According to Vasilevič (1969: 39-41; also summarized in Alekseev 1996: 39f), the Tungus-Manchu peoples take their origins from neolithic hunters living to the south of Lake Baykal. The ancestors of the Manchu split off first from this ancestral group and moved to the Amur-Ussuri region at the end of the first millennium BC, while the ancestors of the Amur and Northern Tungusic groups moved north into the mountainous forests near Lake Baykal, where they were in continued contact with other groups throughout the Neolithic. In the middle of the first millennium AD the arrival of Turkic groups on the shores of Lake Baykal split the ancestors of the Northern Tungus (Evenks and Ėvens) into a western and eastern group; this led to their migration north and initiated the formation of the Evenks and Ėvens as separate peoples without contact with the Tungusic-speaking groups from the Lower Amur.

A different view holds that the ancestors of the Tungus-Manchu peoples originated in Manchuria, since in this region all the different branches of the Tungusic language family are attested (Janhunen 1996: 169). Janhunen suggests a medieval origin of the Northern Tungusic groups on the Middle Amur, who might have dispersed from there under pressure from immigrating Mongolic groups (the later Dagur). Based on Evenki dialectal features (such as the retention of archaic features, or the number of Mongolic lexical copies) Janhunen suggests that the northern expansion of the Evenks and Ėvens (and related Negidals and Solon) took place in two waves, an outer and an inner wave. The outer wave led to the formation

---

[11] It is interesting to note in this respect that the northern Tungusic groups are characterized by high frequencies of the Y-chromosomal SNP M86, which leads to their forming a cluster in multi-dimensional scaling analyses based on pairwise Fst values (data from Karafet et al. 2002, cf. Pakendorf et al. 2007 and Appendix 2).

of the Cisbaikalian Evenks and the Ėvens, while the inner wave resulted in the Transbaikalian Evenks (Janhunen 1996: 169f). Tugolukov (1980) locates the ancestors of the Tungus (presumably implying both Evenks and Ėvens) between the upper reaches of the Verxnjaja Angara and Olëkma rivers (i.e. in a more northwesterly location than Janhunen), where a group of reindeer-herders called *Uvan'* are mentioned in chronicles of the 5th to 7th century AD (Tugolukov 1980: 157). The further expansion of the ancestors of the Evenks and Ėvens to the north is assumed to have taken place fairly late, in the 12th or 13th century AD (Tugolukov 1980: 168; Janhunen 1996: 171). The Northern Tungusic groups spread over their current area of settlement in three waves; in the first wave they settled on the middle reaches of the Lena and the Aldan river before the arrival of the Sakha ancestors in the 13th century; in the second wave they spread down the Lena and up the Aldan under pressure of the immigrating pastoralist Turkic-speaking groups, and lastly the expansion of the Sakha in the 17th and 18th century further displaced Tungusic tribes to peripheral areas (Vasilevič 1969: 17; Tugolukov 1980: 168).

Even though the 'stereotype' of the Tungus is one of reindeer-herding hunters, in historic times Northern Tungusic peoples were classified in three different groups based on what animals they used for transport: horses, reindeers, or dogs (Vasilevič 1969: 19-21). Thus, a subgroup of Evenks in Manchuria, the Oroqen, are classified as Horse Tungus, while the Negidals are classified as Dog Tungus (Janhunen 1996: 109). The 'typical' Evenk and Ėven feature of reindeer-herding is generally regarded as a fairly late development, and is suggested to have been initiated under the influence of horse-breeding (Tugolukov 1980: 157; Janhunen 1996: 171).

1.1.3 The Yukaghirs

The Yukaghirs are a small remnant of what used to be a much larger group of probably related peoples; thus, judging from tribute documents dating to the 17th century, at the time of first Russian contact there were approximately 4,800 Yukaghirs and related peoples settled in a fairly large area of northeastern Yakutia (Dolgix 1960: 615; Figure 1.2); in the first half of the 20th century, there were only approximately 440 left (Evstigneev 2003: 140). The information concerning the current numbers of Yukaghirs and Yukaghir speakers is contradictory: according to Vakhtin (1992), a sociolinguistic survey conducted in 1987 counted approximately 350 Yukaghirs in three villages in the Republic Sakha (Yakutia), of whom about 120 (~ 35%) spoke the language; however, language retention was much higher among Tundra Yukaghirs (approximately 43%) than among Kolyma Yukaghirs

(approximately 22%) (Vakhtin 1992; Vaxtin 2001a: 142ff, 158f). In contrast to these figures, according to the 2002 census there are 1,509 Yukaghirs in the Russian Federation, of which 1,097 live in the Republic Sakha (Yakutia), i.e. a number three times as high as that given by Vakhtin (1992), while the Yukaghir language is claimed to be spoken by 604 individuals (Federal'naja služba gosudarstvennoj statistiki 2004: 19, 113, 124). Compact Yukaghir settlements are found in only three villages in the Republic Sakha (Yakutia): Andrjuškino and Kolymskoe in the Lower Kolyma district, and Nelemnoe in the Upper Kolyma district, (Maslova 2003a: 1f; Maslova, pers. comm.), as well as in two settlements in the Magadan region (Vakhtin 1992).

Traditionally, the southern (Kolyma) Yukaghir groups (who lived on the upper reaches of the Kolyma, Indigirka and Yana rivers) were hunters and fishermen, who used skis, hand-pulled sleds, and dogs for transport purposes. The northern Tundra Yukaghir groups were fully nomadic reindeer herders who had adopted domesticated reindeer from Èvens; their main source of food were wild reindeer, while the domesticated reindeer were used predominantly for transport. A third, small group of Russianized Yukaghirs led a sedentary lifestyle on the Anadyr' river, where they fished and hunted wild reindeer during the spring and autumn migrations (Gurvič & Simčenko 1980: 149ff; Jochelson [1926] 2005: 92ff, 103f).

1.1.3.1 The Yukaghir languages

Although it is assumed that there were several Yukaghir languages spoken at the time of first Russian contact (Gurvič & Simčenko 1980: 147; Kurilov 2005: 9f), nowadays only two Yukaghir languages remain. These are Kolyma (or Southern) Yukaghir and Tundra (or Northern) Yukaghir, which until recently were classified as dialects of one language. In 1987, Kolyma Yukaghir was spoken in the village Nelemnoe in the Upper Kolyma district of the Republic Sakha (Yakutia) by 29 individuals, of whom only nine older people preferred it as their primary means of communication; Tundra Yukaghir was spoken in the villages Andrjuškino and Kolymskoe in the Lower Kolyma district by 93 individuals, of whom only 30 preferred it as their primary means of communication (Vaxtin 2001a: 142ff). The genealogical affiliation of the Yukaghir languages has still not been clarified decisively; although some authors consider Yukaghir as part of the Uralic language family (cf. references in Maslova 2003a: 1), others prefer to consider it a linguistic isolate (Comrie 1981: 10, 258; Abondolo 1998b: 8).

1.1.3.2 The origins of the Yukaghirs

Not much is known about the origins of the Yukaghirs, but in general it is assumed that they represent the descendants of peoples inhabiting northeastern Siberia since at least the Neolithic (Gurvič & Simčenko 1980: 144, 146). According to the scenario proposed by Alekseev (1996: 39), the ancestors of the Yukaghirs originated in the Taimyr Peninsula in neolithic times, with a mixing of cultures from Western Siberia and Yakutia. Approximately in the middle of the second millennium BC the Yukaghir ancestors spread from the Taimyr Peninsula to the east under pressure of immigrating groups (rather speculatively identified by Alekseev as Yenisseic-speakers) and reached Chukotka about 1,000 years later. In the first half of the second millennium AD the expansion of Evenki groups to the northwest cut off the Yukaghirs from Samoyedic-speaking groups in the west and forced them even further to the east, where they ended up surrounded by Chukchi, Koryaks, Ėvens and the ancestors of the Sakha. After contact with Russians in the 17th century they were gradually decimated by attacks of Russian cossacks and Chukchi, by smallpox epidemics and by episodes of starvation (Dolgix 1960: 383, 408, 409, 415; Jochelson [1926] 2005: 99f), and assimilated by their neighbours.

If the genealogical relationship of the Yukaghir languages and the Uralic language family is true, and if the hypotheses about the age and origin of the Uralic languages are correct, then Yukaghirs can justifiably be assumed to have inhabited northern Siberia for a very long time (cf. Fortescue 1998: 183, 193, map 5, 6; Kortlandt [2004] 2006: 4). Thus, the 'Urheimat' of the Uralic language family is assumed to have been located somewhere near the southern end of the Ural mountains, and the primary split of the Uralic language family into the Samoyedic and Finno-Ugric languages is estimated to have taken place at least 6,000 years ago, with the Samoyedic-speakers migrating to the north and east (Abondolo 1998b: 1f). Thus, proto-Yukaghirs would have had to split off from the bulk of the family at least at that time, if not earlier (cf. Kortlandt [2004] 2006: 5). A reason for an even earlier migration of proto-Yukaghirs to the east may lie in the fact that eastern Siberia was not covered by glaciers to the same extent as western Siberia, so that an earlier settlement of the northern regions was possible (Simčenko 1980: 25; Gurvič & Simčenko 1980: 148).

As mentioned in section 1.1.1.2, a genetic feature that unites a large number of peoples of northern Eurasia, and that may have some bearing on the matter of Yukaghir origins, is the Y-chromosomal SNP called Tat C. This is found predominantly in northern Eurasia, with a distribution from Finns and Saami in the west to Eskimos in the east. Finno-Ugric-speaking populations are characterized by high frequencies of this polymorphism (Lahermo et al. 1999), as are the Forest and

Tundra Nenets and the Yukaghirs (Karafet et al. 2002). Fine-scaled analyses of Tat-C-bearing Y-chromosomes show that the Yukaghirs share Tat C haplotypes with other populations (such as Tuvans, Buryats, and Finno-Ugric groups), but not with Sakha; therefore, Tat C in Yukaghirs is not due to recent admixture with Sakha (Pakendorf et al. 2006, 2007). Since the Samoyedic-speaking Nganasans and Selkups lack Tat C (Karafet et al. 2002), a specifically Uralic connection of the Yukaghirs is not evident from the presence of Tat C in the latter; however, the distribution of this polymorphism does show that even in prehistoric times population movements over the vast expanses of Eurasia were possible.

1.1.4 Mongolic groups

Given the large number of Mongolic substance copies in Sakha, it is clear that there must have been a period of intense contact between the Sakha ancestors and one or more Mongolic groups. Mongolic-speaking groups have spread only in historical times with the military expeditions of the Mongol armies; in the 12th century AD they were still settled on the territory of modern-day Mongolia (Janhunen 1996: 160). Nowadays, most Mongolic peoples are settled in a fairly compact area of Central Asia/South Siberia: Mongols inhabit Inner Mongolia in China and the Republic of Mongolia, Buryats are settled in the areas to the west and east of Lake Baykal, and Dagurs inhabit Manchuria. Oirats are settled in western Mongolia and China, with one exception: a subgroup of Oirats, the Kalmyks, migrated to the west in the 17th century and settled along the lower Volga (Comrie 1981: 56). Finally, some outlying groups are settled in China (Santa, Bonan, and Monguor), and one outlying group, the Moghol, is settled in northwestern Afghanistan (Comrie 1981: 55; The Mongolic Languages 2003: xxix).

1.1.4.1 The Mongolic languages

Modern-day Mongolic languages are very closely related, going back to the expansion and dispersion of Mongolic peoples during the Mongol Empire in the 13th and 14th century (Janhunen 1996: 159, 161). Thus, the time depth of the modern-day Mongolic languages is only approximately 800 years. Although there was presumably some linguistic diversity before the rise of Chinggis Khan, in the process of unifying the Mongolic tribes under his authority he also unified the language (Janhunen 1998: 203). In accordance with the origins of modern-day Mongolic diversity at the time of the Mongol Empire, the reconstructed form of

Proto-Mongolic is very close to the languages called Middle Mongolian and Written Mongolian (Weiers 1986: 32; Janhunen 1996: 145f; Janhunen 2003d: 1). Middle Mongolian, which is known from a number of sources written in different scripts from the time of the Mongol domination of China (the Yuan dynasty of the 13th and 14th century), was the unified language of the Mongol Empire. Written Mongolian, which was in use from the 13th century onwards, retains an archaic form of Mongolic which can be considered to reflect some characteristics of Late Pre-Proto-Mongolic, also called Ancient Mongolic (Weiers 1986: 31f; Janhunen 2003d: 2; Janhunen 2003a: 30)[12]. The differences between the modern Mongolic languages are due to the effects of geographical isolation as well as differential substrate and adstrate influences (Weiers 1986: 38; Janhunen 1996: 161).

After the unification by Chinggis Khan, the diversification of Mongolic languages probably began in the period from the end of the 14th century to the middle of the 16th century. Nowadays, there exist ten different Mongolic languages that can be further subdivided into dialects (Weiers 1986: 37). A major split exists between the West Mongolic languages (Oirat with several dialects and Kalmyk with several dialects) and East Mongolic languages, which are divided into three branches: South Mongol, Central Mongol and Northern Mongol or Buryat. The West Mongolic languages Oirat and Kalmyk developed their own written script in the 17th century, Written Oirat, which was in use until the 20th century (Weiers 1986: 42). The East Mongolic languages on the other hand continued to use Written Mongol as a medium of written communication. The South Mongolian dialects are spoken in Inner Mongolia in China (Weiers 1986: 67), while the Central Mongolian dialects are spoken in the Republic of Mongolia; the national language of Mongolia is based on the Khalkha dialect. The North Mongolian dialects are spoken by Buryats to the west, southeast and east of Lake Baykal, with two large dialectal distinctions being recognized, Eastern and Western Buryat (Weiers 1986: 67ff). The Buryat standard language is based on the eastern Buryat dialect Xori (Weiers 1986: 51).

At the periphery of Mongolic settlement several quite divergent languages are spoken that do not fit into the major classification of West vs. East Mongolic. One is Moghol, spoken in Afghanistan, which has undergone considerable Arabic, Turkic and Iranian influence (Weiers 1986: 53). Several peripheral languages are spoken in China in the Gansu-Qinghai area; these are Monguor, Santa, Yellow Uyghur (the Mongolic language of formerly Turkic-speaking Yellow Uyghurs), and Bonan. Lastly, Dagur is spoken in Manchuria (Janhunen 1996: 50f), with one subgroup settled in Xinjiang (Janhunen 1996: 52).

---

[12] It should be noted, however, that Doerfer (1964: 37) disagrees with this view of Written Mongolian as a particularly archaic form of Mongolian, more archaic than Middle Mongolian. In his view, archaic and innovative forms existed side by side in the written language.

1.1.4.2 Origins of the Mongols and the Mongolian Empire

In the first millennium AD the geographic area of present-day Mongolia was inhabited not by Mongolic tribes, but by Turkic tribes, who in the second half of the millennium established large and succesful tribal unions that dominated the area between the Altai-Sayan mountains in the west, Lake Baykal in the north, and northern China in the south. At that time, the Mongolic tribes were located in western Manchuria, possibly in the Greater Xingan mountains, where they may have been hunters and fishers with only rudimentary agriculture (Janhunen 1996: 136f). These Mongolic ancestors must have expanded relatively peacefully into Mongolia before the ascent of the Mongol Empire, because the unification of the Mongolic tribes and the consolidation of their Empire occurred in a territory that coincided with that of current-day Mongolia (Janhunen 1996: 160). Before the process of unification initiated by Chinggis Khan at the turn of the 12th and 13th centuries, the Mongolic peoples were a conglomerate of tribal confederations, with the individual tribes split into clans (Janhunen 1996: 158). Although there were probably dialectal differences between the individual Mongolic tribes in the 12th century, these were not big enough to hinder the communication necessary to unite them in the Mongol Empire; this unification led to the unification of the language as well (Janhunen 1996: 161). The 11th and 12th centuries were characterized by conflicts between the individual Mongolic tribes which were only ended by Chinggis Khan, who in the period from 1197 to 1205 subjugated all the Mongolic tribes, and in 1206 was declared the ruler of all the Mongols (Kämpfe 1986: 184ff). After his political and military victory, Chinggis Khan restructured the Mongol social organization, changing the basis of clans and tribes to one of a military kind. The first foreign military expeditions of Chinggis Khan's subjugated the Turkic Kirghiz and Uyghurs in 1206-1209, after which China was attacked (Kämpfe 1986: 186f). In 1218 a second military campaign was begun with the aim of subjugating the Khwarezm Turks in the west, with Samarkand and Bukhara falling in 1220, and the area up to the Dnjepr being the target of Mongolian expeditions. Chinggis Khan himself died in 1227, but his sons continued his military campaigns, extending the empire over a huge area of Eurasia, from Russia in the west and Iran and Iraq in the south to China (Weiers 1986e, *passim*). After the death of Chinggis Khan's grandson Möngke in 1259 the unified Mongol Empire split into several smaller empires: the Yüan dynasty in China, the Čagatay realm in Central Asia, the Il-Khanate in Iran and Iraq, and the Golden Horde in Russia, all of which ended in the second half of the 14th century. In the Čagatay empire and the Golden Horde Turkic languages soon took over as the main language of communication, while in the Il-Khanate Mongolian was soon replaced by Persian (Weiers 1986d: 62ff).

It is assumed that some Mongolic-speaking groups may have lived near Lake Baykal in the second half of the first millennium AD. These are viewed by some as constituting part of the Buryat ancestors (Nimaev 2004: 25). However, in view of the fact that modern Buryat is an Eastern Mongolian language related to Khalkha-Mongolian and Southern Mongolian dialects, it is clear that the linguistic ancestors of the Buryats must have been in close contact with the other Mongolic tribes in the 13th and 14th centuries, the period of unification of the Mongolic languages under Chinggis Khan and his successors. The Western Buryats are said to represent direct descendants of the Turkic-speaking Kurykans who shifted to the Mongolic language after the migration of the Sakha ancestors to the north (Konstantinov [1975] 2003: 31, 36; Gogolev 1993: 58; Nimaev 2004: 20), while the Buryats as a whole are assumed to have assimilated a number of indigenous Evenk tribes both linguistically and ethnically (Buraev & Šagdarov 2004: 228f).

1.1.5 Potential contact of the Sakha ancestors with the indigenous populations

The Evenks and Ėvens appear to have been settled in Yakutia not much longer than the Sakha themselves, since it is claimed that they migrated to the north only in the 12th century. As highly nomadic hunters and reindeer-herders their lifestyle must have been very different from that of the immigrating cattle- and horse-breeders; however, since the latter depended on hunting and fishing as well as on the meat and milk from their livestock, there may well have been some contact along the rivers.

As to the Yukaghirs, it is not clear whether the immigrating Sakha would have come into contact with them on the middle Lena, or only after their expansion to the northeast. Although it is quite probable that Yukaghirs were initially settled over most of Yakutia, the immigration of the Tungusic-speaking ancestors of the Evenks and Ėvens, who relied on the same game and fish as the Yukaghirs, may well have pushed the latter to the northeast prior to the arrival of the Sakha.

From sections 1.1.1.2 and 1.1.4.2 it follows that there are three possible time periods during which the ancestors of the Sakha may have been in contact with Mongolic-speaking groups: an early period of contact might have taken place between an unknown Mongolic-speaking group and the Turkic-speaking Kurykans, the presumed Sakha ancestors, in the second half of the first millennium AD. However, given the fact that most of the Mongolic substance copies in Sakha appear to stem from a Middle Mongolian or Written Mongolian source of the 13th and 14th centuries, such an early period of contact seems not to have had much lexical impact on Sakha. A second time period may have been the 11th and 12th centuries, when

there was ongoing conflict between the Mongolic tribes; it is not unlikely that some tribes or clans broke away and fled to the area around Lake Baykal to evade this. Finally, the period of the Mongol Empire in the 13th and 14th century was far from peaceful as well; not only were neighbouring tribes and nations conquered, but Mongolic tribes that did not swear allegiance to Chinggis Khan or his successors were punished by military expeditions. So during this period, too, some clans or tribes unwilling to subjugate themselves may have fled to the north; to Lake Baykal but possibly even further north, if the Sakha legends have some connection to actual historical events.

From the Mongolic copies in Sakha it is clear that some contact must have taken place between Sakha and Mongolic-speaking tribes, and from the historic and current settlement of Sakha and Evenks and Ėvens in the same geographical territory, some contact with speakers of Northern Tungusic dialects or languages may well have taken place, too. Thus, prehistoric contact between groups speaking unrelated languages is known to have taken place. In the following section I provide an overview of theories concerning the linguistic results of contact between groups of people speaking different languages.

1.2 Language Contact

Although there were some early general theoretical studies of language contact (most notably Haugen 1950, 1953 and Weinreich 1953), it was the publication of Thomason & Kaufman's seminal monograph *Language Contact, Creolization, and Genetic Linguistics* in 1988[1] that led to a burgeoning of interest in this topic (cf. Ross 2003: 175). In recent years a number of linguists have presented their views on the mechanisms and factors involved in language contact and the possible outcomes (Thomason & Kaufman 1991; Johanson 1992, 1999; Aikhenvald 2003a, b; Ross 1996, 2001, 2003; Heine & Kuteva 2003, 2005, *inter alia*). Different terminologies abound, and although often the terminological differences hide merely shallow distinctions in actual theories, there are some divergent approaches to the matter at hand. This section aims at presenting an overview of current theories and approaches, with the ultimate goal of extracting the terminology and the approach that seem most promising for application in this study.

To facilitate the presentation of the different approaches to language contact, I will here briefly define the terms that I will use in the following discussion; for the reasons behind the choice of each of these terms see section 1.2.8. The transfer of linguistic elements from one language to another will be called copying, and the language from which an element is copied will be termed the model language, while the language doing the copying will be termed the recipient language. From a sociocultural point of view the language spoken within a community that may be emblematic of that community's identity will be called the ingroup language, while the language used for communication with other speech communities will be called the outgroup language. Copying can involve both the transfer of form-meaning units (e.g. morphemes or lexemes), which will be called substance copies, and the transfer of linguistic patterns, which will be called schematic copies. Finally, the large-scale restructuring of the recipient language under the influence of the model language will be called metatypy.

It should also be pointed out at this stage that throughout this thesis I may occasionally talk about 'language contact', or a 'change taking place in language A under influence of language B'. This is not to imply that I think that languages can change of their own accord, independently of any speakers. To me, it is of fundamental importance that languages change through the behaviour of their speakers, either because speakers of different languages are in contact and so have some knowledge of both (or more) of these languages, or because two or more

---

[1] This was reprinted as a paperback in 1991, and in the following I refer only to the paperback version.

languages may be in contact in one speaker's mind. 'Language contact' is only a shorthand expression for such complex psycholinguistic and sociolinguistic scenarios.

1.2.1 The languages in contact

Weinreich (1953: 30) proposes to make two terminological distinctions concerning the languages involved in contact: in cases where substance copies are made, he suggests distinguishing between the *source language* and the *recipient language*, while in cases of structural influence that involve the transfer of schematic copies he proposes to distinguish between the *model language* and the *replica language*. This terminology is taken up by Heine & Kuteva (2003: 531 and 2005: 2) who, in accordance with their focus on contact-induced grammaticalization (i.e. the transfer not of actual material, but of meaning extensions and grammaticalization pathways), adopt Weinreich's distinction between *model language* and *replica language*.

Winford (2005: 376f) bases his approach on that of Van Coetsem (1988) and adopts Van Coetsem's terminology, who follows Weinreich in distinguishing between a *source* or *donor language* (SL) and a *recipient language* (RL). In this framework, linguistic material is always transferred from the source language to the recipient language (Van Coetsem 2000: 51f), while the material being transferred need not be substance copies but can also involve schematic copies.

Johanson (1999: 40) makes a sociocultural distinction between the speaker's *primary code*, that is, the ingroup language (often his mother tongue), and the speaker's *secondary code* which is used for external communication. From a linguistic perspective he distinguishes the *model code*, from which features are copied, and the *basic code*, which does the copying. Ross (1996: 181) likewise makes a sociocultural distinction between a group's ingroup language, called *emblematic language* in his terminology, and the *intergroup language*; it is important to note that the emblematic language is not necessarily used more frequently than the intergroup language. In a later article (2001: 146), Ross changes his terminology, distinguishing between *ingroup lect* and *outgroup lect* in order to make his approach equally applicable to dialects and languages; in 2003 (182) he changes this terminology yet again to *primary lect* for the speaker's emblematic lect and *secondary lect* for the lect used for external communication [i.e. this terminology is very similar to that of Johanson (1999)]. Once again, it is important that some speakers may use their secondary lect more often than their primary lect (Ross 2003: 183).

From a purely sociocultural perspective, Croft (2003: 50) suggests the term *heritage society* (and *heritage language*) for the speaker's ethnically ancestral society and language, while *adoptive society* is the society the speakers are identifying with socially and linguistically. (It should be noted that Croft discusses the development of mixed languages, i.e. only a small subset of all kinds of language contact.) Thomason & Kaufman (1991) do not make any explicit terminological distinction between the languages involved in a contact situation; however, they coin the term *target language* (TL) for the language that a group of speakers is shifting to, and refer to the *source language* as the language that provides the copied material, i.e. the model language in my terminology (Thomason & Kaufman 1991: 39, 114). For the language that receives copies from another language (i.e. the recipient language in my terminology) as well as for the language from which a group of speakers is shifting they have no specific term, but simply refer to the 'native language' or the 'shifting speakers' language', e.g. "Borrowing is the incorporation of foreign features into a group's **native language** by speakers of that language…" (p. 37, emphasis mine); "Often, in fact, the TL adopts few words from the **shifting speakers' language**. […] If the speakers' goal is to give up their **native language**…" (p. 39, emphasis mine).

1.2.2 The types of contact

One of the main distinctions made in all accounts of language contact concerns the types of contact that are possible. These differ between a focus on the kinds of linguistic elements that are copied and a focus on the process of contact. Unfortunately, the terms chosen by authors focussing on the kinds of linguistic elements copied and by those with a focus on the process of contact are often the same (this holds most especially for the widely-used term 'borrowing'), blurring the differences between the approaches and leading to some confusion. I provide an overview over the major terminological differences in Table 1.1 at the end of section 1.2.2.2.

1.2.2.1 Approaches focussing on the type of copies that are transferred

Weinreich (1953: 1, 7) distinguishes between *borrowing* and *interference*, with borrowing involving the transfer of substance copies such as lexemes or morphemes, while interference involves the transfer not of actual formal elements,

but of schematic copies such as structural patterns and semantic meaning. Croft (2003: 51) similarly proposes the term *borrowing* for the introduction of what he calls 'substance linguemes', i.e. form-meaning units, as opposed to *convergence* to designate the introduction of what he calls schematic linguemes (linguistic elements made up of form alone or meaning alone). Heath (1978: 119) distinguishes *direct diffusion* involving the transfer of forms (copied phonemes, morphemes, or lexemes) and *indirect diffusion*, in which only structural patterns are copied: "… a process whereby one language rearranges its inherited words and morphemes under the influence of a foreign model, so that structural convergence results".

Aikhenvald (2003a: 3) emphasizes the need to distinguish between *diffusion of patterns* and *diffusion of form*, since not all linguistic communities are equally accepting of copied forms. Ross (2003: 189), too, points out that lexicon is often emblematic of a speaker's linguistic and ethnic identity and may therefore underlie stricter sociocultural constraints on contact influence than syntax. With respect to diffusion of pattern, Aikhenvald (2003a: 2) distinguishes two kinds of changes: *system-altering changes*, e.g. the introduction of a new category under the influence of a contact language, and *system-preserving changes*, e.g. the extension of already existing categories following the model of a contact language. New categories and new paradigms can be introduced through the reanalysis of existing categories and morphemes, through grammaticalization of new morphemes out of existing language material (Aikhenvald 2002: 60, cf. Harris & Campbell 1995: 50f, 89, 97), or through 'enhancement', "whereby certain marginal constructions come to be used with more frequency if they have an established correspondence in the source language" (Aikhenvald 2002: 238). It is such system-altering changes that can lead to the creation of structurally isomorphic languages in situations of language contact; and such structural isomorphism facilitates the direct copying of morphemes, since these can then fit into equivalent 'slots' in the recipient language (Aikhenvald 2002: 238).

1.2.2.2 Approaches focussing on the processes involved in language contact

Thomason & Kaufman (1991: 37ff), distinguish between *borrowing* and *interference through shift*. In contrast to the distinction made in similar or identical terms by other authors, which concerns the kind of copies that are transferred, in Thomason & Kaufman's approach the terminological distinction concerns the viability of the recipient language: in their terminology, 'borrowing' is the transfer of both substance and schematic copies into a recipient language that is maintained, while in 'interference through shift' both schematic and substance copies enter a

language that is the target of shift by a group speaking another language. That is, the main difference made by Thomason & Kaufman is whether a language is maintained (in which case they call all copies, whether substance or schematic, borrowing) or given up (in which case they talk about interference, either lexical interference or structural interference, cf. Thomason & Kaufman 1991: 40). In both kinds of contact, substance copies and schematic copies can be transferred, but Thomason & Kaufman claim that the order of transfer differs: in what they call borrowing, substance copies, especially lexemes, are introduced first, and schematic copies are made only later, while in what they call interference through shift, schematic copies are transferred first (phonological and syntactic copies first of all), followed by substance copies only at a later stage, if at all. In a later paper Thomason (2003: 692) points out that the term 'shift-induced interference' is misleading, since the phonological and syntactic results of such interference need not necessarily be the result of language shift; however, for lack of a "convenient and fully accurate term for what has been called shift-induced interference" and to avoid "proliferating terms" she proposes to continue using it (p. 692).

Winford (2005: 376f) follows Van Coetsem (1988, cited from Winford 2005; see also Van Coetsem 2000: 32, 53f) in making a functional distinction between the agents of the linguistic transfer; this approach distinguishes between *recipient-language agentivity* (which in this approach is called *borrowing*) and *source-language agentivity* (which in this approach is called *imposition*). The crucial element in this approach is that it is the bilingual speaker's linguistic dominance in one of her two languages that determines the agentivity: if a bilingual speaker adopts elements from her non-dominant source language into her dominant recipient language, 'borrowing' (*qua* Van Coetsem and Winford) has taken place, while if the bilingual speaker adopts elements from her dominant source language into her non-dominant recipient language, 'imposition' has taken place. In this framework, although 'borrowing' involves primarily lexical items, structural features can be borrowed as well; on the other hand, 'imposition' involves mainly phonological and structural elements, but the imposition of lexical items is possible, too. Thus, while the distinction between 'borrowing' (*qua* Van Coetsem and Winford) and 'imposition' seems to match Thomason & Kaufman's distinction between 'borrowing' and 'shift-induced interference' (as pointed out in Thomason 2003: 691), the focus in Van Coetsem's and Winford's distinction is not on the **social context** of the language contact (as in Thomason & Kaufman's approach, where the major distinction is between maintenance of the recipient language and shift), but on the **psycholinguistic context**, with a focus on linguistic dominance in one of the languages of a bilingual speaker.

In the extension of his theory, Van Coetsem (2000) adds a further type of language contact, which he calls *neutralization*. This occurs in the case of symmetrical bilinguals, i.e. when neither of the languages involved in the contact situation is the linguistically dominant one for a given speaker. In cases of neutralization, the outcome of the transfer is determined by the speakers themselves who can freely choose between the features of each of the languages depending on the saliency or frequency of the feature, on social prestige, or what is desirable from a perspective of self-identification. In these situations, "… any of the two languages of the bilingual can serve as RL [recipient language] or as SL [source language]." (Van Coetsem 2000: 42, 50, 85f).

In a similar vein to Haugen's (1950: 211) and Moravscik's (1978: 99, footnote 1) comments that the linguist's use of the term 'borrowing' differs radically from the everyday use of this word, Johanson (1992: 175; 1999: 39f) proposes the term *copying* to describe the transfer of elements between one language and another in order to avoid the metaphors inherent in the traditional terms borrowing, transfer, or interference:

> "In language contact nothing is really borrowed: the 'donor language' is not robbed of any element, and the 'recipient language' does not take over anything that would be identical to an element of the 'donor language'. The same danger is inherent in the term 'transfer'. We avoid the term 'interference' because of its oftentimes negative connotations." (Johanson 1992: 175, my translation[2]; cf. Stolz & Stolz 1996: 95)

Using similar terminology as Van Coetsem, Johanson (1999: 41f) distinguishes between *adoption*, which involves the insertion of a copy of material from the speaker's secondary code (the outgroup language) into his primary code (the ingroup language), and *imposition*, which is the insertion of a copy of material from the speaker's primary code into his secondary code. In Johanson's approach, 'imposition' does not necessarily entail code shift (Johanson 2006: 5). The difference between Johanson's approach and Van Coetsem's and Winford's is that Van Coetsem, and following him Winford, see differences in linguistic proficiency of the bilingual speaker (his 'dominance' in one language) as the major factor influencing the kind of transfer/copying, while Johanson (1992: 170ff; 1999: 41f)

---

[2] Original: "Beim Sprachkontakt wird nichts tatsächlich entlehnt: die „Gebersprache" wird keines Elements beraubt, und die „Nehmersprache" übernimmt nichts, was mit einem Element der „Gebersprache" identisch wäre. Dieselbe Gefahr ist mit dem Terminus „Transfer" verbunden. Den Terminus „Interferenz" vermeiden wir wegen seiner heute oft negativen Konnotationen."

sees sociopolitical dominance of languages as being the major factor[3]: in 'adoption' (*qua* Johanson), a sociopolitically dominated language copies elements from the sociopolitically dominating language, while in 'imposition' (*qua* Johanson) copies from a sociopolitically dominated language influence the sociopolitically dominating one. Both approaches agree that in 'adoption'/'borrowing' primarily lexical items are copied, while in 'imposition' it is mainly phonological and syntactic structural features that are copied. Furthermore, Johanson (1999: 41) makes a linguistic distinction between the types of material copied by referring to the copying of form-meaning units (i.e. substance copies) as *global copying* and to the copying of properties of language (i.e. schematic copies) as *selective copying*. Table 1.1 summarizes the differences in terminology discussed in the previous two sections.

Thus, Thomason & Kaufman, Van Coetsem (and following him, Winford), and Johanson appear superficially to mean the same things when they talk about 'borrowing'/'adoption' vs. 'interference'/'imposition'. All three approaches agree that in the first kind of language contact predominantly substance copies are transferred, while in the second kind of contact schematic copies are predominantly transferred, especially in the initial stages of the process. This superficial similarity in the approaches is further compounded by the overlap in terminology between Thomason & Kaufman and Van Coetsem, who both use the term 'borrowing', and between Van Coetsem and Johanson, who both use the term 'imposition'. However, there are actually fundamental differences between the approaches, since Thomason & Kaufman make a distinction between the maintenance of a language vs. shift to another language, while Van Coetsem focusses on the psycholinguistic issues involved in the contact process, and Johanson focusses on the sociopolitical issues. The terminological confusion is augmented by the fact that other authors use the term 'borrowing' to mean a transfer of substance copies as opposed to a transfer of schematic copies (see also Grant 2003: 251). Given this terminological mess, the term 'borrowing' should rather be avoided; and since both 'interference' and 'imposition' are used by at least two authors with different meanings, they should probably be avoided as well.

---

[3] Van Coetsem (2000: 57) does see social dominance as playing a role in situations of language contact, although not by actually having an impact on the transfer type, but rather by influencing the linguistic dominance of speakers.

Table 1.1 Summary of the terminology used in theories of language contact (in the first half the approaches with a focus on the kind of copies are summarized, in the second half the approaches with a focus on the process are summarized)

Term	Author	Meaning
borrowing	Weinreich	transfer of substance copies
borrowing	Croft	transfer of substance copies
direct diffusion	Heath	transfer of substance copies
global copying	Johanson	transfer of substance copies
interference	Weinreich	transfer of schematic copies
convergence	Croft	transfer of schematic copies
indirect diffusion	Heath	transfer of schematic copies
selective copying	Johanson	transfer of schematic copies
borrowing	Thomason & Kaufman	copies entering a language that is maintained
borrowing	Van Coetsem, also Winford	recipient-language agentivity (transfer of copies from bilingual speaker's non-dominant source language into dominant recipient language)
adoption	Johanson	introduction of material from outgroup language into ingroup language
interference through shift	Thomason & Kaufman	copies entering a language that is the target of shift by a group speaking another language
imposition	Van Coetsem, also Winford	source-language agentivity (transfer of copies from bilingual speaker's dominant source language into non-dominant recipient language)
imposition	Johanson	introduction of material from ingroup language into outgroup language

1.2.2.3 Metatypy

Ross (1996, 2001, 2003) points to the fact that often as a result of language contact you find large-scale morphosyntactic restructuring of the languages involved without concomitant lexical copying or phonological change; that is, the distinction proposed by Thomason & Kaufman between 'borrowing' and 'shift-induced interference' does not adequately describe the result of language contact. For the large-scale restructuring of languages in contact Ross proposes the term *metatypy* (Ross 1996: 182). What Ross designates as metatypy can be considered the result of long-term source language agentivity *qua* Van Coetsem and Winford – Ross stresses the fact that, at least in New Guinea, bilinguals frequently use their outgroup language

more often than their emblematic ingroup language: "Ironically, many speakers are more at home in the intergroup language than in their emblematic language: They use the intergroup language more often, and maintain their emblematic language principally as marker of their ethnicity and for (often limited) use within the village community." (Ross 1996: 181). Thus, to reformulate Ross' approach following Van Coetsem's terms, over a long period of bilingualism, source language agentivity can lead to the restructuring of the non-dominant recipient language on the model of the dominant source language, thus resulting in metatypy.

Although they do not discuss the theoretical implications of their data, Gumperz & Wilson (1971: 164f) find the same mechanism at play in the Indian village of Kupwar:

> "Speakers can validly maintain that they speak distinct languages corresponding to distinct ethnic groups. While language distinctions are maintained, actual messages show word-for-word or morph-for-morph translatability, and speakers can therefore switch from one code to another with a minimum of additional learning." (Gumperz & Wilson 1971: 164f)

Thurston (1987) argues that the same mechanisms have played a role in North-Western New Britain, where languages belonging to different subgroups of Austronesian, as well as one Non-Austronesian language, show very similar syntactic and semantic structures: "[…] in NWNB [North West New Britain] [it is] possible to translate word by word among languages that belong to three different branches of AN and a NAN isolate. In view of the extensive multilingualism and dual-lingualism in NWNB, the implication is that all of these languages share a single semantic and syntactic structure, differing only in the forms encoding items of their lexica." (Thurston 1987: 74). This approach is further elaborated by Ross (2001: 148ff), who suggests that the semantic organization of two languages undergoing metatypy is unified first before syntactic restructuring sets in; Aikhenvald (2002: 228ff) also demonstrates the semantic convergence of Tariana lexicon to East Tucanoan patterns.

It is widely acknowledged that such restructuring in bilinguals answers a need to lighten the cognitive burden inherent in the use of two different languages (e.g. Haase 1992: 167; Ross 1996: 204; Matras 1998: 291; Johanson 1999: 53); this was pointed out initially by Weinreich (1953: 7f), who suggests that *interlingual identification* is the process that drives schematic copying. In such interlingual identification, bilingual speakers identify a structural element in one language with a structural element in the other language and start using the one in lieu of the other.

Heine & Kuteva (2003, 2005) focus on one particular type of contact-induced change, namely contact-induced grammaticalization. Within this narrow framework,

they suggest a distinction between *ordinary contact-induced grammaticalization* and *replica grammaticalization* (Heine & Kuteva 2003: 533, 539; 2005: 81, 92). In ordinary contact-induced grammaticalization, speakers of the recipient language perceive a structure in the model language which they then copy, making use of their own linguistic material; thus, in these cases the contact situation triggers a grammaticalization process which may not necessarily have taken place without the initial contact. In replica grammaticalization, speakers of the recipient language copy not only the pattern of the model language but do so following the same path of grammaticalization as that followed by the model language (at least, as far as linguistically naïve speakers can be aware of such matters). As Heine & Kuteva themselves point out (2003: 555ff, 2005: 100ff), what they call contact-induced grammaticalization, especially replica grammaticalization, is very similar, and often identical to, what has been called polysemy copying or calquing. This view is also argued for by Johanson (in print: 8ff), who maintains that it is not the process of grammaticalization of the model language that is copied, but only the endpoint of the process, since "diachronic processes are not copiable" (Johanson in print: 9).

1.2.3 The role of linguistic structure vs. sociocultural setting in language contact

While Matras (2000) emphasizes the role of structural and functional properties of linguistic elements in language contact ("[…] elements which show structural autonomy and referential stability are more likely to be affected by contact than those which display stronger structural dependency and referential vagueness or abstractness." Matras 2000: 567), Thomason & Kaufman stress the overwhelming role of the sociocultural situation: "[…] it is the social context, not the structure of the languages involved, that determines the direction and the degree of interference." (Thomason & Kaufman 1991: 19). In the recent literature, however, a consensus seems to have been reached that while the sociocultural setting of the contact situation, and especially the intensity and duration of contact, is of primary importance in determining the linguistic outcome of contact, purely linguistic factors such as the structural divergence or similarity of the languages in contact play a role as well (Harris & Campbell 1995: 124f, 131; Johanson 1999: 50, 60; 2002: 306; 2006: 25; Ross 2001: 156, 2003: 176; Aikhenvald 1999: 411). For example, Aikhenvald (2002: 241) suggests that the structural difference between Portuguese and Tariana may have been one of the factors limiting the transfer of schematic copies from the former into the latter, together with the relatively short duration of the contact situation and the complementary distribution of use (diglossia) of the individual languages.

However, Heine & Kuteva (2005: 13) claim that they do not find any correlation between the type of sociolinguistic setting (e.g. sociocultural dominance of one of the languages) and the kind and degree of contact-induced grammaticalization, although they agree that duration and intensity of contact play a role. Stolz & Stolz (1996: 110f) on the other hand stress the importance of the contact situation, especially the degree of prestige of the model language; thus, speakers of American Indian languages in Mesoamerica have copied a large number of discourse particles and conjunctions from Spanish in order to 'exploit the prestige of Spanish'. Matras (1998: 309, 321), however, argues that the frequent copying of such discourse particles should not be ascribed to the prestige of the source language, but rather to the fact that they can be perceived as 'gesturelike devices' and so are easily detached from the content of the utterance.

Johanson suggests that both the sociocultural setting as well as structural features influence the outcome of language contact: "'Attractive' properties may be copied even in the absence of strong social pressure, but the presence of such pressure can ultimately promote copying even of 'unattractive' properties." (Johanson 2002: 310). 'Attractive' properties are such that make them easier to learn and understand, while "less attractive elements are those which have empirically proved to be copied less readily[4]" (Johanson 2002: 309). Winford (2005: 377) emphasizes the importance of the psycholinguistic setting of a bilingual speaker's unequal proficiency in one of his languages over the sociocultural dominance of one language over the other.

While it is often claimed that copying of form-meaning units (especially free lexemes) is easiest (e.g. Weinreich 1953: 56; Gumperz & Wilson 1971: 161; Moravscik 1978: 110; Matras 2000: 567), Ross (2003: 189) and Aikhenvald (2003a:3) point out that in cases where the language is emblematic of a group's identity, the lexicon (as the most salient part of the language for naïve speakers) might be under stronger sociocultural constraints than structural features. Interestingly, in their discussion of the linguistic convergence in the Indian village of Kupwar, Gumperz & Wilson (1971: 161f) find that although copying of lexical and functional items was widespread, cases of copying of suffixes met with disapproval of the speakers. They interpret this as an indication that "such paradigmatically structured inflectional morphs seem to be at the core of the native speakers perception of what constitute 'different languages'" (Gumperz & Wilson 1971: 161f).

---

[4] There appears to be some circularity of argumentation here, in that features that have not been found to be frequently copied are classified as 'unattractive' precisely because they are not copied frequently.

One factor facilitating contact-induced change is whether the feature in question is present already in the recipient language, albeit as a marginal, low-frequency variant. Through contact, such low-frequency variants may rise to higher frequency and eventually even attain the status of the standard form, if they correspond to features in the model language. This is termed *frequential copying* by Johanson (1999: 52; 2002: 306) and *enhancement* by Aikhenvald (2002: 238), while Heine & Kuteva (2005: 50) talk about minor use patterns becoming major use patterns through contact:

> "A widely observable process triggered by language contact concerns infrequently occurring, minor use patterns that are activated because there is a model provided by another language. […] under the influence of the other language they come to be used more frequently and their function tends to be desemanticized – with the effect that they may turn into more widely used major use patterns. This is how new word-order structures can arise, …"
> (Heine & Kuteva 2005: 50)

Conversely, as pointed out by Johanson (in print: 14), frequential copying does not only increase the use of a formerly marginal structure, but it can also decrease the use of a previously common alternative pattern under the influence of the model language. For example, Dutch speakers in Australia are using the definite article *het* less and less, making more use of the article *de*, which is similar to the English definite article *the* (Clyne 2003: 22, 31, cited from Johanson in print: 14).

The amount of time necessary to lead to contact-induced changes is unclear; Aikhenvald (1999: 390) estimates that in the contact situation documented by her in the Vaupés area, Tariana speakers have been in contact with speakers of Tucanoan languages for approximately 400 years. A similar estimate is given for the duration of contact in the oft-cited case of Kupwar (Gumperz & Wilson 1971: 153). On the other hand, in the case of Greek spoken in some regions of Anatolia, the contact of Greek speakers with speakers of Turkish goes back nearly one millennium (Winford 2005: 402). In the Vaupés the strict enforcement of 'linguistic exogamy' (Aikhenvald 1999: 388ff), which leads to widespread multilingualism, clearly plays a role in the degree of contact-induced changes undergone by the Tariana language. Such extensive intermarriage between ethnolinguistic groups has also led to strong influence on genealogically unrelated, neighbouring languages in Arnhem Land, Australia: these have undergone both structural influence ('indirect diffusion' in Heath's terms) as well as copying morphemes and a large number of lexical items (approximately 50% of the lexicon are shared between Ngandi and Ritharngu; Heath 1978, 1981).

1.2.4 The role of social networks in language contact

In a 1985 paper, Milroy & Milroy argue that the social network structure of language communities influences the spread of linguistic innovations. Based on work by Granovetter (1973) and Rogers & Shoemaker (1971; both cited from Milroy & Milroy 1985) they propose that it is weak rather than strong ties between groups that enable diffusion of changes. Strong ties are those in which individuals are emotionally and intimately involved, in which they provide each other with mutual assistance, and on which a large amount of time is spent. Weak ties, on the other hand, are less time-consuming and therefore more numerous, so that more individuals can be reached through weak ties. (Milroy & Milroy 1985: 364 compare the distinction between strong and weak ties to that between friends and acquaintances.) Furthermore, information or innovations passed on through a network of weak ties will be novel at each step, while information or innovations passed on within a network of strong ties will tend not to be novel, since in such a network a large number of individuals have ties with each other, so that the same information will reach a given individual from many associates. Strong ties are found mainly within small groups, while the ties linking different groups are weak ones (Milroy & Milroy 1985: 364). Furthermore, small groups characterized by strong ties are expected to be conservative and not susceptible to outside influences, because the constant contact between members of the group reinforces group norms.

As shown by empirical work by Rogers & Shoemaker (1971, cited from Milroy & Milroy 1985), innovators of cultural, technological, and linguistic change are often marginal members of a group with a large number of weak ties to other groups; these changes are in turn adopted by so-called 'early adopters' who are central members with strong ties within the group and who often provide a model for other non-innovators within that group (Milroy & Milroy 1985: 367). The basic tenet of this proposal is that maintaining strong ties in a social network is a time-consuming business, so that individuals with strong ties will have only few ties. On the other hand, individuals with weak ties will be able to maintain far more of these, since they are not as time-consuming to uphold. Thus, individuals with numerous weak ties will have more opportunity of picking up variant behaviour or speech; therefore, it is precisely the weak ties between groups that can serve as conduits for change (Milroy & Milroy 1985: 365f).

The difficulty with this model is to explain why the 'early adopters', who are central members of the group who conform to group norms, should adopt an innovation from marginal 'innovators'. However, Ross (1997: 231) provides a good explanation for this by pointing out that the way by which innovations may spread through a speech community is an 'invisible hand process'. This is a process which

is determined partly by copying what others do (as happens, for instance, when many individuals take the same shortcut across a patch of lawn, thereby (unintentionally) creating a path), partly by individuals having the same intentions (as happens, for example, when several people stop to watch an accident and so form a circle around the victim without anyone directing this action). Thus, if several marginal 'innovators' adopt a novel form of speech from a neighbouring group, the 'early adopters' may come to copy it because repeated use of the form has made it more acceptable.

Based on research by Trudgill (1986 cited from Ross 1997: 233ff), Ross proposes that one factor that determines the spread of a feature from one community to another is demography: if community A is more numerous than B, then it is more probable that most speakers of B will have direct contact with speakers of A than the other way round, and it is therefore more probable that a feature of A will be copied into B than vice versa. A second factor influencing the spread of features, especially of features that are emblematic of particular groups, is the prestige of that group. Thus, a linguistic feature characteristic of a prestigious group will be copied more readily (as happens, for instance, when emblematic features of the speech of the capital city are copied, such as the uvular /r/ originally characteristic of Parisian French). Milroy & Milroy (1985: 368) also stress the two factors of numeracy and prestige in the spread of linguistic innovations: The 'early adopters' will only adopt an innovation in technology, culture, or language if it has been taken over by a large number of 'innovators', and if the innovation is perceived as being prestigious: "[…] we suggest that persons central to the network would find direct innovation a risky business; but adopting an innovation which is already widespread on the edges of the group is much less risky." (Milroy & Milroy 1985: 368).

The Milroys find historical support for their theory in the comparison of Icelandic and English (Milroy & Milroy 1985: 375ff), suggesting that one of the reasons why Icelandic is so conservative as compared to English is that early Icelandic society was characterized by a very cohesive social network with an emphasis on strong ties between individuals, notwithstanding the very fragmented pattern of settlement with large geographical distances between individual locations. This cohesive social network structure enabled a maintenance of the language norms even in the absence of frequent contact. In England, on the other hand, there were disruptions of society through incursions of foreign peoples, leading to a disruption of strong ties; furthermore, the importance of London as a centre of economic and political power, and thus a magnet for immigration, meant that the society was a lot more mobile, again leading to the formation of weak social ties rather than strong

ones. All this, it is argued, led to changes in English taking place at a more rapid pace than in Icelandic:

> "[…] we have tried to show as explicitly as possible that innovations are normally transmitted from one group to another by persons who have weak ties with both groups. Further, at the macro-level, it is suggested that in situations of mobility or social instability, where the proportion of weak links in a community is consequently high, linguistic change is likely to be rapid. Social groups who contract many weak ties […] are likely to be closely implicated in the large scale diffusion of linguistic innovations." (Milroy & Milroy 1985: 380)

Based on dialect studies in Europe Andersen (1988: 71ff) proposes a two-way distinction of open vs. closed (or central vs. peripheral) and exocentric vs. endocentric speech communities. The distinction between open and closed communities refers to the density of the communicative networks between the community in question and other speech communities: an open community is characterized by a large number of ties with the outside world, while a closed community forms very few ties with other communities. The distinction between exocentric and endocentric communities refers to the speakers' attitudes, to the extent to which they accept linguistic usages of surrounding communities vs. the extent to which they adhere to their own norms. The combination of these features leads to different expectations concerning the acceptance of outside influence:

> "[…] one can expect exocentric closed dialects to accept diffused innovations just like exocentric open dialects, but at a rate which is slower in proportion to the lower density of their inter-dialectal communicative networks. Endocentric open dialects may retain their individuality in the face of relatively extensive exposure to other speech forms whether they form relic areas […] or they represent the dominant norms which are diffused from focal areas. It may be primarily an attitudinal shift from endocentric to exocentric which changes the course of development of a local dialect when it becomes part of a wider socio-spatial grouping and not just the opening up of new avenues of interdialectal communication." (Andersen 1988: 74f).

1.2.5 The individual in language contact

Oksaar (1999: 6) argues that the locus of language change is the multilingual individual: "The bridge between languages, dialects, sociolects is the multilingual individual, being thus the mediator of language contact and also of language change." Based on empirical research in bilingual individuals in different countries, she proposes that such multilingual individuals do not have only two (or more) separate languages/lects, but also an intermediate lect LX, which consists of items

from each of the individual languages, but is characterized by its own norms of use (Oksaar 1999: 9). This LX may thus be the locus where the interlingual identification necessary for metatypy takes place. This is very similar to Myers-Scotton's view of language contact: "Some linguists like to say that to speak of 'language contact' is erroneous, because it is the speakers who are in contact, not the languages. […] what is significant to the structural linguist is that the two languages abut each other. That is, the languages are in contact in the sense they are adjacent in their speakers' mental lexicon and can impinge on each other in production." (Myers-Scotton 2002: 5).

Similar to Oksaar (1999) and Myers-Scotton (2002), Enfield (2003) firmly bases all linguistic processes relevant to language contact in the individual (2003: 3ff). In this approach, language contact takes place via interacting individuals, and individuals' personalities play a role in the diffusion of contact phenomena: reclusive individuals who do not interact with many others will not greatly affect the spread of an innovation, whereas outgoing individuals with a lot of social connections may well be the agents of spread of innovations, be these copies or language-internal developments (Enfield 2003: 11ff). Given this focus on individuals rather than on languages, Enfield suggests that the traditionally stipulated difference between inheritance, copying (in his approach, all kinds of contact-induced changes), and internal innovation are qualitatively much smaller than usually claimed. This approach is very similar to that of Milroy (1997), who also stresses the use of language in social interactions between individuals, and who similarly sees no qualitative difference between internal sound change and copying (Milroy 1997: 316f).

1.2.6 Correlation between the social setting and the kind of contact

There have been numerous case studies of language contact in different parts of the world over the past few decades. In some instances we find long-term contact between speakers of different languages leading to great structural changes, to the extent of achieving a 'morpheme-by-morpheme intertranslatability' between the languages, without concomitant copying of actual substance (form-meaning units), or with only very little copying of substance. Thus in Kupwar in India, where most men residing in the village are able to speak more than one of the languages spoken there, but where each language is emblematic of the social and religious group that speaks it, the languages in contact have undergone nearly complete syntactic and morphological convergence (i.e. metatypy) while retaining their individual lexemes and morphemes: "The sentences in this example are lexically distinct in almost

every respect, yet they have identical grammatical categories and identical constituent structures … It is possible to translate one sentence into the other by simple morph for morph substitution." (Gumperz & Wilson 1971: 154f). However, although bound morphemes, especially inflectional morphemes, are very rarely copied in Kupwar, lexical items, including function words like conjunctions and post-positions, do get copied. Insertion of foreign inflectional suffixes into speech is, however, considered wrong, leading Gumperz & Wilson to conclude that "… wherever social norms favor the maintenance of linguistic markers of ethnic identity, and where there are no absolute barriers to borrowing of lexicon and syntax, these morphophonemic features take on the social function of marking the separateness of two language varieties." (Gumperz & Wilson 171: 161f).

On Karkar Island, however, Ross (1996, 2001, 2003) finds extensive convergence of the semantic and morphosyntactic structures of the languages in contact without concomitant lexical copying; this is similar to the Vaupés river linguistic area described by Aikhenvald (1996, 1999, 2002, 2003a, b). In both of these cases, language is perceived as emblematic of an individual's ethnic identity, and since lexemes are the most salient parts of a language for the native speakers, copying of lexemes is avoided (Ross 2003: 189; Aikhenvald 2003a: 3).

In Arnhem Land, on the other hand, Heath (1978) finds widespread morphosyntactic convergence, i.e. schematic copying ('indirect diffusion' in Heath's terms), copying of bound morphemes ('direct diffusion' in Heath's terms), and a large amount of lexical copying, especially between Ngandi and Ritharngu, two genealogically unrelated languages. These share at least 20% of lexical items in most domains, and in some domains, such as names for trees and shrubs, or terms for human age and sex groupings, the sharing concerns over 50% of all the lexical items (Heath 1981: 349). Heath explains this by the fact that in Arnhem Land language does not serve as a strong marker of social or ethnic identity; thus there is no taboo against the copying of actual forms. At the same time, although speakers of different languages congregated for joint celebrations at certain times of the year, for most of the time a social unit such as a clan or smaller group would have consisted of speakers of one dominant language, so that the amount of daily code-switching necessary would have been a lot less than that found in Kupwar, where men have to switch from language to language on a daily basis (Heath 1978: 142).

"While in the South Asian case direct morphemic diffusion was rare because of pressures to keep the languages, [*sic*] distinct in Arnhem Land there are abundant instances of such diffusion. Whereas in the South Asian case indirect morphosyntactic diffusion has been maximal, in Arnhem Land it has been fairly substantial but far from complete, and we do not find one-to-one morphemic intertranslatability or even a strong tendency in this direction:

> this is presumably due to the lesser extent of code-switching, especially on a day-to-day basis or within single conversations." (Heath 1978: 142f).

Another factor leading to the high rate of lexical copying may have been the extensive intermarriage between ethnic groups in the region, especially between Ngandi and Ritharngu. This led to bilingual families and thus facilitated copying even of core vocabulary (Heath 1981: 359, 365). However, in the Vaupés area linguistic exogamy used to be the norm, but this did not lead to the copying of lexical items (Aikhenvald 1996: 77f, 104; 2002: 21ff, 213ff). It thus becomes clear that one of the major factors influencing the outcome of language contact is the attitude of the speakers. As pointed out by Heath himself, the Arnhem Land contact situation is unusual, precisely because of its lack of social factors influencing the diffusion of linguistic features (Heath 1978: 143).

Based on her work in northwestern Amazonia, Aikhenvald (2003b: 2f) proposes a 'typology of language contact': when several languages are in contact without any one of them being the socioculturally dominant one, the typological patterns of the languages are expected to be enriched. In a situation where only two languages are in egalitarian contact, without either of them dominating the other, a 'mutual adjustment' of the languages with structural levelling is expected. When two languages are in contact, of which one is sociopolitically dominant, then the subordinate language is expected to undergo rapid change with a marked loss of structural patterns.

In a very elaborate model Ross (2003) distinguishes between different results of contact depending on the sociocultural constitution of the communities in contact, following Andersen's (1988) typology of sociospatial and attitudinal differences in speech communities. The theoretical underpinning of the diagnostic 'tools' proposed by Ross (2003) is the social network model presented in an earlier paper (Ross 1997: 213ff): "[…] the social network model, is founded on a transparent fact that the species evolution metaphor ignores – that languages have speakers, and that language resides in their minds. Speakers use language to communicate with each other, and the model treats speakers as nodes in a social network, such that each speaker is connected with other speakers by social (and therefore communication) links." A speech community is defined by Ross as a social entity which is structured in a social network, and as outlined by Ross (1997, 2003) linguistic events can be used to reconstruct prehistoric events in the life of a speech community. Thus, members of a closed and tightknit group (corresponding to Andersen's closed and endocentric community) might attempt to make their lect harder for outsiders to understand and learn, resulting in phonological and morphological complexity (Ross 1996: 183; 2003: 181f); this has been termed *esoterogeny* by Thurston (1987: 38,

58ff). As Ross (2003: 182) points out, it is not clear whether esoterogeny is just the result of internal innovations which can proliferate in small closed communities, or whether it is the result of a reaction to contact, an attempt by the speakers of a language or dialect to enhance the emblematicity of linguistic features that make their lect different from that of outsiders and harder for the outsiders to learn. Interestingly, Kulick (1992: 2f) provides some anecdotal evidence of conscious manipulation of language structures in Papua New Guinea with the purpose of making the particular dialect more different from its neighbours, suggesting that at least occasionally esoterogeny can occur as a result of contact.

Metatypy is expected as the result of contact between an open and tightknit group (i.e. an open and endocentric community in Andersen's terms) and others, that is, in a speech community with many communicative ties with other groups that nevertheless values its ingroup language for its emblematic function. As Ross (2003: 191) argues, a community that is open, looseknit (exocentric) and polylectal is on the verge of losing its identity as a separate community, since the communicative ties within the group may be on a level as those with other groups. Such a community may well shift to the more frequently used outgroup language, occasionally resulting in phonological copies entering the language they shifted to. As to lexical copying, according to Ross' theory this is expected not under language contact, but under culture contact, since such copying can take place without widespread bilingualism (Ross 1996: 209f; Ross 2003: 193).

1.2.7 Achievements in the field of language contact studies

There have been two important lines of progress since the publication of Thomason & Kaufman's widely-read and widely-cited monograph – although one of them appears to have been an independent proposal published in the same year as Thomason & Kaufman (1991) that has not yet received much attention (Van Coetsem 1988 as cited in Winford 2005). What restricts the approach of Thomason & Kaufman (1991) (continued by Thomason 2003) is the classification of all situations of language contact as either language maintenance (involving ever larger degrees of substance and schematic copying) or language shift (involving what they call 'substratum interference'). As has been shown by Gumperz & Wilson (1971), Heath (1978), Aikhenvald (1999, 2003a) and Ross (1996, 2001, 2003), amongst others, linguistic communities are often stably multilingual, with one language (or dialect) serving as the emblematic, identity-giving language and the other(s) serving the needs of communication with neighbouring communities. Both Aikhenvald and Ross clearly show that in such cases the result of contact is not substance copying,

and not necessarily shift (although the Tariana studied by Aikhenvald have recently begun to shift to Tucano), but what Ross terms metatypy. This recognition of a third type of language contact influence is, in my opinion, of fundamental importance, since stable multilingualism is surely widespread in many areas of the world. In addition, Ross (2001, 2003, following Thurston 1987) proposes a fourth type of contact-induced change, namely the complication of the ingroup language in order to make it harder to understand for outsiders ('esoterogeny'); this, however, seems to be of a fundamentally different nature than the other three kinds[5].

The second fundamental insight is the proposal by Van Coetsem (1988), taken up by Winford (2005), that the underlying mechanism of contact-induced change is the relative proficiency of bilingual speakers in one or the other language. This is applicable to all kinds of contact situations, both stable bi- or multilingualism as described by Aikhenvald (2002, *inter alia*) and Ross (1996, 2001, 2003), and sociopolitically biased contact situations such as are the focus of Johanson's work (1992, 1999, 2002: 289). This distinction avoids the issue raised by Thomason (2003: 692) that imperfect learning is involved in 'shift-induced interference', because it assumes the presence of bilingual speakers; in this approach the contact-induced changes are a function of the extent of use of each of the languages.

A further fruitful development in the past 50 years since the publication of Weinreich's monograph (1953) is the paradigm shift from viewing language as a system (Weinreich 1953) to languages as sociocultural entities (Thomason & Kaufman 1991) to languages existing in the minds of speakers (Ross 2001, 2003, Heine & Kuteva 2005). This latter perspective allows the introduction into theories of language contact of psycho- and sociolinguistic insights into language processing (Levelt 1992; Oksaar 1999; cf. Ross 2001: 148) and fine-scaled distinctions of linguistic communities based on their network structure (Grace 1996: 172ff; Andersen 1988; cf. Ross 1997, 2003; Croft 2003) or their self-identification (Le Page & Tabouret-Keller 1985). The most extensively individualistic approach is that suggested by Enfield (2003).

---

[5] It is tempting to speculate in this context that the lexico-semantic divergence of Dolgan with respect to Sakha (Ubrjatova 1966) is due not to linguistic accident alone, but to a process of esoterogeny, with the speakers of Dolgan attempting to delimit their language from the closely-related Sakha language, concomitant with the process of new ethnic identification. However, until the degree of divergence between Dolgan and Sakha has been verified with actual data, this suggestion must remain purely speculative.

1.2.8 Terminology and approach to be followed in this study

Although Thomason (2003: 692) justifiably proposes to rather retain a somewhat misleading term than contribute to a 'proliferation of terms', Johanson (1992: 175) is correct in pointing out that infelicitous metaphors can unduly colour one's perspective of things. Furthermore, as discussed above (sections 1.2.1-1.2.6), frequently the same terms are used with different meanings by different authors; this holds especially true for the term 'borrowing'. The use of these terms therefore carries the potential of serious confusion, since it is unclear which of the meanings is intended; for this reason, I will avoid such terms, even though they may have a fairly long tradition of use. I here propose not to follow any one author in their entire terminology, but rather to 'pick and mix', choosing those terms that seem to me to be best suited to the study of language contact in general and this study in particular.

1.2.8.1 The languages in contact

Of the terms proposed as labels for the languages in contact we have first of all Ross' proposal (2001: 146) to subsume both languages and dialects under the term 'lect', while Johanson (1992, 1999, 2002) uses the general term 'code'. Although the broad term 'lect' to avoid making an unnecessary distinction between dialects in contact and languages in contact is surely a sensible choice for broad comparative studies of different contact situations, given the focus of the present study on contact between different languages, I will continue using the more familiar term language.

Furthermore, there exist on the one hand proposals that focus on the linguistic role played by the languages in contact: a) replica language vs. model language (Weinreich 1953; Heine & Kuteva 2005), b) basic code vs. model code (Johanson 1999, 2002), and c) recipient language vs. source language (Weinreich 1953; Winford 2005), while other proposals focus on the sociolinguistic situation of the contact: 1) emblematic language (later: ingroup language) vs. outgroup language (Ross 1996, 2001), 2) primary code/lect vs. secondary code/lect (Johanson 1999, Ross 2003), and 3) heritage society (and concomitantly, language) vs. adoptive society (Croft 2003). The use of separate terms to designate the languages involved in contact situations from a linguistic and from a sociocultural perspective is surely fruitful – if enough is known about the sociocultural background of the contact situation to be able to make such distinctions. (It is here assumed that given some knowledge of the state of a certain feature not only in the proposed contact languages, but also in their relatives, an assignment of languages to the linguistic

roles of 'model' and 'recipient' will most often be possible, cf. Heine & Kuteva (2005: 33). If the analysis of one specific language should show up changes in this language relative to its sister languages, and if these changes can be shown to be due to contact, then this language is by definition the recipient language, cf. section 1.4.2). Given Johanson's correct admonishment that in cases of language contact no material actually leaves the 'source' or 'donor' language, the term 'model language' is clearly preferable to 'source language'. As 'replica language' conveys to me the impression that the language is a wholesale replica of the model, I prefer the term 'recipient language' (I here assume that a language can **receive a copy** from the model language, not the original item). To distinguish the two languages from a sociocultural point of view I prefer 'ingroup language' vs. 'outgroup language' over 'primary' and 'secondary lect/code', since the latter terms convey the impression that the primary lect or code is used more frequently than the secondary lect/code – an impression intended by neither Ross (2003) nor Johanson (1999).

1.2.8.2 The processes involved in language contact

As to the process involved in language contact situations, here I propose to follow Johanson's terminology of 'copying' (Johanson 1992, 1999), making a distinction however not between 'global' and 'selective copying' (terms that to me are not intuitively comprehensible), but rather, following Croft (2003), making a distinction between 'substance copies' (i.e. copied form-meaning units such as lexemes or morphemes) and 'schematic copies' (e.g. the copying of form alone, extensions of meaning of specific categories, or the development of previously non-existent categories, based on a model language). Within schematic copies it might be useful to distinguish between system-preserving and different kinds of system-altering copies (Aikhenvald 2003a: 2).

Although I consider the psycholinguistic approach of Van Coetsem (1988, 2000) valuable, with its focus on the linguistic dominance of bilingual speakers, I will restrict myself to referring to 'model-language agentivity' and 'recipient-language agentivity', avoiding the cover terms proposed by Van Coetsem ('borrowing' and 'imposition') for the reasons discussed in section 1.2.2.2. Following Van Coetsem and Winford (2005) from a functional perspective, recipient-language agentivity is the process that takes place when recipient-language dominant bilinguals import elements (predominantly substance copies) from the model language into the recipient language. Model-language agentivity is the process that takes place when model-language dominant bilinguals introduce elements from the model language into the recipient language; in this case, these are

very often schematic copies. Large-scale restructuring of the recipient language in stable bilingual settings will be designated 'metatypy', following Ross (1996, 2001, 2003).

The process involved in schematic copying is one of 'interlingual identification' (Weinreich 1953: 7f; Johanson 1999: 53; Ross 2001: 148ff), where speakers of the recipient language identify certain structural elements of the model language as being equivalent to elements in their language and copy them to make the languages structurally more similar; this facilitates ease of production and/or perception in bilingual situations. Substance copies are often made from elements that are not present in that form in the language, i.e. they fill a gap; however, in heavy bilingualism it may also be that substance elements are used interchangeably and that then one gets replaced by the other. Schematic copies, too, can lead to the filling of a 'structural gap' – although whether this is a causal factor in the copying process is still unclear (cf. Harris & Campbell: 128ff).

1.2.8.3 Summary of chosen terminology

From a sociocultural perspective we can distinguish between the ingroup language and the outgroup language, while from a linguistic perspective we can distinguish two processes: 1) recipient-language agentivity (recipient-language dominant bilinguals introducing primarily substance copies into the recipient language), and 2) model-language agentivity (model-language dominant bilinguals introducing mainly schematic copies into the recipient language). Model-language agentivity can subsume system-altering and system-preserving copies. However, although in recipient-language agentivity mainly substance copies are introduced into the recipient language, schematic copies can be introduced as well; likewise, although in model-language agentivity it is primarily schematic copies that are inserted into the recipient language, this does not exclude the occasional transfer of substance copies.

1.3 Previous studies concerning language contact in Sakha

Given the fact that the Sakha are are known to have immigrated into the area they inhabit nowadays from a more southerly area of settlement, and that they are now surrounded by speakers of very different languages, it is not surprising that this is not the first study dealing with the effect language contact may possibly have had on the Sakha language. However, most of the previous work has focussed on the Sakha lexicon and the impact substance copies from Mongolic and Tungusic languages have had on this.

As early as the 19th century, the first linguistic study of the Sakha language found evidence of a large amount of lexical copies from Mongolic. Thus, in the introduction to his Sakha grammar, Böhtlingk ([1851] 1964: XXIX) states that Sakha can definitely be classified as a member of the Turkic language family, albeit a very divergent one. He also points out that the large number of lexical and morphological copies from Mongolic support the assumption that the Sakha and Buryats lived in intimate contact ("in inniger Verbindung") for some time (p. XXXVII). Although Böhtlingk provides a brief list of lexical copies from Mongolic to illustrate how these are phonologically integrated into the Sakha system of vowel harmony (p. 120), and throughout the grammar compares the Sakha roots and suffixes with Tatar and Mongolian forms, he does not discuss the issue of language contact in any more detail. In another early study, Radloff (1908) finds that of 1748 Sakha lexical roots, 32.5% are of Turkic and 25.9% of Mongolic origin, while he is unable to trace the origin of 41.6%. However, he recognizes Mongolic suffixes in a number of these, and therefore suggests that they probably have a Mongolic source, too (Radloff 1908: 2). After a brief survey of the Sakha grammar, Radloff comes to the conclusion that Sakha was initially a 'mixed language' that was mongolicized and, at an even later stage, turkicized (p. 51).

One of the first serious and notable investigations of the impact of language contact on Sakha is Kałużyński's monograph *Mongolische Elemente in der jakutischen Sprache* published in 1962. Here, Kałużyński provides a detailed analysis of the substance copies from Mongolic languages found in the Sakha-Russian dictionary compiled by Pekarskij ([1907-1930] 1958-1959). He refutes Radloff's assumption of Sakha being a mongolicized language that was turkicized only later, by showing that the copies from Mongolic entered the language later than the inherited Turkic elements (p. 8). Kałużyński deals exclusively with substance copies, but he does mention one syntactic copy from Mongolic as well, namely the use of the numeral 'two' to conjoin noun phrases, e.g. *aγa iỹe ikki* [father mother two] 'mother and father' (p. 119). Kałużyński comes to the conclusion that the bulk of the Mongolic copies in Sakha were adopted during the Mongol Empire and the

immediately subsequent period, between the 12th/13th and the 15th/16th centuries (p. 119). Judging from the nature of the copies, he concludes that the Sakha must have been part of the Mongol Empire, and that they were socially and politically subordinate to the Mongols (p. 120). Finally, as it is impossible to trace all substance copies in Sakha to a single Mongolic language, he concludes that the Mongolic model language either does not exist anymore nowadays, or that the language contact took place over such an extended period of time that speakers of Sakha were in contact with speakers of several different Mongolic dialects. One of these may well have been an older form of Buryat (p. 126). Kałużyński continued to conduct etymological studies of Sakha until the mid-1980s, most of which are compiled in the collection of his writings on Sakha, *IACUTICA*, published in 1995. One of these is his very useful presentation of some Tungusic lexical copies in Sakha (Kałużyński [1982] 1995: 225-232).

Other studies dealing with contact influence in Sakha are Antonov (1971), Romanova, Myreeva & Baraškov (1975), Rassadin (1980), and Popov (1986). All of these have a focus on the substance copies (mainly lexical copies) from other languages that can be found in Sakha. Antonov (1971) discusses the origin of Sakha lexical items divided by lexical domain, and within each domain by model language (Turkic, Mongolic, Evenki). Contrary to Kałużyński, he comes to the conclusion that the ancestors of the Sakha must have left the sphere of Mongol influence and migrated to the north prior to the rise of the Mongol Empire, i.e. before the 12th century; however, this is based not on a phonological analysis such as that performed by Kałużyński (1962), but on a purported lack of terms characteristic of the Mongol Empire (Antonov 1971: 165).

Romanova et al. (1975) highlight the 'mutual influence of Evenki and Sakha'. While they deal quite extensively with the Sakha influence on the Evenki dialects spoken in Yakutia, the section on the Evenki influence on Sakha is much shorter (less than 20 pages). This deals predominantly with some phonological influence to be found mainly in the northern, especially the northwestern dialects of Sakha (p. 145-157); but two suffixes copied from Evenki into the standard Sakha language and one suffix copied into two dialects are discussed as well (p. 157f), as are lexical copies from Evenki (p. 158-160). Structural influence from Evenki on Sakha is completely ignored, although the authors do provide an analysis of the calques from Sakha found in the language of Evenki folktales. Malchukov (2006) sketches some of the structural influence of Sakha on the Tungusic languages spoken in Yakutia, and discusses internal relative clauses in more detail, the structure of which he suggests was copied from Tungusic into Sakha rather than the other way around (pp. 130-133). Finally, Rassadin (1980) and Popov (1986) discuss

copied lexical items in Sakha; Rassadin bases his discussion on Kałużyński's (1962) data, while Popov analyzes words of 'unknown origin', i.e. words that preceding researchers had not been able to etymologize.

As becomes clear from the above discussion, although there have been several book-length monographs concerned with the role language contact played in the development of the Sakha language, most previous studies were concerned solely with analyzing substance copies in the language. There have been several suggestions of schematic copies (mainly from Evenki, but occasionally from Mongolic) found in Sakha; however, no data are presented to support these suggestions. Thus, based on the number of copied verbs in Sakha, Širobokova (1980: 140) suggests that Mongolic languages exerted substrate influence on Sakha: "The deep penetration of Mongolian elements in the Yakut language […] could only be the result of protracted bilingualism, since Turks do not borrow verbs." (translation mine[1]). Furthermore, it has been suggested that the change of [s] to [h] is due to Evenki substrate influence (Ubrjatova 1985a: 46), that the loss of the Turkic Genitive case in Sakha may be due to Tungusic influence (Schönig 1993: 157), that the extension of the Dative case to a marker of stative location may be due to either Tungusic or Mongolic influence (Poppe 1959: 680; Schönig 1990: 95), that the Sakha Comitative and Partitive case were copied from Evenki[2] (Ubrjatova 1956: 91; 1985a: 46; Schönig 1990: 95f), and that the subject agreement marking on canonical converbs can be ascribed to Tungusic influence as well (Ubrjatova 1956: 91; Johanson 2001: 1732). However, without a presentation and discussion of actual data, it is hard to evaluate such claims.

Stachowski & Menz (1998: 417) write: "There is considerable older Mongolic and later Russian influence [on Sakha], and a still little explored impact of Tungusic and Yeniseian substrate languages." This study aims at contributing to our knowledge of the impact of Tungusic languages on Sakha. Given the extensive literature on substance copies in Sakha, the focus here will be on some of the possible schematic copies from Evenki.

---

[1] Original: "Глубокое проникновение монгольских элементов в якутский язык […] могло быть только результатом длительного двуязычия, так как тюрки глагол не заимствуют."

[2] Schönig does give a very brief comparison of the function of the Tofa and Sakha Partitive case and the Evenki Indefinite Accusative, based on language descriptions, and is cautious about the possibility of Evenki contact influence: "Until there are reliable investigations about the use of these 'partitive' cases in both languages the question of such an influence remains open." (footnote 1 on p. 96)

1.4 Aims of this study and methodology adopted

1.4.1 Aims

As has been shown above (section 1.1.1.1), the Sakha language, although clearly belonging to the Turkic language family, differs greatly from its relatives. Thus, it has copied a large amount of lexical items as well as morphemes from Mongolic (Kałużyński 1962, *passim*), it has undergone a number of sound changes, and it shows divergent morphosyntactic features as well. It is known from archaeological and ethnographic data that the Sakha migrated north from a more southerly area of settlement (presumably close to Lake Baykal) several hundred years ago (Gogolev 1993; Alekseev 1996; cf. section 1.1.1.2). This long separation from fellow Turkic speakers may have led to the development of a number of independent innovations in Sakha[1] and thus to the divergence from other Turkic languages. On the other hand, the migration brought Sakha speakers into the vicinity of speakers of Tungusic languages (predominantly Evenks, but also Ėvens) as well as Yukaghir languages; thus, the influence of contact in the development of Sakha idiosyncrasies may have played a role as well.

Of course, to postulate contact influence in the development of certain features of a language is to postulate that the speakers of these languages were in contact with each other:

> "Linguistic change is initiated by speakers, not by languages. […] Linguistic changes, whether their origins are internal to a variety or not, are passed from speaker to speaker in social interaction. As for *language contact*, it is not actually languages that are in contact, but the speakers of the languages. […] the term 'language contact' therefore really means 'contact between speakers of different languages'." (Milroy 1997: 311, italics original)

In a non-literate society, such contact between speakers can only take place in direct interaction. This implies that the speakers of the languages interacted socially; the social interaction may have been sporadic and casual, or it may have been very intense, leading to intermarriage and the adoption of cultural practises of the neighbouring group. In the absence of historical data, it is very difficult to know what kinds of interaction a group such as the Sakha may have engaged in. After their migration north, they may have remained isolated from their neighbours, since their subsistence pattern of cattle- and horse-breeding would have necessitated their

---

[1] In this section, when I refer to Sakha as being divergent from the other Turkic languages, it is intended to include Dolgan as well. Although Dolgan has had a history of its own, and thus a study of the contact influence it has undergone during its development is required, most of the features that distinguish Sakha from Common Turkic appear to be shared by Dolgan.

settling in areas rich in grass, while the hunting and reindeer-herding Evenks, Ėvens, and Yukaghirs were nomads following the migration routes of wild reindeer, or settled along rivers rich in fish. It is also possible that during the historical expansion over the territory they occupy today, the Sakha were able to settle in regions depopulated by smallpox and measles, as described by Dolgix (1960: 385, 398, 408, 415, 443, 446f, 452f, 470). But it may also have been the case that the Sakha intermarried with the indigenous groups[2] after their migration north and after their expansion. It is unclear whether the differences in lifestyle (nomadic vs. settled, hunters and reindeer-herders vs. cattle- and horse-breeders) and language would have presented a barrier to intermarriage; given the fact that other 'more likely' marriage partners of the Sakha (i.e. other settled cattle- and horse-breeders) would have been lacking after their migration to the Lena, it is not unlikely that some amount of intermarriage took place, unless the immigrant group was large enough to furnish an autochtonous pool of marriage partners. That this, however, was not the case, at least with respect to the paternal half of the immigrating population, is clear from the genetic analyses (Pakendorf et al. 2002, 2006).

Contact influence has been postulated for a number of features that distinguish Sakha from other Turkic languages (cf. section 1.3); for example, the changes in the case system have been variously claimed to be the result of Evenki influence (Poppe 1959: 680f; Ubrjatova 1985a: 46, 118; Schönig 1990: 50; Nevskaya 2001: 299), while Mongolic influence has been suggested as an alternative for the extension of the Dative case to encompass locative functions (Poppe 1959: 680). Since Evenks were widespread in the area in which the Sakha initially settled, and into which they subsequently expanded (Dolgix 1960, map; cf. Figure 1.2), and since there exist claims of groups of Evenks shifting to the Sakha language and culture (Seroševskij [1896] 1993: 230f; Dolgix 1960: 369, 461, 486; Tugolukov 1985: 220), it is not surprising that influence of Evenki on the Sakha language is often assumed. However, in the absence of precise historical data, it is difficult to obtain true insights into the language contact situation that may have existed in the past. This is especially difficult (if not impossible) if language shift has taken place, because, if the shift was complete, no trace of the substrate language remains for comparison with structurally divergent features of the language that was the target of the shift (Thomason & Kaufman 1991: 111). In these cases, genetic studies may be of help, because a shifting group that has completely merged with the

---

[2] I here refer to Evenks, Ėvens and Yukaghirs as the 'indigenous groups' the Sakha would have come into contact with. Although the Tungusic-speaking groups may have immigrated to Yakutia not very long before the arrival of the Sakha, it is assumed they were already present in the area prior to the latter event (cf. section 1.1.2.2).

group whose language it adopted is expected to leave a detectable genetic trace in the genepool of the new population (e.g. Nasidze et al. 2004).

It is thus the aim of this study to combine both molecular anthropological and linguistic analyses to evaluate the extent to which the Sakha came into contact with the indigenous populations of the area in which they are currently settled, both from a physical (i.e. as regards admixture) and from a sociocultural perspective (as shown by linguistic contact influence). This combined approach will hopefully not only provide further evidence relating to Sakha prehistory, but will also enable further insights into the processes involved in language contact, since the combination of genetic and linguistic data can show up a correlation, or lack thereof, between physical and sociocultural contact. Thus, the molecular genetic analyses permit an estimate of the extent of genetic admixture that has taken place between the Sakha and the indigenous northeastern populations; furthermore, the use of mtDNA and Y-chromosomal analyses permits a differentiated view of whether such admixture was sexually biased, i.e. whether it was predominantly indigenous men or predominantly indigenous women who intermarried with the Sakha. On the other hand, the kinds of contact influence observed in the Sakha language may be able to provide some insight into the kind of sociocultural contact the populations were engaged in (cf. section 1.4.3).

The basic hypothesis with which I began this study in 2001 was that there had been substantial admixture in the maternal line from Evenks into Sakha (Pakendorf et al. 2003). I therefore expected to find evidence of substrate influence from Evenki in the Sakha language (Pakendorf 2001). Since the data on which my previous results were based were very limited, I included more samples of Sakha men from different regions of Yakutia as well as samples from some Evenk, Ėven, and Yukaghir groups in the genetic analyses (cf. section 2.2 and Pakendorf et al. 2006, 2007) to enable a better view of the genetic prehistory of the population. As shown by the current molecular anthropological analyses, however, the mtDNA lineages shared between the Sakha and the Tungusic-speaking groups, which led to the previous hypothesis of Evenk admixture in Sakha, are shared with South Siberian Turkic-speaking groups as well, implying that these populations may have shared a maternal gene-pool during the period when both the Northern Tungusic groups and the Sakha ancestors were still settled near Lake Baykal. Thus, admixture with Evenks after the migration of the Sakha to Yakutia, which is the focus of this investigation, cannot be shown in this extended study; however, it cannot be entirely excluded, either (Pakendorf et al. 2006). These inconclusive results of the genetic studies place a greater burden on the linguistic analyses for the elucidation of the prehistoric contact situation the Sakha may have found themselves in.

Given the results from my previous study (Pakendorf et al. 2003), which appeared to show strong signs of Evenk admixture in the maternal line, and given

the historical and current distribution of the Evenks and the Sakha, the focus of this study is for the most part directed towards the elucidation of contact influence from Evenki in the Sakha language, as well as further genetic analyses to elucidate the genetic prehistory of the populations of Yakutia. However, in one respect the Sakha differ greatly from most Tungusic-speaking groups and appear genetically close to Uralic-speaking peoples: the Sakha have the world's highest frequency of the Y-chromosomal SNP variant 'Tat C', which is hardly found in Tungusic-speaking groups, but is found in fairly high frequency in Uralic groups, from the Finns in the West to the Nenets in the East (Zerjal et al. 1997; Lahermo et al. 1999; Karafet et al. 2002; Pakendorf et al. 2006, 2007). This might be an indication of some Samoyedic substrate in Sakha, traces of which might possibly remain in the language. However, the Nganasans and Selkups, who are currently the easternmost Samoyedic-speaking groups, lack this polymorphism (Karafet et al. 2002; cf. section 1.1.1.2), complicating the picture somewhat. On the other hand, should there have been a substrate that was completely absorbed genetically by the incoming Sakha ancestors, there may be traces of Samoyedic substrate influence in the language that might still be detectable. I will return to the possibility of such a Samoyedic substrate in the Sakha language in the discussion (cf. section 5.2).

1.4.2 Methodology adopted for the assessment of linguistic contact influence

Since the extent of substance copies from Mongolic and Tungusic languages has been the subject of several previous studies (cf. the references in section 1.3), I focus here on the assessment of several features of Sakha that may represent schematic copies from the neighbouring languages. However, the interpretation of the kinds of contact the ancestors of the Sakha were engaged in cannot be complete without inclusion of lexical evidence; therefore, the evidence provided by the substance copies is reviewed in chapter 4.

In assessing the amount and kind of contact the Sakha language may have undergone from Tungusic languages, it is obviously of great importance to establish a) whether the feature in question is present in other Turkic languages, b) whether it is present in the Tungusic languages the Sakha speakers most probably would have been in contact with (Evenki and Èven), and c) whether it is found in other Tungusic languages. Only if a feature found in Sakha is not present in Turkic languages, but is found in Evenki and Èven as well as in other Tungusic languages, can I follow the heuristic proposed by Heine & Kuteva and conclude that the feature in Sakha is due to contact influence from Evenki or Èven:

"If there is a linguistic property x shared by two languages M and R, and these languages are immediate neighbours and/or are known to have been in contact with each other for an extended period of time, and x is also found in languages genetically related to M but not in languages genetically related to R, then we hypothesize that this is an instance of contact-induced transfer, more specifically, that x has been transferred from M to R." Heine & Kuteva (2005: 33)

In order to keep the amount of features analyzed in this study to a manageable level, only those in which Sakha differs from other Turkic languages were chosen for analysis. Since these features all distinguish Sakha from the South Siberian Turkic languages, which are the closest geographical relatives of Sakha, I assume that any contact influence that may have led to their development took place after the Sakha separated from the bulk of the Turkic speakers, after their migration to the north. Most of these features have been suggested as being due to contact influence (mainly from Evenki; cf. section 1.3 and the individual sections in chapter 3). Thus, this study is not only an attempt at elucidating Sakha prehistory from a combined linguistic and molecular anthropological perspective, but it is also an evaluation of the proposals made by others as to which features in Sakha are due to contact influence.

However, it may well be that Sakha and Evenki share a linguistic feature, but that this feature is found in neither the Turkic languages nor the Tungusic languages (cf. section 3.2.3). In such a case, although it is quite likely that contact between the languages was involved in the development of the feature, it may be impossible to judge the direction of influence. In such instances, I propose to follow Heath's method of 'internal reconstruction' (1978: 23, 74f):

"… if M_1 is a morpheme found in language X_1 and Y_1, but not in other members of either the X or Y groups and not reconstructable for Proto-X or Proto-Y, we can be fairly sure that diffusion has taken place but we have no comparative evidence bearing on the directionality problem. […] If, in the case of X_1 and Y_1, we can show by internal reconstruction that M_1 is likely to be relatively archaic in X_1 and shows no evidence of being archaic in Y_1, then we can conclude that X_1 was the probable source language and Y_1 has done the borrowing. Internal reconstruction of this type involves consideration of irregular allomorphic specialisation, unusual functional specialisation and/or restrictions, degree of integration into the morphosyntactic system, and the like." (Heath 1978: 23)

Siberian languages share some typological features [such as having for the most part SOV word order, being predominantly suffixing, and marking the possessor on the possessum with affixes (Dryer 2005: map 81, 26, and 57)]; this

sharing has been interpreted as indicating 'centuries of interaction and common development' (Anderson 2004: 2). Should a feature found in Sakha, but not in other Turkic languages, be widespread amongst Siberian languages, this would complicate the assignment of contact influence to one specific model language. In order to evaluate the prevalence of the features analyzed in this study amongst the languages of Siberia, I examine the respective features in a sample of Siberian languages in addition to assessing their value in the Turkic and Tungusic language family.

Of course, some changes may be due to internal developments rather than to contact influence. It is hard to distinguish between the two kinds of change from a purely linguistic perspective (i.e. disregarding possible genetic evidence for intimate contact between the speakers of the languages), but one approach advocated by Gensler (1993: 33f, 46) is to evaluate the cross-linguistic frequency of specific linguistic traits in a world-wide sample. Linguistic features that are shared by a large number of languages world-wide are more likely to have arisen through internal developments than features that are cross-linguistically rare. Such cross-linguistically rare features ('quirks') that are shared by genealogically unrelated languages are thus of much greater diagnostic value for the elucidation of prehistoric language contact. It is therefore desirable to have a reasonably large cross-linguistic sample in which the putative contact-induced features are examined in order to assess their world-wide frequency and their diagnostic value. However, the examination of several linguistic features in a typologically valid sample is a time-consuming undertaking. Given the extensive nature of the current project (brought about by the double amount of labour required by the dual approach of combining both genetic and linguistic analyses in one study), such a typologically valid cross-linguistic study of the features analyzed here is not feasible, even though I recognize the value of such an approach. Where possible, the World Atlas of Language Structures (edited by Haspelmath et al. 2005) is consulted; otherwise, the determination of relative frequency of the features examined here can only be judged in the perspective of the Siberian area.

1.4.3 Using language contact to draw inferences about population prehistory

It is the basic tenet of this study (as also proposed by Ross 2003: 192ff) that the different kinds of contact-induced changes outlined in section 1.2 may allow one to make inferences about the prehistory of a population that is assumed to have been in contact with populations speaking different languages. As mentioned above (section 1.4.2), differences between Sakha and its linguistic relatives can be taken as an indication that language contact may have taken place in the past (cf. Johanson 1999: 53; Heine & Kuteva 2005: 33), and since the perspective taken in this approach is to analyse the kinds of copies found in Sakha, Sakha can be defined as the recipient language with regard to the contact situations it was involved in.

Recipient-language agentivity involves primarily substance copies, while model-language agentivity involves primarily schematic copies (Van Coetsem 1988 as discussed by Winford 2005). Since in Van Coetsem's approach recipient-language agentivity is the term used to designate psycholinguistic dominance of a bilingual in the recipient language, while model-language agentivity designates psycholinguistic dominance of the model language, the kind of copies found in Sakha will allow me to deduce which language was in predominant use in the ancestral Sakha community, i.e. which language was used by a large number of speakers as their dominant language.

If I should find a large number of substance copies in Sakha, this would indicate that the speakers were dominant in Sakha (since in this analysis Sakha is identical to the recipient language), while conversely a large number of schematic copies would provide an indication of model-language dominance in the Sakha speech community. This claim of course rests on the assumption that a given change is due not only to a small but influential group of speakers (individuals with a lot of connections in the social networks) being bilingual and dominant in a certain language, but rather that we can obtain some insight into the state of language use for the group as a whole.

If only a small group of Sakha speakers were dominant in their ingroup language, the majority of the Sakha community would have been dominant in the outgroup language; in such a case, we would expect to find at least some changes due to model-language agentivity, i.e. schematic copies rather than substance copies due to recipient-language agentivity. If, on the other hand, only a small group of speakers were dominant in the model language, i.e. if the majority of the community were dominant in Sakha, this would imply that the community as a whole would have been relatively closed (*qua* Andersen 1988), and in such a group Sakha would have been in predominant everyday use by the majority of speakers. This assumption, however, precludes the existence of a small group of model-language

dominant bilinguals with extensive connections within the Sakha community, since individuals with extensive connections within their native community would be involved in extensive interactions within their community and would thereby probably be dominant in Sakha.

I therefore assume that if I should find a large number of substance copies in Sakha, the Sakha ancestors were involved in contact with the model language, but with dominance of their ingroup language in the community as a whole. Conversely, should I find a large number of schematic copies in Sakha this would imply that the Sakha ancestors were involved in contact with speakers of the model language and that the Sakha speakers were dominant in the model language at the time of contact.

Language shift can be detected by phonological influence in the recipient language (Thomason & Kaufman 1991: 39, 121; Ross 2003: 193). However, this holds only for cases of shift where the shifting group was large, or where the shift took place rapidly (Thomason & Kaufman 1991: 119f), so that the shifting speakers were not able to fully acquire the outgroup language they were shifting to.

1.4.4 Caveats

There are some caveats to be mentioned at the outset: first of all, genetic admixture will only be detectable when the two parental populations were sufficiently distinct from each other. If not, admixture cannot be proved, nor can it be disproved (cf. Pakendorf et al. 2006 and chapter 5), at least with the fairly restricted polymorphisms analyzed here (cf. section 2.2 and Pakendorf et al. 2006). Thus, the conclusions one can draw from such a study will be limited by the degree of genetic differentiation of the populations concerned. Furthermore, the conclusions one can draw from molecular anthropological studies depend heavily on the samples included for comparison. This holds especially true for such geographically widespread and fragmented populations as the Evenks and Ėvens, in which different subgroups can differ from each other quite substantially (Pakendorf et al. 2007). Thus, it may well be that I cannot detect conclusive signs of genetic admixture with the comparative samples included here, while inclusion of samples from different subgroups might provide a different picture. Another factor that may complicate the evidence derived from molecular anthropological studies is that genetic drift can erase traces of population affinities. Since drift has more of an impact in small populations (cf. Appendix 1, section 6), and the individual Tungusic-speaking groups were always fairly small (e.g. Dolgix 1960: 447, 454, 465f, 484), genetic drift may have had such an impact on the Evenks and Ėvens as to make judgements of their population affinities difficult (Pakendorf et al. 2007).

Similarly, there are some caveats regarding the linguistic side of the investigation as well. As with the lack of distinction between the genetic ancestors of the populations in contact, it may be very difficult to find evidence of linguistic contact influence in languages that are structurally quite close. Given the general typological similarity of Sakha and the Tungusic languages (e.g. SOV word order, suffixing agglutinative morphology, similar means of subordination by the use of participles and converbs), large-scaled structural changes (such as those found by Ross in the structurally very divergent languages Takia and Waskia) are not to be expected. Furthermore, although I was able to base my analysis of Sakha on actual data collected in the field (cf. section 2.1.1), for the evaluation of linguistic features found in other languages I was restricted to consulting grammars of the languages concerned. Although I tried to consult more than just one grammar where possible, this restriction limits my approach to the perspective and interpretation of language data offered by the writers of those grammars. This approach is also limited in that I have to base my judgement on synchronic language data. This may not provide a true picture of the historic distribution of the speakers of the languages, especially of such dialectally diverse and highly mobile peoples as the Evenks and the Èvens. Thus, Dorian's (1993:133) warning needs to be heeded in this study: "Unless one has personal experience of a contact setting, it is all too easy to read of influence from 'English', 'Spanish', or any other language very well known in a standardized form, and to assume that what we know as the standard form can be used in assessing the source, direction, and degree of the influence." (see also Johanson 2006: 7). Lastly, this study is restricted to the investigation of possible contact influence in the development of a limited number of features of Sakha, chosen because of their difference from Turkic languages. It can therefore not lay any claim to being exhaustive, and further investigations may well lead to somewhat different conclusions.

Taking all these caveats into consideration, I nevertheless believe that the task I have set myself is not impossible. However, I have tried to be as careful as possible in my evaluation of the possible contact-induced developments in Sakha – to the extent that it may be difficult to see the conclusions for the number of hedges I have raised. But I feel that it is better to err on the side of caution than to rashly assign all the features that are superficially shared by Sakha and the Tungusic languages, or Evenki, to contact influence.

1.4.5 The structure of this thesis

In chapter 2, I give an overview over the sources of the linguistic samples used in this study, as well as the provenance of the genetic samples analyzed. I furthermore give important information on the transcription used, as well as providing an overview of the grammars of Eurasian languages most frequently consulted (together with the abbreviations used in chapter 3 to refer to these grammars).

Chapter 3 is the most extensive chapter of this thesis. Here, I present a detailed discussion of the features examined, their presence or absence in the Turkic, Tungusic, Mongolic and other languages of Siberia, as well as my evaluation as to whether these features in Sakha may be due to contact influence or whether they represent an internal innovation. Since this judgement is frequently not at all straightforward, the individual sections of chapter 3 are quite extensive; however, it was deemed necessary to give detailed arguments to let the readers judge for themselves whether my conclusions are correct.

Chapter 4 provides a very brief overview over the substance copies found in Sakha and some phonological changes associated with them, based predominantly on work by other authors. In chapter 5 I discuss the genetic and linguistic results in the light of the prehistoric population contact the Sakha engaged in. and offer an outlook for further studies that may still be necessary.

The genetic results have been published in relevant scientific journals (Pakendorf et al. 2006; Pakendorf et al. 2007). Since this thesis has been written in fulfillment of the requirement for a Ph.D. in Linguistics, the focus here is on the linguistic aspects of this work. The genetic results are therefore not included in the body of this thesis, but a summary of the main findings is provided in section 5.4. For details, readers are referred to the original articles. In order to facilitate an understanding of the the results presented there, as well as in the discussion in chapter 5, a brief introduction to the most important issues in Molecular Anthropology is provided in Appendix 1. Appendix 2 shows a figure not included in Pakendorf et al. (2007), while a table showing the case suffixes in the simple and possessive declension in Sakha, and a table showing the case suffixes in the Tungusic language family have been added for reference in Appendix 3 and 4.

60

2 DATA AND METHODS OF ANALYSIS

The data analyzed in this thesis were collected during three periods of fieldwork in the Republic Sakha (Yakutia). The first field trip took place from the end of February until mid-October 2002, when I collected genetic samples in seven districts (cf. section 2.2) and linguistic data in four of these districts (cf. section 2.1). The districts were chosen so as to represent the four dialectal groups defined for the Sakha language (Voronkin 1999: 154f; cf. section 1.1.1.1). From the central dialect group, the Taatta district was chosen for linguistic fieldwork, since this is the district most of the early Sakha intellectuals came from, and the subdialect of this district may therefore have had a strong impact on the standard language. The Suntar district was chosen to represent the Vilyuy dialect group, since it was recommended to me as having retained the traditional culture most strongly. From the northwestern dialect group the Olenëk district was chosen, since of the two districts on the territory of the Republic Sakha (Yakutia) belonging to this dialect group, it is more easily accessible to foreigners than the Anabar district, which lies in the border zone and hence requires a special entry permit. Lastly, the Verxojansk district was chosen from the northeastern dialect group, since this is one of the few regions outside the Lena-Amga-Aldan triangle already inhabited by the Sakha at the time of first Russian contact (cf. Figure 1.2).

The second field trip took place in November and December 2003. The linguistic part of this trip was restricted to one month in the Verxojansk district, followed by genetic sample collection in the Èveno-Bytantaj and Tompo districts (cf. section 2.2). The third field trip, which took place in August 2006, was restricted to clarifying some open questions concerning the Sakha language during three weeks in the Verxojansk district; unfortunately, due to bad weather conditions a planned sojourn in the Èveno-Bytantaj district to clarify some questions concering Èven was not possible.

2.1 Linguistic data and methods

2.1.1 Linguistic data

During the first field trip I started off the linguistic data collection in the Verxojansk district, where I stayed in the village of Tabalaax[1] (~ 1,000 inhabitants) for three weeks in April/May, followed by a further week in the town of Verxojansk (~ 1,900 inhabitants) (cf. Figure 2.1). In June I went to the Suntar district, where I spent five weeks in a small village called Xadan (~ 600 inhabitants). I spent five weeks from the end of July until the end of August in the Olenëk district, where I spent one week in Xaryalaax (~ 800 inhabitants), two weeks in Žilinda (~ 900 inhabitants), and another two weeks in the district capital Olenëk (~ 2,100 inhabitants). Finally, I spent nearly four weeks in September in the Taatta district, the first three days in Xarbalaax (~ 1,300 inhabitants) on the Amga river, and over three weeks in Bajaga (~ 1,000 inhabitants). In 2003, I restricted my linguistic research to a four-week sojourn in Tabalaax (Verxojansk district) in November, and in 2006, I spent three weeks in Tabalaax in August.
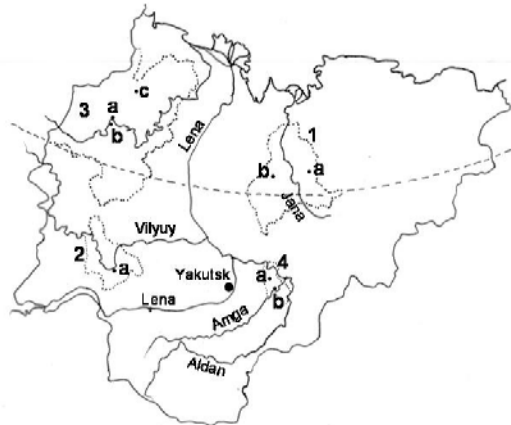
Figure 2.1: Map of the Republic Sakha (Yakutia) showing the four districts in which linguistic fieldwork was conducted. 1: Verxojansk district, 1a: Tabalaax, 1b: Verxojansk; 2: Suntar district, 2a: Xadan; 3: Olenëk district, 3a Olenëk, 3b: Xaryalaax, 3c: Žilinda; 4: Taatta district, 4a: Bajaga, 4b: Xarbalaax. The dashed line represents the Arctic Circle.

---

[1] Most villages in Yakutia have two names; one in common usage and one official name found in maps and publications. Here, I give the commonly used names. All estimates of number of inhabitants given here are taken from Safronov (2000) and are based on the 1989 census; however, in my estimate they do not differ greatly from current numbers.

In each of the villages, I worked closely with native speakers of Sakha, mainly women whose age ranged between 29 and 74, all of whom were fluent in Russian (cf. Table 2.1). From these women, I elicited translations of Russian sentences into Sakha and back translations of Sakha sentences into Russian. These Sakha sentences were partly obtained from consultants, and also constructed by myself with the help of Sakha grammars. In addition, I re-elicited selected sentences from additional native speakers in each of the villages.

Table 2.1: My primary consultants[2]

Location	Name	Age[3]	Education and profession
Tabalaax	E.I. Migalkina	*1973	higher, teacher of Russian
	I.I. Stručkova	~ 40	higher, teacher of Sakha
	M.N. Slepcova	*1978	intermediate specialized, primary school teacher
Verxojansk	L.M. Postnikova	50	intermediate specialized, retired fel'dšer[4]
Xadan	M.E. Ivanova	32	higher, handicraft teacher
	O.G. Arxipov	49	intermediate specialized, forestry worker
	L.A. Arxipova	50	higher, teacher of geography
	R.I. Arxipova	50	higher, teacher of French and Russian
Xaryalaax	N.N. Tipjanova	~ 60	higher, retired teacher of Russian
	T.V. Xristoforova	17	high school student, 11th grade
	X.K. Matveeva	74	intermediate specialized, retired fel'dšer
Žilinda	T.E. Nikolaeva	51	higher, teacher of Russian
Oleněk	M.A. Anisimova	73	intermediate specialized, retired culture worker
Bajaga	O.N. Makarova	48	higher, teacher of Russian
	A.R. Raxleeva	72	intermediate specialized, retired fel'dšer

However, my main corpus of data consists of recorded, transcribed and translated texts (mainly personal life stories) from 15 native speakers of Sakha[5], of whom six were men, nine women (cf. Table 2.2). The age of these speakers ranged

---

[2] In addition, the following women were kind enough to transcribe texts for me: A.V. Semënova in Žilinda, V.V. Blaxirova in Xarbalaax, and M.N. Slepcova in Tabalaax.

[3] Since I worked with Elizaveta Migalkina in 2002, 2003 and 2006, and with Marija Slepcova in 2003 and 2006, I give the year of their birth. For my other consultants, I indicate their age at the time I worked with them.

[4] A fel'dšer is a medical worker with a qualification intermediate between a nurse and a doctor.

[5] I actually recorded more texts, but was unable to transcribe them all due to lack of time and technical problems.

from 63 to 95 years of age, and eight were completely monolingual Sakha speakers. I recorded the stories on a Sony MZ R-900 minidisc recorder with a Philips SBC ME670 microphone, using the LP4 mode. Since most informants told me their life stories, the texts are roughly comparable in contents and register; however, R.E. Xatylaev from Bajaga is a marked exception, since he recounted the history of the region and described some aspects of traditional culture. Initially, I transcribed the texts with the help of my consultants; however, starting in Žilinda in August 2002 I asked native Sakha speakers to transcribe the stories for me (in standard Sakha orthography), and then converted this into my Latin-based transcription. I translated the texts to the best of my ability with the help of the Sakha-Russian dictionary (Slepcov 1972) and then clarified all my questions and checked the accuracy of my translation with my consultants in each of the villages.

Table 2.2: The people who told me their life stories

Location	Name	Age	Knowledge of Russian	Abbreviation[6]
Tabalaax (2002)	X.L.E. &	66	none	Efmy
	P.I.E.	77	none	
(2003)	A.I. Čirikov	88	none	Chir
	P.V. Potapov	79	none	PotP
Xadan	A.S. Pavlova[7]	90/95	probably none	P90/P95
	P.Ja. Ivanova	70	slight	IvaP
	E.S. Ymyčanov	70+	slight	YmyE
Xaryalaax	X.K. Matveeva[8]	74	good, quite some code-switching	MatX1/MatX2
Žilinda	M.A. Nikolaeva &	80	none	Afny
	A.A. Anisimova	82	none	
	P.N. Lukinov	75	hardly any	LukP
Xarbalaax	P.I. Bestinova	63	none	BesP
	A.G. Malyševa	71	relatively good	MalA
Bajaga	R.E. Xatylaev	69	slight	XatR
	A.R. Raxleeva	72	good, a lot of code-switching	RaxA

---

[6] This refers to the reference given for each example in chapter 3.

[7] Although A.S. Pavlova had already died, her son O.G. Arxipov kindly let me record two brief video recordings he had made of his mother on her 90th and 95th birthday, respectively.

[8] I made two recordings of X.K. Matveeva on two different days; the first time she told me her life story, the second time she told about daily life as a nomadic reindeer herder.

In addition to the spoken texts recorded by myself in Yakutia I also made use of the text included in the first grammar of Sakha published by Otto Böhtlingk in 1851 (facsimile reprinted in 1964). This text, the memoirs of Uvarovskij, a Russian cossack born and raised in Žigansk and near Yakutsk, is the earliest textual data on Sakha available. It differs markedly from my modern-day data by being a written text, and therefore containing much longer and more complex sentences than the spoken narratives I collected; however, since it represents an early stage of the language before implementation of prescriptive grammar rules, I considered it an important complement to my own data. Due to time constraints, at time of writing I had managed to analyze only 268 sentences, approximately one third of the entire text.

In chapter 3, examples taken from the texts are indexed with an abbreviation of the name of the speaker as shown in Table 2.2 followed by the number of the utterance for that speaker. Examples taken from Uvarovskij's narrative are referenced with the abbreviation [Uvar]. Examples taken from the elicited data are indexed as 'translations' of Russian sentences into Sakha or as 'back translations' of Sakha sentences into Russian, followed by the abbreviation of the district name where the example was obtained. These abbreviations are as follows: Oln = Olenëk district, Sun = Suntar district, Tat = Taatta district, and Ver = Verxojansk district.

2.1.2 Linguistic methods

As mentioned in section 2.1.1 the stories I recorded were transcribed either by myself with the aid of my consultants, or by native Sakha speakers and then converted from the Cyrillic-based modern Sakha orthography to my Latin-based transcription. The transcription I chose is generally broad and follows the standard orthography, with the exception that I indicate differences of pronunciation that might be of relevance to the topic at hand, e.g. the use of word-initial [h] instead of [s]. I follow the transcription of Stachowski & Menz (1998), except that for ease of use I represent long vowels by a colon instead of a macron, and I use a plain *j* to represent the voiced palatal affricate. The letter *y* represents the palatal oral glide, while *ỹ* represents the palatal nasal glide present in words such as *iỹe* 'mother'.

Since Sakha is characterized by both palatal and labial vowel harmony as well as consonantal sandhi across morpheme boundaries, morphemes may have widely varying surface forms depending on the context in which they occur. Thus, the past participle can take on the form *–bït–*, *–mup–*, *–pik–* and a large number of others. I therefore follow Turcological notational conventions by giving the underlying archiphonemic form for morphemes in isolation, e.g. *–BIt* for the past

participle. In these archiphonemes, an upper-case *I* stands for any high vowel, while an upper-case *A* stands for any low vowel; upper-case consonants indicate consonants that undergo assimilation to the root or stem, while lower-case consonants are invariant. (Thus, the assimilation undergone by stem-final consonants under influence of following morphemes is not indicated in the archiphonemic notation.)

The transcribed texts were entered into the programme SHOEBOX vers. 5.0 (SIL International), where they were provided with interlinearized glosses following the 'Leipzig Glossing Rules' (http://www.eva.mpg.de/lingua/files/ morpheme.html). Some glossing abbreviations, however, were shortened from those in the 'Leipzig Glossing Rules' to make the glosses more readable. Furthermore, for case suffixes of the possessive declension (which are fused too much to be easily segmentable) the gloss 'POSS' for possessive was omitted; thus, these glosses contain only the information on case plus person/number, e.g. 'INS.3SG' indicates the third singular possessive form of the Instrumental case marker. The auxiliary verb in converbal constructions was glossed according to the aspectual nuance expressed by the overall predicate, e.g. *öl–ön χa:l–la* [die–PF.CVB **RES**–PST.3SG] 'died', where the verb *χa:l* 'remain' adds an aspectual nuance of resultativity. As can be seen from the examples in chapter 3, the interlinearization contains an additional line between the text and the glosses consisting of the archiphonemic representation of the text. This was included in the examples since, despite the fact that Sakha is typologically an agglutinative language, it is not always possible to separate morphemes cleanly. For example, the Imperfective Converb, which is marked by just a low vowel (archiphonemically *–A*), merges with the final vowel of verbs ending in a long vowel or diphthong, resulting in a long high vowel. Thus, the Imperfective Converb of the verb *o:nŋo:* 'to play' is *o:nŋu:*. In such a case, it is not possible to assign the long high vowel to either the root or the morpheme, and I deemed it more accurate to leave the actual text untouched and indicate morpheme boundaries in a separate line, even though this led to a superficial lengthening of the thesis.

In order to evaluate the degree of contact influence undergone by the Sakha language, I compared it to other Turkic and Siberian languages. For this I had to rely on published descriptions. I tried to examine the features under consideration in at least one language from every major branch of Turkic, following Johanson (1998b). In addition, I included Old Turkic as the oldest documented form of Turkic to provide me with an indication of what the Sakha ancestor may have looked like. For the southwestern Turkic branch, I mainly consulted grammars of Turkish and occasionally Turkmen, while for the northwestern branch I relied primarily on

grammars of Kazakh and Kirghiz. The southeastern branch was represented by Uzbek in my analysis, while from the northeastern branch I included Khakas, Tuvan, and Tofa. Furthermore, I included the Turkic outliers Chuvash and Khalaj in my comparisons. Since Dolgan, which is the closest linguistic relative of Sakha and even considered a Sakha dialect by some, is said to have undergone considerable influence from Evenki, I included it as well as a measure of what changes language contact might achieve. In addition to consulting grammars of Turkic languages to obtain a perspective on the ancestral state of the features examined for Sakha, I also studied the features in Mongolic languages, in Evenki, Ėven, and other Tungusic languages (especially Udihe, Oroč, Nanay, and Manchu, for which quite comprehensive descriptions are available), as well as in other Siberian languages to obtain a broader perspective on the possible areality of the feature of interest. Wherever possible, I myself segmented and glossed examples taken from grammars and other descriptions that did not include glosses. Where I was unable to find a translation for a root or a gloss for a morpheme, this is marked with two question marks in the glosses; translations for roots or glosses of which I am not entirely sure are marked with one question mark, e.g. [that? man wife–POSS.3SG girl–POSS.DU:SG–INS ?? beat?–PST.DU.OBJ.3SG] 'that man beat his wife together with his daughter'. The glosses of other authors were adapted to the style and abbreviations adopted in my own work. In general, I followed the transcription of the original author, with the exception of such grammars that were written in Russian and where the transcription was in Cyrillic; these were latinized. All transcriptions were adapted to be somewhat consistent with the one chosen for Sakha; e.g. the palatal glide, which is often transcribed as *j*, was changed to *y* to distinguish it from the palatal voiced affricate, for which I have chosen the transcription *j*. The following is a list of the grammars or grammar sketches consulted most frequently for each language with the abbreviations used to refer to them in chapter 3.

Table 2.3: Grammars consulted

Language	Grammar consulted	Reference
Turkic languages		
Old Turkic	Erdal, M. (2004): A Grammar of Old Turkic. Leiden, Boston: Brill.	Erd
Turkish	Kornfilt, J. (1997): Turkish. London: Routledge.	Krn
	Göksel, A & Kerslake, C. (2005): Turkish: A Comprehensive Grammar. London, New York: Routledge.	G/K
Turkmen	Clark, L. (1998): Turkmen Reference Grammar. Wiesbaden: Harrassowitz Verlag.	Clk
Kazakh	Sovremennyj kazaxskij jazyk. Fonetika i morfologija (1962). Alma-Ata: Izdatel'stvo Akademii nauk Kazaxskoj SSR.	SKJ
Kirghiz	Imart, G. (1981): Le Kirghiz. Description d'une langue de littérisation récente. Aix-en-Provence: Publications de l'Université de Provence.	Imr
Bashkir	Grammatika sovremennogo baškirskogo literaturnogo jazyka (1981). Moskva: Izdatel'stvo 'Nauka'.	GSBJ
Uzbek	Bodrogligeti, A.J.E. (2003): An Academic Reference grammar of Modern Literary Uzbek. München: LINCOM Europa.	Bdr
Tuvan	Anderson, G.D. & Harrison, K.D. (1999): Tyvan. München: LINCOM Europa.	A/H
Khakas	Anderson, G.D. (1998): Xakas. München: LINCOM Europa.	And
Tofa	Rassadin, V.I. (1978): Morfologija tofalarskogo jazyka v sravnitel'nom osveščenii. Moscow: Izdatel'stvo 'Nauka'.	Ras
Sakha	Grammatika sovremennogo jakutskogo literaturnogo jazyka. Fonetika i morfologija (1982). Moskva: Izdatel'stvo 'Nauka'.	GSJa
	Böhtlingk, O. ([1851] 1964): Über die Sprache der Jakuten. Facsimile reprint. The Hague: Mouton & Co.	Btl
Dolgan	Ubrjatova, E.I. (1985): Jazyk noril'skix dolgan. Novosibirsk: Izdatel'stvo 'Nauka', Sibirskoe otdelenie.	Ubr
Khalaj	Doerfer, G. (1988): Grammatik des Chaladsch. Wiesbaden: Otto Harrassowitz.	Drf

Mongolic languages		
Written Mongolian	Poppe, N. [1954] 1991: Grammar of Written Mongolian. Wiesbaden: Otto Harrassowitz.	PopWM
Khalkha Mongolian	Kullmann, R. & Tserenpil, D. (2001): Mongolian Grammar. Ulaanbaatar: Academy of Sciences, Inst. of Language and Literature.	K/Ts
Buryat	Poppe, N.N. (1960): Buriat Grammar. Bloomington: Indiana University, The Hague: Mouton & Co.	PopB
	Skribnik, E. (2003): Buryat. In: Janhunen, J. (ed): The Mongolic Languages. London, New York: Routledge: 102-128.	Skr
Dagur	Tsumagari, T. (2003): Dagur. In: Janhunen, J. (ed): The Mongolic Languages. London, New York: Routledge: 129-153.	Tsum
Tungusic languages		
Evenki	Nedjalkov, I. (1997): Evenki. London: Routledge.	Ned
	Bulatova, N. & Grenoble, L. (1999): Evenki. Munich, Newcastle: LINCOM EUROPA.	B/G
	Konstantinova, O.A. (1964): Èvenkijskij jazyk.	Kon
Èven	Novikova, K.A. (1960): Očerki dialektov èvenskogo jazyka. Ol'skij govor. Čast' 1. Moscow, Leningrad: Izdatel'stvo Akademii Nauk SSSR.	NovI
	Benzing, J. (1955): Lamutische Grammatik. Wiesbaden: Franz Steiner Verlag.	Ben
	Malchukov, A. (1995): Even. Munich, Newcastle: LINCOM EUROPA.	Mal
Nanay	Avrorin, V.A. (1959): Grammatika nanajskogo jazyka. Tom pervyj. Moscow, Leningrad: Izdatel'stvo Akademii Nauk SSSR.	AvrI
	Avrorin, V.A. (1981): Sintaksičeskije issledovanija po nanajskomu jazyku. Leningrad: 'Nauka'.	AvrIII
Orok	Petrova, T.I. (1967): Jazyk orokov (ul'ta). Leningrad: Izdatel'stvo 'Nauka', Leningradskoe otdelenie.	Pet
Udihe	Nikolaeva, I. & Tolskaya, M. (2001): A Grammar of Udihe. Berlin, New York: Mouton de Gruyter.	N/T
Oroč	Avrorin, V.A. & Boldyrev, B.V. (2001): Grammatika oročskogo jazyka. Novosibirsk: Izdatel'stvo SO RAN.	A/B
Manchu	Gorelova, L.M. (2002): Manchu Grammar. Leiden, Boston, Köln: Brill.	Gor

Uralic languages		
Northern (Obdorsk) Khanty	Nikolaeva, I. (1999): Ostyak. Munich: LINCOM Europa.	Nik
Eastern (Vakh) Khanty	Tereškin, N.I. (1961): Očerki dialektov xantyjskogo jazyka. Moscow, Leningrad: Izdatel'stvo Akademii Nauk SSSR.	Trš
(Sosva) Mansi	Riese, T. (2001): Vogul. Munich: LINCOM Europa.	Rse
	Rombandeeva, E.I. (1973): Mansijskij (Vogul'skij) jazyk. Moscow: Izdatel'stvo "Nauka".	Rmb
Nganasan	Tereščenko, N.M. (1979): Nganasanskij jazyk. Leningrad: Izdatel'stvo "Nauka", leningradskoe otdelenie.	Ter
Southern Selkup	Bekker, È. G. (ed): Morfologija sel'kupskogo jazyka. Južnye dialekty. 1995. Tomsk: Tomskij gosudarstvennyj pedagogičeskij institut.	MSJ
Northern Selkup (Taz dialect)	Kuznecova, A.I.; Xelimskij, E.A.; Gruškina, E.V. (1980): Očerki po sel'kupskomu jazyku. Moscow: Izdatel'stvo Moskovskogo universiteta.	OSJ
Other Siberian languages		
Kolyma Yukaghir	Maslova, E. (2003): A Grammar of Kolyma Yukaghir. Berlin, New York: Mouton de Gruyter.	Mas
Chukchi	Dunn, M.J. (1999): A Grammar of Chukchi. Australian National University: PhD thesis.	Dnn
Itelmen	Georg, S. & Volodin, A.P. (1999): Die itelmenische Sprache. Wiesbaden: Harrassowitz Verlag.	G/V
Siberian Inupik Eskimo	Menovščikov, G.A. (1980): Jazyk èskimosov beringova proliva. Leningrad: Izdatel'stvo "Nauka", leningradskoe otdelenie.	Men
Nivkh	Gruzdeva, E. (1998): Nivkh. Munich: LINCOM Europa.	Grz
	Mattissen, J. (2003): Dependent-Head Synthesis in Nivkh. A Contribution to a Typology of Polysynthesis. Amsterdam/Philadelphia: John Benjamins Publishing Company.	Mat
	Panfilov, V.Z. (1962): Grammatika nivxskogo jazyka. Čast' 1. Moscow, Leningrad: Izdatel'stvo Akademii Nauk SSSR.	PanI
Ket	Werner, H. (1997): Die ketische Sprache. Wiesbaden: Harrassowitz Verlag.	Wer

	Dul'zon, A.P. (1968): Ketskij jazyk. Tomsk: Izdatel'stvo Tomskogo Universiteta.	Dul
Several Tungusic and other Siberian	Jazyki narodov SSSR, V: Mongol'skie, tunguso-man'čžurskie i paleoaziatskie jazyki. Leningrad: Izdatel'stvo 'Nauka', Leningradskoe otdelenie. 1968	JaN5

2.2 Genetic data and methods

In March and April 2002, I undertook three sample collecting trips to the Taatta, Čurapča, Suntar, Njurba, and Megino-Kangalas[9] districts (cf. Figure 2.2) together with Dr. Al'bina Danilova from the Institute of Health, Yakutsk. During these collection trips, we were supported by the medical administration of each district; the samples were collected in village hospitals with the assistance of the local staff. Dr. Danilova, being a doctor, took 10 ml of whole blood from healthy, unrelated male volunteers after informed consent. The blood samples were frozen immediately after collection by storing them at ambient temperature and were transported to Yakutsk in their frozen state. DNA was extracted at the Institute of Health by T.Sivceva and A.Fedorov, and I obtained an aliquot of each DNA sample upon leaving the Republic Sakha (Yakutia) in October 2002.

During the rest of my 2002 field trip I collected cheek swab samples in the villages in which I was staying; in this, I was generally assisted by my hosts. However, in the Verxojansk district, I collected samples not only in Tabalaax and Verxojansk where I was doing linguistic fieldwork, but also in Adyča and in Boronuk, while in the Taatta district I collected samples in Čymnaji, Čyčymax and Xarbalaax along the Amga river with the generous assistance of Egor and Valentina Blaxirov. During this phase of collection I generally visited people in their homes; however, in Tabalaax, Adyča, and Boronuk the local village administrations let me make use of an office in the administration buildings, while in Verxojansk I collected samples at the offices of Verxojansktorg and the ŽKX (Žiliščno-kommunal'noe xozjajstvo). Unfortunately, especially in the Verxojansk and Olenëk districts a number of the volunteers were related, so that not all samples could be used for this study, as can be seen from Table 2.4.

---

[9] Most of the samples collected in the Megino-Kangalas district were not used in this study, since the number of Sakha samples was already quite extensive.

Since the climate and ecology of the Olenëk district is unsuitable for cattle- and horse-breeding, the inhabitants practise Evenk-style reindeer herding and hunting. They identify themselves as Evenks; however, as far back as the grandparental generation, they have spoken only Sakha (personal observation). An ethnographic monograph, the data for which were collected in the 1940s, described this group as Sakha, and traced most of the clans which are present in the district nowadays to Sakha who had immigrated in the past centuries (Gurvič 1977). Tugolukov (1985: 220) says that this group of Sakha-speaking reindeer herders resulted out of large-scale intermarriage between Sakha and Evenks; however, in the ensuing discussion of individual clans, most of those present nowadays whose samples were included in the collection (Betu, Čordu, Osogostox) are described as being immigrants from central Yakutia with Sakha roots (partly with Tungusic admixture). In order to distinguish this group from Sakha-speaking cattle and horse pastoralists and from Evenki-speaking reindeer herders, the samples from the Olenëk district were classified as 'Sakha-speaking Evenks' in all the analyses ('Yakut-speaking Evenks' in Pakendorf et al. 2006, 2007). This is a purely descriptive term and is not meant to reflect any opinion on their ethnic affiliation.

Furthermore, in 2002 I obtained a number of samples from the DNA archive of the Insitute of Health: 53 Yukaghirs from Andrjuškino in the Lower Kolyma district and Nelemnoe in the Upper Kolyma district, as well as 42 Evenks from Iengra in the Nerjungri Territory, Republic Sakha (Yakutia). Unfortunately, these were samples from nuclear families, and a large number of them were women, so that very few of the samples could be used in this study; the number of samples that could be used for the Y-chromosomal analyses was particularly small (cf. Table 2.4). One maternal lineage from the Olenëk district was Yukaghir and was included with the Yukaghir sample in the analyses, while one maternal lineage from Megino-Kangalas, and two paternal lineages from the Verxojansk and Olenëk districts were Evenks and were included with the Iengra Evenk sample in the analyses. On account of the very low sample sizes, it was unfortunately not possibe to analyze the Tundra Yukaghir samples (from Andrjuškino) and the Kolyma Yukaghir samples (from Nelemnoe) separately. In addition, I obtained six samples from the Vilyuy district from the Institute of Health (classified as belonging to the Vilyuy dialect group, together with the samples from the Suntar and Njurba districts).

In December 2003 I undertook two trips to collect samples from Ėvens; in the Ėveno-Bytantaj district (cf. Figure 2.2) I collected samples in Sakkyryyr, Kustuur, and Džargalaax with the kind assistance of the district administration – these were later classified as Western Ėvens following the dialectal division of the Ėven language. In the Tompo district I collected samples in the village of Topolinoe

with the assistance of my host, Ol'ga Nikolaeva; these were classified as Central Ėvens. In some cases, the maternal lineage came from the western dialect group, while the paternal lineage came from the central dialect group. Four maternal lineages from Andrjuškino were Ėvens and were included with the Central Ėven samples in the analyses.
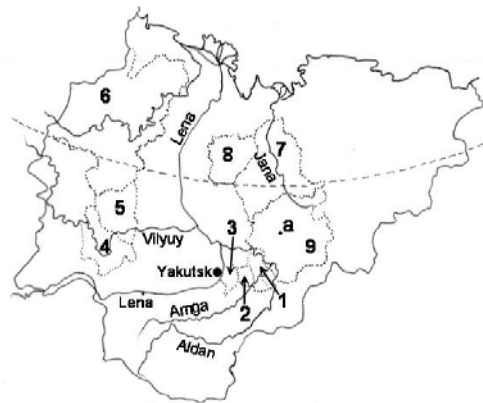
Figure 2.2: Map of the Republic Sakha (Yakutia) showing the districts in which genetic fieldwork was conducted. 1: Taatta district; 2: Čurapča district; 3: Megino-Kangalas district; 4: Suntar district; 5: Njurba district; 6: Olenëk district; 7: Verxojansk district; 8: Ėveno-Bytantaj district; 9: Tompo district (9a: Topolinoe). The dashed line represents the Arctic Circle.

Innokentij Novgorodov (then affiliated with the Institute of Humanitarian Studies in Yakutsk) helped me greatly by undertaking to collect samples from Evenks, Tuvans, and Sakha belonging to the *akan'e* dialects (cf. section 1.1.1.1). In the summer of 2003 Dr. Novgorodov collected 40 samples from unrelated healthy Evenk men in the southern part of the Evenk National District, and 60 samples from unrelated healthy Tuvan men in Kyzyl and the Ovjursk district, Republic Tuva. In March 2004 he collected samples from the Aldan and Namcy districts in the Republic Sakha (Yakutia). The Sakha dialects spoken in these districts traditionally show *akan'e*, and it has been suggested that this feature is due to a Mongolian substrate (Ubrjatova 1960: 42; Voronkin 1999: 57ff).

As far as possible, all the samples were collected from unrelated, healthy male volunteers after informed consent was obtained. Each volunteer was asked

about his parents' and grandparents' birthplace, ethnic affiliation and first and second language in order to exclude individuals with known admixture.

DNA was extracted from the cheek swab samples in Leipzig. The mitochondrial hypervariable region I (HVR1) was sequenced in all samples representing unrelated mtDNA lineages; furthermore, the mtDNA haplogroup assignment for these samples was confirmed by typing the relevant single nucleotide polymorphisms (SNPs; performed by Patricia Heyn under my supervision). For all samples representing unrelated Y-chromosomal lineages, a set of 15 Y-chromosomal SNPs known to be of interest in Eurasia was typed in a more or less hierarchical manner (although the Sakha samples were typed first for the presence or absence of Tat C) and nine Y-STRs (short tandem repeats) were typed in all of these samples as well (part of the Y-STR typing was done by Matti Heino under my supervision). For details concerning the SNPs and STRs as well as the lab methods see Pakendorf et al. (2006) and Pakendorf et al. (2007).

A number of standard statistical and phylogeographic analyses (diversity calculations, estimates of pairwise Fst values, Multidimensional Scaling analyses (MDS), and MJ-networks) were performed with the HVR 1 sequences, the Y-chromosomal SNP frequency data, and the Y-STR haplotypes on the background of the individual SNPs. For details see Pakendorf et al. (2006, 2007).

Table 2.4: Number of samples collected in each location

	N samples collected	N samples used for analyses	
Location/ethnic group		mtDNA	Y-chrom.
Taatta district	37	29	29
Čurapča district	27	16	18
Megino-Kangalas district	26	8	8
Aldan district	21	20	20
Namcy district	20	18	21[10]
Suntar district	32	28	29
Njurba district	26	21	21
Viljuj district	6	6	6
Verxojansk district	44	32	32

Table 2.4: Number of samples collected in each location, cont.

---

[10] A sample collected in the Verxojansk district originally came from the Namcy district and was included in this count.

Total Sakha	**233**	**178**	**184**
Olenëk district: Sakha-speaking Evenks	51	32	33
Western Èvens	29	24	22
Central Èvens	24	26[11]	24
Iengra Evenks	42	23	9[12]
Yukaghirs (Andrjuškino)	25	13[13]	6
Yukaghirs (Nelemnoe)	28	9	7
Evenks (Evenk National District)	40	39	40
Tuvans	61	59	55
Total samples	**533**	**403**	**380**

[11] Including the four samples from Andrjuškino.
[12] Including the two paternal lineages from the Verxojansk and Olenëk district.
[13] Including the one maternal lineage from the Olenëk district.

76

3 ANALYSIS OF DIVERGENT TRAITS OF SAKHA

In this chapter I present a detailed analysis of some features in which Sakha differs from its Turkic sister languages. These are the shift of [s] to [h], which has been completed in intervocalic position, but may still be ongoing in word-initial position; a number of differences in the case system, the development of a category of Distant Future Imperative, changes in the system of possessive marking, and the possibility of marking subject agreement on converbs. Thus, I analyze features from several subsystems of the grammar: phonology, morphosyntax, and syntax. For most, if not all, of these features contact influence has been suggested, generally from Evenki. However, as will be seen, that conclusion cannot always be upheld.

3.1 The shift of [s] to [h] in Sakha

Sakha has undergone a large number of sound changes, making mutual comprehension with other Turkic languages quite difficult. Amongst others, Common Turkic word-initial [s] has become lost (Johanson 1998b: 103), e.g. *u: < *su:w* 'water', *ü:t < *sü:t* 'milk' (Tenišev 2001: 88, 448f), while intervocalic [s] sometimes changed to [t], e.g. *iti: < *isiɣ* 'heat, hot' (Tenišev 2001: 19f); in word-final position, however, [s] was retained (Ščerbak 1970: 162).

The phoneme /s/ found in modern-day Sakha is partly a secondary development, deriving from several different Common Turkic phonemes (cf. Anderson 1998a: 14; Johanson 1998a: 103f). Thus, Sakha /s/ developed from Turkic /č/, e.g. *saχ < *čaq* 'strike fire' (Tenišev 2001: 373) and *üs < *üč* 'three' (Sevortjan 1974: 641), as well as from Turkic /š/ (which probably did not occur in initial position; Johanson 1998a: 105), e.g. *bas < *baš* 'head' (Sevortjan 1978: 85ff). In intervocalic position, in Sakha both Turkic [č] and [š] ultimately developed into [h] via [s], following a general development of intervocalic [s] to [h] which has taken place in Sakha. Thus, Sakha *bīhaχ* 'knife' corresponds to Turkic *bičaq* (Johanson 1998a: 103) and *kihi* 'person, man' corresponds to Turkic *kiši* (Tenišev 2001: 325). Turkic word-initial [y] also developed into Sakha [s], as shown by such examples as *süreχ < *yürek* 'heart' and *suol < *yo:l* 'way' (Tenišev 2001: 276, 531). Ubrjatova (1960: 68) suggests that the change from [y] to [s] took place via a chain of changes, from [y] to [j] to [č] to [s], while Širobokova (2005: 171) suggests that [y] changed to [s] via [ç]. She bases this suggestion on the word pair *bīyïl* 'this year' and *ehi:l* 'next year' (cf. *sïl* 'year'), where the former appears to have retained the Turkic form *yïl* 'year', while the latter might represent a frozen form of the [ç] stage. Thus, three distinct but connected sound changes have taken place in the history of Sakha:

first of all, a change of initial [s] to Ø, possibly via an intermediate [h] (cf. section 3.1.5); secondly, a change of [č], [š] and initial [y] to [s], and lastly a change of [s] to [h] in intervocalic position.

Since the pharyngeal fricative *h* is found in a wide area of central and northeastern Siberia (Menges 1978: 254; Širobokova 1980: 145; Skribnik 2004: 156f), it has been suggested that the development of this sound may be due to substrate influence in the languages of this region (Širobokova 1980: 145; Skribnik 2004: 157, 159). Thus, for Sakha Skribnik (2004: 159) writes, summarizing the opinion of Nadeljaev (1986, cited in Skribnik 2004): "[…] the Turkic component of this language is assumed to be the language of the *Qurïqan* Turks, which in the Circumbaikal-Sprachbund was adopted by an ethnic group whose AAB [articulatory-acoustic base] amongst other features rendered initial /s/ as /h/." (translation mine[1]). Menges (1978: 254), Ubrjatova (1985a: 46), and Johanson (pers. comm.) postulate Northern Tungusic influence in the development of [s] to [h] in Sakha. In the following I will analyze the evidence for contact influence in the development of this feature.

3.1.1 The shift of intervocalic and word-initial [s] to [h]

The allophone [h], which is a variant of the secondary /s/ that in Sakha developed out of Turkic /č/ and /š/ and word-initial [y] (Tekin 1976: 113), is one of the features that distinguish Sakha from the other Turkic languages. It regularly occurs as a variant of /s/ in intervocalic position, not only through diachronic development as described above, but also through morphophonological alternations. Thus, word-final [s] changes to [h] if followed by a suffix beginning with a vowel, e.g. *ïs* 'plant[PRXIMP.2SG]' vs. *ïh–a:r* [plant–DSTIMP.2SG] 'plant later', *uos* 'lip' vs. *uoh–um* [lip–POSS.1SG] 'my lip'. Suffix-initial [s] changes to [h] if it follows on a stem ending in a vowel, e.g. *ïnaχ–sït* [cow–AGNR] 'cow-herder' vs. *sïlgï–hït* [horse–AGNR] 'horse-herder'. This alternation has been anchored in the modern orthography, with the Cyrillic letter *c* for [s] and the Latin letter *h* for [h], e.g. ба**с** *bas* 'head', ба**h**ым *bahïm* 'my head'.

However, the intervocalic change of [s] to [h] is not the only instance of this allophonic variation in the Sakha language. In word-initial position, too, [s] frequently changes to [h] in spoken speech, although the standard language

---

[1] Original: "[…] die türkische Komponente dieser Sprache solle die Sprache der *Qurïqan*-Türken sein, die im Circumbaikal-Sprachbund von einer Ethnie übernommen wurde, deren AAB unter anderem /s/ im Anlaut als /h/ wiedergab."

prescribes [s] in this position. This can partly be accounted for by fast speech phenomena: when following upon a word ending in a vowel, the environment of the sibilant is in practice intervocalic, e.g. *armiya–ɣa sulu:spa–lï:* [army–DAT service–VR.IPF.CVB] → *armiyaɣa **hu**luspalï:* 'serving in the army'. However, the intervocalic environment arising from fast speech is not the only factor affecting the use of word-initial [h]. Rather, this is a common feature of colloquial speech, and words such as the negative noun *huoχ* (often used with a meaning of 'no') or *höp* 'OK' are heard more frequently with [h] than with [s], even in isolation. There are, however, geographical differences in the frequency of word-initial [h], with a high frequency in the northern dialects, especially the extreme northwest and northeast, as well as in the Suntar district on the Vilyuy river (Voronkin 1999: 82; cf. Table 3.1 in section 3.1.2).

In Dolgan, [s] is restricted mostly to word-final position or adjacent to consonants, e.g. *ïpsï:* 'crack', *kï:s* 'girl'. In word-initial position and between vowels it generally changes to [h], e.g. *a**h**–a* [food–POSS.3SG] 'his/her food', ***h**ïmnagas* 'quiet, submissive' (Literary Sakha: *sïmnaɣas* 'soft') (Ubr: 31). This widespread occurrence of word-initial [h] in Dolgan is reflected in the orthography, which spells all such words with Latin *h* rather than Cyrillic *c*. This contrasts with Sakha, where people generally write *s* in initial position, even when transcribing spoken speech in which the word-initial [s] was replaced by [h]. The Dolgan alphabet even has a letter for capital *h* (an enlarged lower-case *h*, not the Latin capital *H*), which in Sakha is used only in printed headings that appear in block capitals and in the alphabet listed at the beginning of dictionaries, although occasionally a normal lower-case *h* is used here as well.

3.1.2 The distribution of word-initial [h] amongst different Sakha dialects

As mentioned in section 3.1.1, the occurrence of word-initial [h] in Sakha shows some geographical patterning (Voronkin 1999: 82). How striking the differences between different districts actually are, however, becomes truly apparent from my corpus, as shown in Table 3.1. I have here analysed a restricted number of sentences from the narratives I recorded from each speaker. For the speaker 'Efmy' from the Verxojansk district I analysed the first 51 sentences and then a further 45 sentences from the middle of the very long recording, since in the beginning she had attempted to speak standard Sakha (which is characterized by strict adherence to word-initial [s]), but had relaxed after a while. I therefore felt that the later sections of the recording might be a more accurate representation of her speech than the initial ones; the difference is obvious from the table. For the speaker 'Chir'

questions and interjections by two young women (his granddaughter and my primary consultant) were included in the analysis. The second column (school) lists the number of years of primary schooling the speakers have had; for those with 7 years of schooling, additional years of college-level training generally followed. The third column (N) gives the number of occurrences of the phoneme /s/ in word-initial position analysed for each speaker (between 41 and 79 per speaker and text), followed by the proportion of these occurrences in which word-initial [s] follows upon a word ending in a consonant (indicated as –C s– in the column heading) or in a vowel (–V s–); as well as the proportion of occurrences in which word-initial [h] follows upon a word ending in a consonant (–C h–) or in a vowel (–V h–).

Table 3.1: The proportion of word-initial [s] and [h] in informal Sakha speech

	school	N	–C s–	–V s–	–C h–	–V h–
Taatta district						
Bajaga: XatR	7 years	46	0.37	0.20	0.15	0.28
Bajaga: RaxA	7 years	62	0.48	0.18	0.16	0.18
Xarbalaax: MalA	7 years	56	0.23	0.09	0.36	0.32
Xarbalaax: BesP	0	56	0.04		0.43	0.53
Average (w/o BesP)		164	0.37	0.15	0.22	0.26
Average (with BesP)		220	**0.28**	**0.11**	0.28	0.33
Verxojansk district						
Efmy	3 years					
First half:		26	0.11	0.04	0.31	0.54
Second half:		36			0.5	0.5
Average:		62	0.05	0.01	0.42	0.52
Chir	0	46	0.15		0.28	0.57
PotP	0	55	0.13	0.02	0.43	0.42
Average:		163	**0.10**	**0.01**	0.39	0.50
Olenëk district						
LukP	7 years	59			0.44	0.56
MatX1		50	0.06		0.44	0.50
MatX2		51			0.45	0.55
Average:	7years	101	0.03		0.45	0.52
Afny	0	67	0.01		0.51	0.48
Average:		227	**0.02**		0.46	0.52

Suntar district						
P90		41	0.07		0.56	0.37
P95		54			0.54	0.46
Average:	0	95	0.03		0.55	0.42
IvaP	5 years	79	0.04		0.50	0.46
YmyE	4 years	48	0.04		0.73	0.23
Average:		222	**0.04**		0.57	0.39

As can be seen from Table 3.1, there is a clear difference in the frequency of word-initial [s] between the different districts (this difference is highly significant: χ^2 = 166.31, d.f. = 3, P = 7.9 x 10^{-36}): in the Taatta district (which belongs to the Central dialect group and has had a strong influence on the literary language) 39% of all word-initial /s/ is realized as [s], while in the Verxojansk district (belonging to the northeastern dialect group), we find only 11% [s] on average, as opposed to a mere 2-4% word-initial [s] in the Olenëk district (northwestern dialect group) and the Suntar district (Viljuj dialect group). It is also noticeable that in the latter two districts we find no instance of word-initial [s] following after a word ending in a vowel, and in the northeastern Verxojansk district, this is on average only 1%; in the central Taatta district, however, we find 11% [s] following after a vowel. The difference in frequency of word-initial [s] as opposed to [h] between the Olenëk and Suntar districts is non-significant (χ^2 = 1.46, d.f. = 1, P = 0.23), while the Taatta district is highly significantly different from all other districts, as is the Verxojansk district.

In the Taatta district, there is a difference between the two speakers recorded in the rather remote village Bajaga, with 57% and 66% word-initial [s], and the two speakers recorded on the Amga river. One of these pronounces word-initial [s] in 32% of cases, while the other shows a pronunciation falling into the range of the Suntar and Olenëk districts, with only 4% word-initial [s], none of which follows on words ending in a vowel[2]. At first glance, this might seem to be explainable by differences in education – the latter speaker had no schooling at all, while the others had had at least seven years of schooling. However, the speakers 'MatX' and 'LukP' from the Olenëk district also had seven years of schooling plus further training, and yet use hardly any word-initial [s], while the speakers 'Chir' and 'PotP' recorded in the village of Tabalaax in the Verxojansk district had no schooling at all, but do use

---

[2] However, it should be noted here that the Chi-Square contingency tests were performed on the total data from the Taatta district, including the speaker 'BesP', and all comparisons were still exceedingly highly significant.

word-initial [s] relatively frequently. Thus, differences in schooling do not make a difference in the realization of word-initial /s/ overall, and the difference between 'BesP' and her district compatriots is therefore more probably due to idiosyncratic factors.

These highly significant differences in the average frequency of word-initial [h] between different dialects are rather surprising and call for an explanation. Voronkin suggests that in the central dialectal area previous word-initial [h] has been replaced by the [s] prescribed by the standard language, while the archaic feature of word-initial [h] was retained in peripheral dialects: "Today it [the change of [s] to [h] in initial position] is restricted by the influence of the norm of the literary language, which regulates the writing of *s* at the beginning of a word; that is why in the central zone […] this process has already begun to diminish, being conserved to different degrees at the periphery […]" (Voronkin 1999: 82, translation mine[3]). This implies that the sound change had progressed equally far in all dialects before being reversed in some areas. However, since the change appears to have started very recently (c.f. section 3.1.3), it is hard to imagine that it would have progressed to a near complete shift of [s] to [h] in the central districts before the introduction of a literary standard in the early 1930s began to exert its prescriptive influence. Furthermore, since the literary language is based on the central dialects, it is not entirely clear why word-initial [s] would have been chosen to represent the standard when word-initial [h] was already wide-spread in these dialects. It rather appears to be the case that the change took place with different speed in different regions of Yakutia, and perhaps even started at different times in different areas. It is possible that contact or substrate influence played a role in the development of this change, as will be discussed below (cf. section 3.1.6).

3.1.3 The timing of the sound changes involving [s] in Sakha

From the description in the beginning of section 3.1 of the sound changes that have taken place in Sakha, it is clear that the loss of Common Turkic word-initial [s] must have taken place before Common Turkic [č], [š] and word-initial [y] had changed to [s], because otherwise the resulting word-initial [s] in Sakha would have been lost as well (cf. Poppe 1959: 678). Since word-initial [s] is retained in all Mongolic copies in Sakha, e.g. *sa:ra:* 'be doubtful, hesitate' < Literary Mongolian

---

[3] Original: "Ныне оно ограничено действием норм литературного языка, регламентирующих написание *с* в начале слова, поэтому в центральной зоне […] данный процесс уже пошел на убыль, консервируясь в разной степени на периферии […]".

saɣara 'be doubtful, hesitate' (Ubr: 35, Kałużynski 1962: 48, 28), it is clear that the loss of Common Turkic initial [s] must have taken place before the time of Sakha-Mongolic contact. Similarly, since Mongolic [č] is retained in copied lexical items in Sakha, e.g. *čaɣïlɣan* 'lightning' < Literary Mongolian *čakilɣan* 'lightning' (Kałużynski 1962: 45, 46), while it changed to [s] in Turkic roots, the sound change of Common Turkic [č] to [s] must also predate the time of contact with Mongolic. However, in Sakha Mongolic [š] has regularly changed to [s], and to [h] in intervocalic position, e.g. *sijiŋ* 'bed-wetter' < Mongolic *šijiŋ* 'bed-wetter', *tahïy* 'hit' < Mongolic *taši* 'hit with the flat of the hand' (Kałużynski 1962: 49). This implies that the change of Turkic [š] to [s] and [h] in Sakha took place later than the time of Mongolic contact[4]. Mongolic word-initial [y] regularly changes to [j] in Sakha, as does Russian [y] in copies, e.g. *jaŋ* 'illness, epidemic' < Literary Mongolian *yaŋ yara* 'sexually transmitted disease', *ja:t* 'poison' < Russian *yad* 'poison' (Kałużynski 1962: 50). Thus, the change of Turkic word-initial [y] to [s] in Sakha must have predated Mongolic contact. This is estimated to have taken place from the 12th or 13th century until the 15th century or possibly even later (Kałużynski 1962: 122), so that the loss of the Common Turkic initial [s] and the change of [č] and word-initial [y] to [s] must have taken place before the 12th or 13th century. Intervocalic [s] in Mongolic copies has, however, undergone the sound change to [h], e.g. *nüher* 'big, solid, weighty' < Literary Mongolian *nöser* 'sluggish, plump, too heavy' (Kałużynski 1962: 49). Thus, this sound change must have taken place later than the 12th to 15th century. The change from [š] to [s] must have taken place after contact with Mongols, but before the change of intervocalic [s] to [h], since intervocalic [s] derived from Mongolic [š] changed to [h], as described above.

   The shift of intervocalic [s] to [h] can be shown to have happened fairly recently, since most of the Siberian explorers who collected word lists in the 17th and 18th century still wrote down words that have an intervocalic /s/ with an *s*:

> "The available word lists show *s* with hardly any exceptions, and also the Russians in Yakutsk and even Uvarovskij, who always pronounces *h* in these cases, write *s*. The Yakuts who Middendorff met noted that even when they said *kihi* etc, it should actually be *kisi*. The weakening of *s* to *h* can be traced to the end of the previous century, since Sauer (p.7) lists *buhak* in addition to

---

[4] Actually, since Mongolic [š] is derived from [s] before [i], one could also postulate a retention of Mongolic [s] in these words. However, since the sound change of [s] to [š] had already taken place in Middle Mongolian (Kałużynski 1962: 49), which coincides with the time of Sakha-Mongolic contact, one can plausibly assume a change of [š] to [s] in Mongolic copies in Sakha.

busak (our *bïsax*); Witsen always writes *s*." (Btl: 158, footnote 79; translation, italics and Latin transcription mine[5]).

Nicolaes Witsen was a Dutch diplomat who visited Moscow from 1664 to 1665; after his return to Holland, several items from Siberia were sent to him from Russia, and he published an account of Russia and Siberia based on these materials in 1692 (Nikolaeva 2006: 10f). Thus, Witsen's word list used by Böhtlingk dates from the second half of the 17th century, and, as noted by Böhtlingk, all cases of intervocalic /s/ are written with an *s*, e.g. *besen* 'fifty' (modern-day *bies uon*), *bisigi* 'we' (modern spelling *bihigi*) (Witsen [1692] 1705: 677). The next word list was collected by von Strahlenberg, who spent 12 or 13 years in Siberia as a prisoner of war after the Battle of Poltava in 1709 until he left Russia in 1723 (Krueger [1962] 1997: 305). From the word list he published (von Strahlenberg 1730, as reprinted by Krueger [1962] 1997) it appears that in the first half of the 18th century the shift of intervocalic [s] to [h] had not yet taken place; thus, von Strahlenberg gives *kisill* for modern-day *kïhïl* 'red' and *ir cksi* for modern-day *er kihi* 'man' (Krueger [1962] 1997: 309). Martin Sauer wrote the description (first published in 1802) of the expedition to northeastern Siberia which took place from 1785-1794 under the command of Joseph Billings. It should be noted that (at least in the German translation of this book) all words with intervocalic /s/ are written with an *s*, e.g. *kasil* 'red' (modern-day *kïhïl*) and *ki-esse* 'evening' (modern-day *kiehe*) (Sauer 1803: 320, 321). The only exception in this list of 271 Sakha words, of which 31 have an intervocalic /s/, is the word for 'knife' mentioned by Böhtlingk in the quote above, where Sauer gives both *busak* and *buhak* (Sauer 1803: 323). Thus it would seem that the change of intervocalic [s] to [h] had only just begun in the late 18th century. Alexander Theodor von Middendorff's expedition took place from 1842-1844 (Severnaja Ènciklopedija 2004: 562), and Uvarovskij worked with Böhtlingk in the middle of the 19th century (from 1847 (Btl: LIII) until 1851 at the latest, when Böhtlingk's grammar was published). As can be seen from the above quotation, Uvarovskij consistently pronounced intervocalic /s/ as [h]. It is therefore clear that the shift of intervocalic [s] to [h] took place very rapidly in the first half of the 19th century.

As to the change of word-initial [s] to [h], in Witsen's wordlist (Witsen 1705: 677) all words with initial /s/ are spelt with an *s*. Similarly, von Strahlenberg writes

---

[5] Original: "Die vorhandenen Wörtersammlungen weisen fast ohne Ausnahme c auf und auch die Russen in Jakutsk und sogar Uwarowskij, der in diesem Falle immer h spricht, schreiben c. Die Jakuten, mit denen Middendorff zusammenkam, bemerkten, wenn sie auch ҟihi u.s.w. sprachen, dass es eigentlich ҟici heissen müsste. Die *Schwächung* des c zu h lässt sich bis an's Ende des vorigen Jahrhunderts verfolgen, da Sauer [...] neben *busak* (unser бысах) auch *buhak* aufführt; Witsen schreibt immer s." (italics original)

the words for 'hundred' and 'earth' as *suss* and *sirr*, respectively, i.e. with an [s] (Krueger [1962] 1997: 307, 308), and in Sauer's word list (Sauer: 1803: 317-324) we find 34 words with an initial /s/, all of which are spelt with an *s*, such as *serai* 'face' (*sirey* in the modern spelling) and *sacin, soyin* 'summer' (modern-day *sayïn*). Böhtlingk, too, lists words only with word-initial *s*, as does Pekarskij ([1925-1927] 1959), with the exception of *suoχ* 'no', for which he gives the variant *huoχ* (Pekarskij [1926] 1959: 2354). Although Pekarskij's dictionary was published at the beginning of the 20ᵗʰ century (between 1907 and 1930), he collected the data for it during his period of exile in the Taatta district from 1881 to 1899. Although both Böhtlingk and Pekarskij write intervocalic [s] as *s* (i.e. they do not mark the allophonic variation in their orthography), Pekarskij at least is very conscientious about listing phonetic variants of words, and judging from his inclusion of the variant *huoχ* for *suoχ*, he would probably have noted other variants with word-initial [h] as well. From this it would appear that the change of word-initial [s] to [h] took place only in the 20ᵗʰ century, at least in the central districts.

According to Böhtlingk (Btl: 158), the change of intervocalic [s] to [h] started in Yakutsk: "Since not very long ago, *s* is spoken by many Yakuts, namely by the inhabitants of Yakutsk, as *h* between two vowels." (Btl: 158, translation and italics mine[6]). However, if it really originated in the speech of Yakutsk inhabitants, it must have spread quite rapidly to other areas as well; thus, Maak, who visited the Viljuj district (*Viljujskij okrug*) in 1854, gives variants for the words *basïba*/*paxïba* 'thank you' (modern-day *bahï:ba*) and *sasïl*/*saxïl* 'fox' (modern-day *sahïl*) (Maak [1883-1887] 1994: 299, 317). In a section on Sakha riddles, he writes *kixi* 'person, man' (i.e. modern-day *kihi*) in three riddles, but in two other riddles writes *kisi*. For the hearsay marker (modern-day *ühü*) he always writes *isju*; in one riddle he writes *abasï kïs–a* [devil girl–POSS.3SG] 'devil's daughter', in another he writes *abaxï kïx–ïn* [devil daughter–GEN]. Most of the morphologically conditioned changes of [s] to [h] he writes with a *x*, e.g. *tjux–er* (modern-day *tüh–er*) [fall–PRSPT] 's/he falls'. Similarly, the change of [s] to [h] across word boundaries appears to have been taking place in the Viljuj dialect at that time, since Maak writes *bexon* for *bies uon* [five ten] 'fifty' (Maak [1883-1887] 1994: 307ff). However, it appears that at the time of Maak's expedition, [s] was still quite frequently pronounced in intervocalic position, while all the words with an initial /s/ are spelt with an *s*.

The following chronology for the sound changes affecting Sakha sibilants can therefore be postulated: Turkic [s] in word-initial position was lost first of all, followed by the change of Turkic [č] and word-initial [y] to [s]. This took place

---

[6] Original: "C wird seit nicht gar langer Zeit von vielen Jakuten, namentlich von den Bewohnern der Stadt Jakutsk, zwischen zwei Vocalen wie h gesprochen."

before the 12th or 13th century. Subsequently, after contact with Mongols, Turkic and Mongolic [š] changed to [s]; in intervocalic position, this and the [s] derived from [č] and [y] began to change to [h] in the late 18th century, a change that was completed by the middle of the 19th century. The change of word-initial [s] to [h] is the most recent, taking place in the central districts in the 20th century, and has not been completed yet in Sakha, though it has in Dolgan.

3.1.4 [h] in the languages of Siberia

As mentioned above, the presence of the pharyngeal fricative *h* is considered by some researchers as an areal feature of Siberian languages (Menges 1978: 254; Širobokova 1980: 145; Skribnik 2004: 156f). In most Turkic languages, [h] does not occur in word-initial position in originally Turkic words (Ščerbak 1994: 45). Amongst the South Siberian Turkic languages, Khakas and Tuvan lack a pharyngeal fricative [h] (And: 3, A/H: 6), although according to Ščerbak (1994: 45) this sound does occur in more than 30 lexical items in Tuvan dialects. Tofa, on the other hand, has a phoneme /h/ that can occur both word-initially and medially, but not in word-final position. In intervocalic position [h] is an allophone of /q/ after a pharyngealized vowel (Rassadin 1971: 59f). Bashkir, too, has a phoneme /h/; this has developed from Common Turkic /s/ in word- and syllable-initial position, e.g. Tatar *sal* vs. Bashkir *hal* 'raft'. The phoneme /s/ in Bashkir developed from Common Turkic /č/ via /c/, e.g. *aɣas* < *aɣač* 'tree' (Serebrennikov 1973: 10; GSBJ: 54, 60).

A pharyngeal fricative /h/ is lacking in the Mongolic languages, with the exception of Buryat (Sanžeev 1953: 88, 90; Skribnik 2004: 156f). In Buryat, /h/ developed from Mongolic /s/, e.g. *han/g* < *sang* 'store house', while modern Buryat /s/ developed from Mongolic /c/, e.g. *san/g* < *cang* 'cymbal(s)' (Sanžeev 1953: 87f; Skr: 106).

The Amur Tungusic languages Udihe, Oroč, and Nanay lack a phoneme /h/ (N/T: 51; AvrI: 37; A/B: 24). In Udihe, [h] only occurs in ideophones and interjections (N/T: 51). Manchu, too, lacks a pharyngeal fricative *h* (Gor: 85f). In the Northern Tungusic languages Evenki and Ėven, however, [h] occurs as an allophone of /s/ in different position in different dialects. The classification of Evenki dialects is based on the distribution of [s] vs. [h] in word-initial and intervocalic position, while the classification of Ėven dialects is based partly on this feature (cf. section 1.1.2.1 and section 3.1.6.).

Kolyma Yukaghir lacks both /s/ and /h/ (Mas: 30), while Tundra Yukaghir has a phoneme /s/ (Maslova 2003a: 3; Krejnovič 1958: 10). Nganasan has two

phonemes /s/ and /h/, both of which occur in word-initial and in intervocalic position (Helimski 1998: 484). The Nganasan word-initial [h] is derived from Proto-Samoyedic *p–, while Nganasan word-initial /s/ is a retention of Proto-Samoyedic *s–. In Forest Nenets, however, Proto-Samoyedic *s– developed to [h] (Sammallahti 1988: 497f). Selkup, Mansi, and Khanty have a voiceless alveolar sibilant /s/, but no pharyngeal fricative /h/ (OSJ: 120; Rse: 14; Nik: 6). Itelmen has a voiceless sibilant /s/ and a voiceless uvular fricative, which in word-initial position is realized as a pharyngeal fricative [h] (G/V: 18, 23f). Chukchi lacks a /h/, and the voiceless sibilant /s/ was originally distinctive of men's speech only (Dnn: 43, 59). Siberian Inupik Eskimo has a phoneme /s/, but no /h/ (Men: 27). Nivkh has both a /s/ and a /h/ (Grz: 10; Mat: 36). Ket has a voiceless sibilant /s/ that occurs in all positions and a voiceless pharyngeal fricative /h/ that is, however, restricted to the word-initial position before vowels (Dul: 49; Wer: 19f; Vajda 2004: 5).

We thus see that the distribution of the pharyngeal fricative *h* is fairly widespread amongst Siberian languages, being found in at least one language of most language families. In the Turkic language family, it is present in Tofa, Sakha, Dolgan, and Bashkir, while in the Mongolic family it is restricted to Buryat. Evenki and Ėven are among the Tungusic languages that have this feature, while Nganasan and Forest Nenets are the Samoyedic languages that have it; as does the Chukotko-Kamchatkan language Itelmen and the isolates Nivkh and Ket. Bashkir, however, does not belong to the Siberian languages, since it is spoken to the southwest of the Ural Mountains.

The pharyngeal fricative is of varied origins in the Siberian languages that have it. Thus, in Tofa, [h] in initial position is derived from the fortis aspirated uvular stop [kʰ] (Tenišev 2002: 625, 628), while in Nganasan, [h] appears to derive from Proto-Samoyedic *p, and in Itelmen it is an allophone of the uvular fricative. In Nivkh, medial [h] derives from [s], which in turn derives from palatalized [t'] and [č]; synchronically, [t'] and [č] alternate with [s] (Frederik Kortlandt, pers. comm.; Mat: 44f). In Forest Nenets, Sakha, Dolgan, Buryat, Evenki and Ėven [h] is either historically derived from or is in current allophonic variation with [s]. This distribution is strongly reminiscent of contact, since none of these languages inherited the pharyngeal fricative from their proto-language, and Evenki is in contact with all the other languages except Forest Nenets. However, although one may postulate Evenki or Northern Tungusic contact or substrate influence in Buryat, Sakha and Dolgan (cf. Menges 1978: 254; Ubrjatova 1985a: 46; Lars Johanson, pers. comm), this does not solve the question of the ultimate origin of this sound in this area, as Evenki did not inherit it either. Some postulate an unknown substrate (cf. Skribnik 2004: 145; Janhunen 1996: 172); but the languages that might plausibly

have furnished this, Samoyedic and Yukaghir (cf. Fortescue 1998: maps 5, 6), lack a pharyngeal fricative as well; the development of [h] from [s] in Forest Nenets is an individual feature of this language and not characteristic of the family (Janhunen 1998a: 462; Mas: 30; Krejnovič 1958: 10). Another option is that the substrate may have been a Yenisseic language. Proto-Yenisseic is reconstructed as having had a pharyngeal fricative at least in word-initial position (Werner 2005: 222f), and it is assumed that the Yenisseic languages were previously spoken in the area of the Altay-Sayan mountains (Alekseenko 1980: 127; Fortescue 1998: map 6). This makes a Yenisseic substrate plausible for at least some of the languages of Siberia that have a pharyngeal fricative. On the other hand, the possibility that the development of [h] from [s] may have been a language-internal process must also be taken into account, especially since the change of [s] to [h] is not uncommon (cf. section 3.1.5).

3.1.5 The change of [s] to [h] in cross-linguistic perspective

Of the sound changes that affect [s], the shift from [s] to [h] (and further to Ø) is cross-linguistically the most commonly attested one, found for example in Ancient Greek, Armenian, Celtic and Indo-Iranian languages, in Austronesian languages, in African languages, as well as in Finnish, Hungarian, Korean, Buryat, Evenki and Èven (Serebrennikov 1973: 12f; Merlingen 1977: 192ff; Ferguson 1990: 63; Frederik Kortlandt, pers. comm.). Based on the sequence of shifts, Ferguson (1990: 65ff) distinguishes two types. The first pathway of change, which he terms the 'Greek type', starts with a shift of [s] to [h] in intervocalic position, proceeds to word-initial [s], and affects preconsonantal and (possibly) word-final [s] only in the final stages. In contrast, in the second 'Spanish' type, the shift starts in syllable-final positions, and affects word-initial [s] only in the last step. It is clear that the change from [s] to [h] that has been taking place in Sakha since the late 18th century belongs to the Greek type, since the change started in intervocalic position (Ferguson 1990: 67f).

Since the change from [s] to [h], and further to complete loss, is so common, it is quite plausible to assume that the loss of Common Turkic word-initial [s] in Sakha took place via [h], even though no traces of the intermediate stage remain. This could have been a purely language-internal development, since this sound change is quite natural; however, given the distribution of [h] in the languages of eastern Siberia (cf. section 3.1.4), one could also postulate some substratal influence, as done by Širobokova (1980: 145) and Skribnik (2004: 157, 159). Similarly, for the recent and ongoing shift of [s] to [h] one might again assume a purely language-

internal development. However, the very striking differences in pronunciation of [s] in word-initial position between different dialectal areas imply that more than just natural, language-internal processes may be at work, as will be discussed below.

3.1.6 Possible contact influence on the shift of [s] to [h] in Sakha

Several researchers have proposed that substrate influence led to the change of [s] to [h] in Sakha. One of the candidates for influencing this change are the Northern Tungusic languages, as suggested by Menges (1978: 254), Voronkin (1999: 81), and Ubrjatova (1985a: 46), while others have suggested a historical connection with Buryat, possibly through a shared substrate (Afanas'ev 1965: 71; Širobokova 1980: 145). As mentioned above (section 3.1.4), Buryat, too, has undergone a shift of [s] to [h] as well as a parallel change of [c] to [s] and [š] before vowels other than *i (Skr: 106). Since the Sakha ancestors are presumed to have come from the area around Lake Baykal that is nowadays settled by Buryats, the influence of a shared substrate in these two languages is quite plausible. This shared substrate has been suggested as a dialect of Evenki (Buraev & Šagdarov 2004: 230; Lars Johanson, pers. comm.), while Janhunen (1996: 172) proposes that the shift in Sakha, Buryat and Evenki is due to a common substrate in all three languages, rather than to Evenki influence in Sakha and Buryat. Buryat itself is ruled out as a source of contact influence in Sakha, since the change in Buryat can be dated to the turn of the 17th/18th centuries (Širobokova 1980: 145), a time when the speakers of both languages were already separated. Although a shared Circumbaykal substrate (for example a Yenisseic language, cf. section 3.1.4) may explain the loss of Common Turkic word-initial [s] in Sakha, and possibly the change of [č] to [s], it cannot have been involved in the shift of intervocalic and word-initial [s] to [h] observable nowadays. The historical word lists show that this change began only after the end of the 18th century and even later, when the Sakha had long left the Baykal region and were already spread over Yakutia. This change may, however, have developed under Tungusic influence, as will be discussed here.

The Evenki language is split into 51 dialects [Ned: xix; B/G: 3), which are grouped into three big dialect groups based on phonological and geographical grounds (cf. section 1.1.2.1): The northern dialect group is spoken in the Evenk National District, around Tura (on the Nizhnjaja Tunguska river), Chirinda (close to Lake Essej), and Ekonda (on the source of the Viljuj river). This dialect group is characterized by both word-initial and intervocalic [h]. The southern dialect group is spoken near Lake Baykal, in the villages Vanavara (on the Podkamennaja

Tunguska), Bajkit and Poligus (on the river Chunja, which merges with the Podkamennaja Tunguska). There are two subgroups of this dialect; one is characterized by [s] both in word-initial and intervocalic position, while the other pronounces /s/ as [š]. The eastern dialect group is spoken in the Far East, Vladivostok and Khabarovsk regions as well as in the south of Yakutia. This features word-initial [s] and intervocalic [h] (Sunik 1962: 22; Ned: xixf; B/G: 3; Atknine 1997: 117).

As shown in Figure 1.2, the Sakha were concentrated along the confluence of the Lena, Aldan and Amga rivers at the time of first Russian contact. If the present-day distribution of Evenki dialects is comparable to that several centuries ago, the Sakha ancestors would have first come into contact with Evenki groups speaking dialects similar to the Vitimo-Olëkminsk and Aldan dialects. These belong to the eastern dialect group (Romanova 1962: 8), which are characterized by word-initial [s] and intervocalic [h]. These are also the Evenk groups the central Sakha would have stayed in contact with throughout the centuries, and this is the sound pattern favoured by the Sakha standard language, which is reflected in the speech of most of my informants from the Taatta district. The Sakha migrating to the west and northwest (e.g. to the Viljuj and Olenëk rivers) would have come into contact with Evenk groups belonging to the northern dialect group characterized by both word-initial and intervocalic [h]. Those Sakha migrating to the northeast, e.g. towards the Yana river, would have come into contact with Èvens and Yukaghirs. All Èven dialects on the Eurasian mainland have only [h] in initial position (with the exception of copied lexemes); the pronunciation of /s/ in medial and final position, however, distinguishes the dialects spoken in Yakutia from those spoken further east. The Èven dialects spoken in Yakutia are characterized by medial and final [h], while the eastern dialects have [s], e.g. western and central Èven *ahi* vs. eastern *asi* 'woman' (NovI: 18, 72, 73 footnote 59).

The data presented in section 3.1.2 on the realization of /s/ in word-initial position in informal spoken Sakha show a clear correlation with the Evenki dialect group originally spoken in the area: the centrally located Taatta district lies in an area that was formerly inhabited by Evenki groups probably belonging to the eastern dialect group, such as the Aldan Evenki dialect – this Evenki dialect group is characterized by word-initial [s] and intervocalic [h], which is what we find in the Sakha dialect spoken in this area as well. The Suntar and Olenëk districts, which are located in the west and northwest of Yakutia, are adjacent to areas inhabited by Evenks speaking northern Evenki dialects, and were probably inhabited by Evenks until the 18th or 19th century. These dialects are characterized by both word-initial and intervocalic [h], just like the Sakha dialects spoken in this region. However, the

correlation between the Sakha dialect and the presumed source of influence does not hold for the northeastern dialect, as reflected by the recordings I made in the Verxojansk district. The western Èven dialects, which are spoken in the area to the west and north of the Verxojansk district, are characterized by [h] in initial, medial, and final position (NovI: 73, footnote 59). One would thus expect the Sakha dialect spoken in contact with these Èven dialects to feature a frequency of word-initial [h] similar to that in the Suntar and Olenëk districts; and yet, the average frequency of word-initial [s] here is 11%, much higher than would have been expected[7]. This is even more surprising given the fact that Sakha were settled in this area at the time of first Russian contact, longer than Sakha have been settled on the Viljuj river and in the northwest.

Nevertheless, the findings for the Taatta, Suntar and Olenëk districts appear to lead to the conclusion that even though the shift from [s] to [h] is cross-linguistically frequent, and may therefore have happened independently of any contact influence on the Sakha language after their migration to the Lena river, it was a process that was at the very least reinforced by contact with the Evenki dialects spoken in the areas of settlement, if it was not actually initiated by this contact. However, there are some problems with this conclusion, as will be discussed below.

3.1.7 Problems with the Tungusic contact scenario

There are two major problems with the scenario sketched above, namely that contact with Evenks speaking different dialects of Evenki led to the change, first of intervocalic, and later of word-initial, [s] to [h]. The first problem concerns the distribution of the Evenki dialects. The above scenario relies on the modern-day distribution of these dialects, which fit the Sakha data very well. However, Evenks were a highly mobile people, and it is rather unlikely that their current distribution has remained unchanged over centuries. For instance, the Manchurian Reindeer Tungus are known to have migrated from Yakutia across the Amur river to China shortly before 1830 under pressure of the expanding Sakha (Janhunen 1996: 110). Dolgix (1960: 620) summarizes the immense population movements that took place after Russian colonization by saying that the Evenks nowadays living on the Olenëk originated on the Aldan river, while the Evenks now living in the Ob river basin and in the Tomsk district came from the Angara river; similarly, the Evenks from the

---

[7] However, it should be noted that the speakers recorded in the Verxojansk district showed a significantly lower frequency of word-initial [s] than those recorded in the Taatta district.

Angara migrated to the Podkamennaja Tunguska. Vasilevič (1969: 25) claims that in the mid-19th century the Evenks living in the Viljuj river basin were immigrants from the Angara river. If these scanty indications of quite substantial changes in the distribution of Evenk clans and subtribes are true, then it may to be rather unwise to base the explanation of contact influencing the change from [s] to [h] on the current distribution of Evenki dialects.

The other problem concerns the very recent origin of the shift from [s] to [h] in Sakha. In intervocalic position the first indication of this is the alternative spelling of the word for 'knife' as *busak* or *buhak* in Sauer's word list of the late 18th century, as mentioned in section 3.1.3; by the mid-19th century this feature is quite regular in the speech of Sakha living in Yakutsk, but still irregular in the dialect of the Sakha living on the Vilyuy river. But if the sound change were truly due to contact with speakers of Evenki (or to Evenki substrate), one would expect it to have happened much earlier, after the immigration of Sakha speakers to the Lena-Amga-Aldan area, and not some 300 or 400 years after the migration. The same holds for the change of word-initial [s] to [h], which appears to fit nicely with contact with Evenks in the west and northwest of Yakutia. However, the only mention of this change in the historical literature is in Böhtlingk (Btl: 159), who says that the negative noun *suoχ* is pronounced *huoχ* even when not following upon a word ending in a vowel; similarly, Pekarskij ([1926] 1959: 2354), who collected the data for his dictionary in the late 19th century, gives *huoχ* as an alternative for *suoχ*. According to the scenario sketched in section 3.1.6, the shift of word-initial [s] to [h] should have started earlier in the western and northwestern districts of Yakutia than in the central districts; however, there is no indication of this in the data published by Maak, even though Sakha speakers migrated to the Viljuj river in the second half of the 17th century (Dolgix 1960: 470f) – 150 to 200 years before Maak documented their language (cf. section 3.1.3). An additional argument against Evenki substrate influence in the development of this feature in Sakha is the lack of conclusive genetic evidence for this (Pakendorf et al. 2006). Although some admixture in the maternal line cannot be ruled out, a shift of whole groups of Evenks (i.e. of men as well as women) to the Sakha language and ultimately Sakha ethnic identity is not discernible.

It is nevertheless true that the distribution of a change of [s] to [h] in the languages of Siberia is intriguing and appears to be strongly indicative of some form of substrate or contact influence, as has been repeatedly claimed in the literature. However, as has been argued here, the currently observable changes of [s] to [h] in Sakha and Buryat are hardly reconcilable with a common substrate, since the changes took place very late in both languages. The change that led to the loss of

Turkic word-initial [s] in Sakha, on the other hand, might have been the result of substrate or contact influence. However, that is very hard to prove nowadays, since the change was completed long before any documentation of Sakha occurred.

Given the facts that the distribution of Evenki dialects in historical times may well have been very different from that nowadays, that the sound change from [s] to [h] happened so recently, and that there is no genetic evidence of an Evenki substrate in Sakha, contact or substrate influence may not be the correct explanation for this feature, after all. Since the change from [s] to [h] is cross-linguistically common, an internally motivated sound change may well be the best explanation. The rapid spread of the intervocalic shift of [s] to [h] may have been aided by social prestige, since according to Böhtlingk this change originated amongst the inhabitants of Yakutsk, the dominant town in northeastern Siberia at that time. That linguistic innovations can spread quite widely and quite rapidly when they are associated with prestigious lects has been shown previously, for instance with the spread of the uvular /r/ through France and parts of Europe after it originated in Paris (Chambers & Trudgill 1998: 170f; Ross 1997: 231f). This initial weakening of [s] to [h] in intervocalic position may have in turn led to further weakening of word-initial [s] in some dialect areas. That this sequence of change is quite natural is demonstrated by Ferguson (1990), who finds that it has happened in other languages, too.

However, although for Sakha an internally motivated change from [s] to [h] appears rather more probable than a change induced by contact with speakers of Evenki, the same cannot necessarily be said for Dolgan. As mentioned in section 3.1.1, not only the shift of intervocalic [s] to [h], but even the shift of word-initial [s] to [h] has been completed in Dolgan, and has even been included in the standard orthography. It is assumed that the initial stages of the formation of the Dolgans as a separate ethnolinguistic group began in the 18th century, and were completed in the 19th century (*Severnaja Ėnciklopedija* 2004: 257). Thus, the Sakha speakers whose language was later adopted as the language of the newly formed ethnic group must have left the immediate vicinity of the Sakha proper before the change of [s] to [h] was initiated there. Since Dolgans are thought to be of mixed origins, with Evenks constituting the main core of the population (Gracheva 1994: 99f, *Severnaja Ėnciklopedija* 2004: 257), Evenki substrate influence in the development of [s] to [h] in Dolgan cannot be excluded.

3.2 Divergent features of the Sakha case system

The case complement of Sakha differs from that found in other Turkic languages. In general, Turkic languages distinguish six cases: the unmarked Nominative and the morphologically marked Genitive, Dative, Accusative, Locative, and Ablative (Johanson 1998a: 39). In contrast, eight cases are recognized for Sakha (cf. Table 3.2): the zero-marked Nominative, plus the suffixally marked Accusative, Dative, Instrumental, Ablative, Partitive, Comitative and Comparative. As can be seen from Table 3.2, Sakha differs from the other Turkic languages by having lost the Genitive and Locative case, by having innovated the Partitive and Comparative case, and by having retained a distinction between Instrumental and Comitative. In the following, I will undertake to analyze whether these differences may be due to contact influence in Sakha.

Table 3.2: Comparison of common case forms across some Turkic languages[1]

Case	Old Turk.	Sakha	Tuv.	Kaz.	Uzb.	Turk.	Khal.[*]	Chuv.[**]
Nom	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø
Gen	*-(n)Xŋ* *-nXŋ*	——	*-NIŋ*	*-NIŋ*	*-nI(ŋ)*	*-(n)In*	*-ü:n* *-üy* *-i:(n)*	*-(ə)n/* *-yən/* *-nən*
Dat	*-kA*	*-GA*	*-KA*	*-GA*	*-GA*	*-(y)A*	*-KA*	*-(n)A/*
Acc	*-(X)g* *-nI*	*-(n)I*	*-NI*	*-NI*	*-nI*	*-(y)I*	*-i* *-U*	*-(y)A*
Part	——	*-TA*	——	——	——	——	——	——
Loc	*-dA*	——	*-DA*	*-DA*	*-dA*	*-DA*	*-čA*	*-RA*
Abl	*-dIn* *-dAn*	*-(t)tAn*	*-DAn*	*-DAn*	*-dAn*	*-DAn*	*-dA* *-dAn*	*-RAn*
Instr	*-(X)n* *-(I)n*	*-(I)nAn*	=bile	-ben/ -men	bilan	=(y)lA	*-lA(n)*	*-pA*
Comit	*-lXgU* *-lUgUn*	*-LI:n*						
Comp	——	*-TA:γAr*	——	——	——	——	——	——

Abbreviations of language names: Old Turk. = Old Turkic; Tuv. = Tuvan; Kaz. = Kazakh; Uzb. = Uzbek; Turk. = Turkish; Khal. = Khalaj; Chuv. = Chuvash

[*]In Khalaj, the different allomorphs represent dialectal differences.
[**]In Chuvash, the allomorphs are phonologically conditioned.

---

[1] This table shows the simple (non-possessive) declension; for the Sakha case forms of the possessive declension see Appendix 3. Here, only those cases which are relevant to the discussion of contact influence in Sakha are summarized. Some languages, e.g. Old Turkic and Chuvash, can be analyzed as having more cases than those listed here.

3.2.1 The loss of the Genitive in Sakha

One of the salient differences of Sakha from other Turkic languages is the loss of the Common Turkic Genitive case, a difference that has been suggested as a result of Tungusic influence in Sakha (Schönig [1988] 1990: 50, footnote 43; Stefan Georg pers. comm.).

3.2.1.1 The genitive case in Turkic languages

Old Turkic has a Genitive case marked by the suffix *–(n)Xŋ* in the early runic inscriptions, and *–nXŋ* in later texts (Erd: 168ff). It marks the possessor of a possessive noun phrase, while the possessum agrees in person and number with the possessor, e.g. ***meniŋ*** *yutuz–um* [1SG.GEN wife–POSS.1SG] 'my wife' (Erd: 381). However, in a number of cases, the possessor does not carry the Genitive suffix, e.g. *teŋri yer–i* [god land–POSS.3SG] 'divine land' (Erd: 382). According to Doerfer, Nominative case-marked modifiers in possessive constructions are very frequent in Old Turkic: "The Nominative as modifier is exceedingly frequent in [Khalaj], about as frequent as in [Old Turkic]" (Drf: 76, my translation[2]). This usage is similar to the unmarked possessor in abstract relational noun phrases in several modern Turkic languages, as exemplified below. However, it does not appear to be the case that Genitive-marked possessors indicate specificity of the possessor, while unmarked possessors are generic, as is the case in current-day Turkic languages that make this distinction (Erd: 381f). The Genitive can also be used to mark the subject of relative clauses, e.g.

(1)
   meniŋ   *kol–mïš*   *kut–um*   *büt–me–zün*
   1SG.GEN   ask.for–PSTPT   blessing–POSS.1SG   materialize–NEG–IMP.3SG
   'May the blessing for which I prayed not materialize (if…)'

                                                          [Erd: 440]

As can be seen in Table 3.2, the Turkic case forms, including the Genitive suffix, have been quite stable throughout the history of the language family. In modern Turkic languages the Genitive has the forms *–nIn*, *–nI*, *–(n)Iŋ* and *–(n)In* (Johanson 1998). Like in Old Turkic, the Turkish Genitive suffix *–(n)In* (Krn: 212) is also used to mark the possessor of a possessive noun phrase, while the possessum agrees in number and person with the possessor e.g. *Hasan–**ın** kitab–ı* [H.–GEN

---

[2] Original: "Der Nominativ als Attribut ist im C. ungemein häufig, etwa wie im Alttü."

book–POSS.3SG] 'Hasan's book' (Krn: 230); *ben–**im** ev–im* [1SG–GEN house–POSS.1SG] 'my house' (G/K: 183). The Genitive also marks the subject of certain subordinate clauses, e.g.

(2)

*Turgut–**un***	*gel–me–sin–i*	*isti–yor–um*
T.–GEN	come–NR–POSS.3SG–ACC	want–AOR–1SG

'I want Turgut to come.'

[G/K: 182]

The Kazakh Genitive suffix is *–NIŋ*, and it, too, is used to mark the possessor in a possessive noun phrase, with the noun denoting the possessum taking suffixes agreeing in person and number with the possessor, e.g. ***meniŋ** kitab–ïm* [1SG.GEN book–POSS.1SG] 'my book'. However, the Genitive suffix is dropped in cases of abstract, generic possession and part-whole relations, e.g. *qoy et–i* [sheep meat–POSS.3SG] 'mutton' (SKJ: 161).

The Uzbek Genitive case is marked by the suffix *–nI(ŋ)*, and it is similarly used to mark the possessor in a possessive noun phrase, in which the noun referring to the possessum takes possessive suffixes agreeing with the possessor, e.g. *Žobirov–**niŋ** odam–i* [J.–GEN man–POSS.3SG] 'Jobirov's man' (Bdr: 75, 76). As in Kazakh, an 'unmarked Genitive' is used for abstract possessors and part-whole relations, e.g. *boš yorig–i* [head wound–POSS.3SG] 'head wound' (Bdr: 77).

In Tuvan and Khakas, the Genitive case suffix has the form *–NIŋ*, and it is likewise used to mark the possessor in possessive noun phrases, as exemplified by Tuvan *baškï–**nïŋ** bažïŋ–ï* [teacher–GEN house–POSS.3SG] 'the teacher's house' (A/H: 20). The possessum is marked for person and number of the possessor (And: 7f). The Tuvan Genitive can also mark the subject of subordinate clauses, e.g.

(3)

*dayïn–**nïŋ***	*čüge egel–e:n–in*	*bil–bes*	*men*
war–GEN	why begin–PST–ACC.3SG	know–NEG.AOR	1SG

'I don't know why the war started.'

[A/H: 20]

In Tofa, too, the Genitive functions as the modifier in adnominal constructions, e.g. *bo kiši–**niŋ** a't–ï* [this person–GEN horse–POSS.3SG] 'this person's horse' (Ras: 37), although as in other Turkic languages the Nominative is used in constructions of abstract possession and part-whole relations, such as *ibi düg–ü* [reindeer fur–POSS.3SG] 'reindeer fur' (Ras: 36).

The Genitive in Khalaj predominantly has the form *–I:n*, while in Chuvash it is *–n/–ən/–yən/–nən*. The functions are the same as in the other Turkic languages,

namely to mark the possessor in a possessive noun phrase (Drf: 79f; Benzing [1942] 1993: 49f). In Chuvash, the possessum agrees with nominal possessors (Benzing [1942] 1993: 51), while in noun phrases with pronominal possessors the possessum often lacks agreement with the possessor, the relationship being marked solely by the Genitive case-marking on the pronoun. This holds especially for the first and second person plural possessor (Krueger [1961] 1997: 113, 118). However, in Khalaj, the Nominative case can also be used to mark the modifier in possessive attribution. There seems to be no difference between the use of the Nominative or Genitive in such constructions, although the Nominative is used more frequently than the Genitive. The possessum is generally marked for agreement with a 3rd person possessor, both following a Nominative-marked modifier as well as following a Genitive-marked modifier (Drf: 76, 80).

In contrast to the above, the possessor in a Sakha possessive noun phrase is always unmarked, even in specific possessive constructions, while the possessum carries a possessive suffix agreeing with the person and number of the possessor, e.g. *uču:tal jie–te* [teacher house–POSS.3SG] 'the teacher's house', *min jie–m* [1SG house–POSS.1SG] 'my house'. The word order in possessive constructions is fixed, with the possessor preceding the head noun. This construction is used not only for actual possession, but for all kinds of adnominal relations, e.g. *süöhü sït–a* [livestock smell–POSS.3SG] 'the smell of livestock', *haχa jon–o* [Sakha people–POSS.3SG] 'Sakha', *min töbö–m* [1SG head–POSS.1SG] 'my head'. However, when the modifying noun expresses the material out of which the head noun is made, then the two nouns are simply juxtaposed, without possessive suffixes on the head noun, e.g.

(4)

en	*ta:s*	*jieγe*	*handa:ran*	*oloroŋŋun*
en	ta:s	jie–GA	sanda:r–An	olor–An–GIn
2SG	stone	house–DAT	glow–PF.CVB	sit–PF.CVB–PRED.2SG

'You live like a prince in a brick house.'

[IvaP155]

Furthermore, relative clauses and subordinate clauses in Sakha often lack an overt subject (which in relative clauses is implicit in the possessive marking on the head of the relative clause, and in subordinate clauses is implicit in the possessive marking on the nominalized verb – 5a, b), or the subject is in the Nominative case (5c, d). When it is the subject that is relativized, it appears in the case governed by the matrix verb (5e, f):

(5a)

če	*taptï:r*	*taptï:r*	*jonnorugar*	*bieren*
če	tapta:–Ar	tapta:–Ar	jon-LAr–(t)IgAr	bier–An
well	love–PRSPT	love–PRSPT	people–PL–DAT.3SG	give–PF.CVB

bieren	*kebiste*	*emie*	*bez*	*kopeyki*
bier–An	kebis–TA	emie	bez	kopeyki
give–PF.CVB	PFV–PST.3SG	also	without	kopeck.GEN

'Well, he gave and gave to the people that he liked, also for nothing.'

[IvaP57]

(5b)

χas	*uon*	*ča:nńïk*	*čeyi*	*örörbün*	*aybït*
χas	uon	ča:nńïk	čey–(n)I	ör–Ar–BIn	ay–BIt
how.many	ten	teapot	tea–ACC	put.to.boil–PRSPT–ACC.1SG	creator–1PL

beyete	*biler*	*künŋe*
beye–(t)A	bil–Ar	kün–GA
self–POSS.3SG	know–PRSPT	day–DAT

'... how many dozens of teapots I put to boil each day the Lord alone knows.'

[MalA91]

(5c)

en	*Ivanovtar*	*saŋa*	*tuttubut*	*jielerin*
en	Ivanov–LAr	saŋa	tut–(I)n–BIt	jie–LArIn
2SG	I.–PL	new	hold–REFL–PSTPT	house–ACC.3PL

körbütüŋ	*duo*
kör–BIt–(I)ŋ	duo
see–PSTPT–POSS.2SG	Q

'Have you seen the new house that the Ivanovs built for themselves?'

[translation Ver]

(5d)

min	*Saxa*	*tïlïn*	*Saxa*	*oloɣun*	*maygïtïn*
min	Saxa	tïl–(t)In	Saxa	oloχ–(t)In	maygïn–(t)In
1SG	Sakha	tongue–ACC.3SG	Sakha	life–ACC.3SG	similarity–ACC.3SG

bilerbin	*isten*	*bu*	*toyon*	*miɣigin*
bil–Ar–BIn	ihit–An	bu	toyon	mi:gin
know–PRSPT–ACC.1SG	hear–PF.CVB	this	chief	1SG.ACC

kördüön	*ïlbïta*
kördö:–An	ïl–BIt–(t)A
request–PF.CVB	take–PSTPT–POSS.3SG

'Hearing that I knew the Sakha language and the Sakha ways of life, this lord took me to assist him.'

[Uvar177]

(5e)

*beɣehe kulu:p kiriliehiger itirik olorbut **kihini** bügün*
beɣehe kulu:p kirilies–(t)IgAr itirik olor–BIt kihi–(n)I bügün
yesterday club porch–DAT.3SG drunk sit–PSTPT person–ACC today

χontuoraɣa kördüm
χontuora–GA kör–TI–(I)m
office–DAT see–PST–POSS.1SG

'I saw the man who had been sitting drunk on the porch of the club at the (administration) offices today.'

(5f)

*Nastyanï ïtïrbït **ït** onno sïtar*
Nastya–(n)I ïtïr–BIt ït onno sït–Ar
N.–ACC bite–PSTPT dog there lie–PRSPT

'There lies the dog that bit Nastya.'

[translation, Ver]

However, some vestiges of the old Turkic Genitive are still found in Sakha. Thus, in a sequence of possessive noun phrases with a third person possessor, the intermediate noun phrase carries the old Genitive suffix, which is formally identical to the third person Accusative suffix in the possessive declension (Stachowski & Menz 1998: 428; cf. Appendix 3), e.g.

(6)

kergenim […] *ńu:čča tïl**ïn** uču:tala*
kergen–(I)m […] nu:čča tïl–(t)In uču:tal–(t)A
spouse–POSS.1SG […] Russian word–GEN teacher–POSS.3SG

'My husband […] (was) a teacher of Russian.'

[RaxA121]

3.2.1.2 The genitive case in Mongolic languages

Proto-Mongolic had a Genitive case marked by the suffixes *–yin/*–Un/*–U, which marked attributive dependence in noun phrases. Most modern Mongolic languages, with the exception of Dagur and the languages of the Gansu-Qinghai region in China, have retained this Genitive case (Janhunen 2003d: 14). Thus, the Khalkha Genitive is marked by the suffixes *–ï(n)/–iy(n)/–n*, and it expresses the dependent noun in a noun phrase, with the head noun being unmarked, e.g. *usn–**ï***

šuvu: [water.OBL–GEN bird] 'water bird', *Ba:tar–ïn exner* [Baatar–GEN wife] Baatar's wife' (K/Ts: 81). It also marks the subject of a relative clause, e.g.

(7)

> *Dorž–iyn xudalda–ž av–san mašin–ïg bi üz–sen*
> D.–GEN sell–IPF.CVB take–PSTPT car–ACC 1SG see–PSTPT
>
> 'I saw the car that Dorž bought.'

<div align="right">[K/Ts: 82]</div>

In Buryat, the Genitive expresses adnominal attributes, such as possession, e.g. *ax–iin nom* [older.brother–GEN book] 'brother's book', general attribution, e.g. *namar–ai üder* [autumn–GEN day] 'autumn day', and partitive constructions, e.g. *xübüü–d–ei negen* [son–PL–GEN one] 'one of the children'. (Skr: 108f). "The genitive also marks the primary actants of deverbal nouns (and non-finite predicates), e.g. *uhan–ai uradxal* [water.OBL–GEN flow(ing)] 'flow of water', *shuluun nüürhen–ei maltalga* [stone.OBL coal.OBL–GEN digging] 'mining of coal'." (Skr: 109) The noun referring to the possessum is unmarked, as in Khalkha Mongol.

According to Tsumagari (Tsum: 138), Dagur (a Mongolic language spoken in Inner Mongolia that has been in intensive contact with speakers of the Tungusic language Solon) has lost its separate Genitive case, using a single case form (called 'Connective') to express both an attribute in adnominal constructions and definite direct objects, compare *mo:d–i: larc* [tree–CON leaf] 'leaf of a tree' and *ter xu:–y ši: tani–bei–š ye:* [that person–CON 2SG know–NPST–2SG Q] 'Do you know that man?' (Tsum: 138). However, it has retained the distinction between the Genitive and Accusative case in the singular personal pronouns (Tsum: 141). Possession is marked by possessive suffixes on the possessed noun that agree in person and number with the possessor, e.g. *aca:–yi–na:ny mor–iny* [father–CON–POSS.1PL.INCL horse–POSS.3SG] 'our father's horse' (Tsum: 142). Thus, Dagur has partly restructured its possessive constructions on the pattern of Tungusic. However, the data is somewhat unclear, in that three different authors give slightly differing descriptions. Thus, Todaeva (1986) claims that possessively marked head nouns are used only in conjunction with personal pronouns, e.g. *ši: mini: keku–min ul tani–n– ši:–ye:* [2SG 1SG.GEN son–POSS.1SG NEG know–NPST–2SG–Q] 'Don't you know my son?' (Todaeva 1986: 48), while Martin ([1961] 1997: 25) describes Dagur as having a Genitive case distinct from the Accusative (with a suffix ending in –*n*, as is common in Mongolic languages) for nouns as well as for pronouns. It is not clear what the reason for this discrepancy in the descriptions is, but one possibility might be a lack of time spent with native speakers of the language, or possibly Tsumagari documented a later stage of an ongoing process.

Khamnigan Mongol, on the other hand, which is another Mongolic language in close contact with a Tungusic language (in this case, Khamnigan Evenki) has not lost its Genitive case (Janhunen 1990: 55ff). However, Khamnigan Mongol has developed possessive suffixes out of Genitive-marked personal pronouns (Janhunen 1990: 64ff). In possessive constructions with pronominal modifiers there are three alternative means of expressing the possessive relation: a) with a Genitive-marked personal pronoun as modifier and unmarked head (*minii ijii* 'my mother'), b) with Genitive-marked personal pronoun followed by a possessive-marked head (*minii ijii–mini* 'my mother'), or c) with possession marked on the head solely with the possessive suffix (*ijii–mini* 'my mother'). According to Janhunen (1990: 66), variant a) might be due to influence from Standard Mongolian, in which Genitive-marked personal pronouns function as possessive pronouns, but the head noun does not agree with the possessor (Vietze 1988: 77). However, it should be noted that all Mongolic languages have possessive suffixes or clitics which are derived from Genitive case-marked pronouns; in Buryat it appears to be possible to use them in constructions together with a possessive pronoun, e.g. *(minii) axa–mni* [(1SG.POSS) elder.brother–POSS.1SG] 'my elder brother (Skr: 112), e.g. in a very similar way to the above Khamnigan Mongol examples. It is thus not really clear to what extent Khamnigan Mongol differs from other Mongolic languages. However, the possessive suffixes or clitics of Khalkha do not appear to co-occur in constructions with nominal possessors (K/Ts: 107, Vietze 1988: 84), in contrast to the situation in Dagur (cf. the example above). Unfortunately, it is unclear whether the Khamnigan possessive suffix can occur with a nominal possessor or not, as Janhunen (1990: 65f) gives only examples with the possessively marked noun in isolation, or with a pronominal possessor, as in b) above.

3.2.1.3 The genitive case in Tungusic languages

As will be presented in more detail in the following sections, both the Northern Tungusic and Amur Tungusic languages have no case-suffix to mark the attributive possessor in a possessive noun phrase (Evenki is an exception, as will be discussed below.) However, all languages have a suffix *–ŋi* that attaches to possessors in predicate position, which is variably called the 'direct possessive' (B/G: 14f), 'independent possessive form' (Rišes & Cincius 1952: 714), 'emphatic possessive' (Mal: 11), and 'predicate possessive' (AvrI: 186). This suffix is formally identical to the Tungusic Alienable Possessive suffix *–ŋi* (cf. section 3.4.2.1); however, the latter attaches to the possessum, while the former attaches to the possessor. It therefore seems justified to differentiate between the two suffixes (cf.

Boldyrev 1976: 33f). In addition to attaching to predicate nominal possessors, this suffix also attaches to the oblique stems of personal pronouns; in all the Tungusic languages pronouns derived in this way are used predicatively (cf. below), while in Evenki this additionally derives possessive pronouns (used attributively). Given the predominant use of this suffix to mark predicate possessors, I will refer to it as 'Predicate Possessive' in the following discussion.

3.2.1.3.1 Possessive constructions in Northern Tungusic languages

In Evenki attributive possessive constructions the possessor is unmarked, while the possessum agrees with the possessor in person and number, e.g. *ollomimni jav–in* [fisherman boat–POSS.3SG] 'the/a fisherman's boat' (Ned: 158). However, as mentioned, a construction that additionally marks the possessor (suggested to be the remnant of an old Genitive, Ned: 158, B/G: 14) is also possible, e.g. *atïrka:n*[3]*–ŋi gerbi–n* [old.woman–PRDPOSS name–POSS.3SG] 'the name of the old woman', *Ivul–ŋi oro–r–in* [Ivul–PRDPOSS reindeer–PL–POSS.3SG] 'Ivul's reindeer' (Ned: 158). According to Nedjalkov (Ned: 158), this was preferable in the first half of the twentieth century, but is rarely used nowadays. Writing in the first half of the twentieth century Sunik (1948: 287) claims that such constructions are relatively rare. Avrorin (1956: 97, 98) also emphasizes that the Predicate Possessive forms in attributive use are rare in spoken speech; thus, Nedjalkov's observation on their preferred use in the first half of the 20th century[4] may be due to the fact that he based his judgement on the analysis of folklore texts, which may show a stylistic difference to everyday spoken speech. Bulatova & Grenoble (B/G: 14f) describe this suffix as having a long vowel (not indicated by Nedjalkov) which distinguishes it from the Alienable Possessive suffix[5] (cf. section 3.4.2), while Boldyrev (1976: 33f) describes the Predicate Possessive suffix in Evenki as being homonymous with the Alienable Possessive suffix. As argued above, given the fact that the Alienable Possessive suffix attaches to the possessum, while the Predicate Possessive suffix attaches to possessors, it might be better to distinguish between the two, regardless

---

[3] Nedjalkov (Ned: 158), Bulatova & Grenoble (B/G: 14; cf. ex. 8a below) and Konstantinova (Kon: 64; cf. ex. 9 below) all spell this word differently – with or without a final long vowel, and also with a front or back high vowel. For the sake of consistency, I follow the orthography of Myreeva (2004: 63), who gives *atïrka:n*, in all examples

[4] From footnote 9 in Beisenherz (2001: 45) it becomes clear that the Predicative Possessive suffix was only infrequently used, i.e. that it was preferred in comparison to modern usage, but not preferred in comparison to the unmarked possessor.

[5] Somewhat surprisingly Whaley, Grenoble & Li (1999: 301) do equate this suffix with the Alienable Possessive suffix.

of whether there is any phonological difference between them. The Predicate Possessive suffix can be used both with (8a) and without (8b) further possessive suffixes:

(8a)

  *tar   učak atïrka:n–**ŋi:**–v*
  that   deer wife–PRDPOSS–POSS.1SG

  'That deer belongs to my wife.'

(8b)

  *er kniga kuŋaka:n–**ŋi:***
  this book   child–PRDPOSS

  'This book belongs to the child.'

<div align="right">[B/G: 14]</div>

Although the examples provided by Bulatova & Grenoble show the marked possessor only in predicative function, the examples given by Konstantinova (Kon: 64f) all show the marked possessor in attributive function, with the possessum agreeing with the possessor in case, person and number (9). The difference in use between a possessive NP with unmarked possessor and that with a possessor carrying the Predicate Possessive suffix is that in the latter the possessor is emphasized (Kon: 65). Note that in this example the Predicate Possessive suffix *–ŋi:* is followed by the Accusative case suffix:

(9)

  *xuna:t atïrka:n–**ŋi:**–va*   *avsa–kan–ma:–n*
  girl   old.woman–PRDPOSS–DEF.ACC box–DIM–DEF.ACC–POSS.3SG

  delum ga–da–n
  secret   take–PST–3SG

  'The girl secretly took the old woman's box.'

<div align="right">[Kon: 64]</div>

In attributive possessive phrases in Ėven the possessor is unmarked, while the possessed noun agrees in number and person with the possessor, e.g. *ju: heye–**n*** [house roof–POSS.3SG] 'the roof of the house' (NovI: 141, 146). However, as in Evenki, the suffix *–ŋi:* is added to nouns designating humans when they express possessors in predicate position, as independent constructions in answer to the question 'whose', or when the possessor is postposed (10a). In the Ola dialect, it can appear in subject or complement position when the possessum is deleted (10b), and also in attributive position (10c); the latter function is not found in literary Ėven (NovI: 151f). It should be noted that the three examples of this last function given by

Novikova are all with a pronominal possessor, so that it is not clear whether nominal possessors can carry the Predicate Possessive suffix in attributive position as well.

(10a)

*tara:k   ju:–n,   etïke:–**ŋi:**,   xo: irbe:t*
that   house–POSS.3SG old.man–PRDPOSS very old

'That house, the old man's, is very old.'

[Rišes & Cincius 1952: 714]

(10b)

*noŋïn ja:–l–ni   em–re   min–**ŋi**–l*
3SG   relative–PL–POSS.3SG come–NFUT.3PL 1SG.OBL–PRDPOSS–PL

timinak   em–ji–r
tomorrow   come–FUT–3PL

'His relatives have come, mine will come tomorrow.'

(10c)

*min–**ŋi**–l–bu   ja:–l–bu   em–u–de–ten*
1SG–PRDPOSS–PL–ACC relative–PL–ACC come–E–PURP–3PL

gö:–li
say–PRXIMP.2SG

'Tell my relatives to come.'

[NovI: 151]

In both Evenki and Ėven, the Predicate Possessive suffix *–ŋi:* is identical in form to the interrogative pronoun *ŋi:* (Ned: 214 renders the interrogative pronoun with a short vowel, corresponding to his rendering of the Predicate Possessive suffix with a short vowel). The Predicate Possessive suffix forms possessive pronouns from the oblique stems of personal pronouns, e.g. *bi:* 'I', *min–du:* [1SG.OBL–DAT] 'to me', *min–**ŋi:*** [1SG.OBL–PRDPOSS] 'mine, my' (Rišes & Cincius 1952: 726ff; Ned: 200, 207; B/G: 21ff). In the southern and northern dialects of Evenki, only possessive pronouns are used in attributive function, e.g. ***min–ŋi*** *ju–v* [1SG.OBL–PRDPOSS house–POSS.1SG] 'my house', while in the eastern dialects personal pronouns are used instead: ***bi*** *ju–v* [1SG house–POSS.1SG] 'my house' (Ned: 210). According to Boldyrev (1976: 16f), in Evenki pronominal possessors in attributive function originally stood in the Nominative case, and the possessive pronouns represent a later development. The possessive pronouns also occur predicatively, e.g.

(11)

>   *tar   purta   **min–ŋi**   bi–si–n*
>   that   knife   1SG.OBL–PRDPOSS   be–PRS–3SG

>   'That knife is mine.'

<div align="right">[Ned: 124]</div>

In Èven, pronominal possessors in attributive function do not take the Predicate Possessive suffix, but remain in the bare oblique stem, while possessive pronouns with the Predicate Possessive suffix are used predicatively or in independent constructions, as are nouns with this suffix (Rišes & Cincius: 727f). Compare: *bi:* 'I' (Mal: 12), ***min** teti–β* [1SG.OBL clothes–POSS.1SG] 'my clothes' (NovI: 142), and *tarak kini:γa **min–ŋi*** [that book 1SG.OBL–PRDPOSS] 'that book is mine' (NovI: 151). It has been suggested that both Evenki and Èven dialects in contact with Sakha are losing the possessive pronoun in attributive function, using the unmarked personal pronouns (corresponding to the Nominative case) instead (Malchukov 2006: 124). However, the Western Èven dialect of the Bulun and Verxojansk districts appears to have used full possessive pronouns in attributive function as well, e.g. ***miŋ–ŋi** haga–vu* [1SG.OBL–PRDPOSS collar–POSS.1SG] 'my collar' (Sotavalta & Halén 1978: 12).

Negidal lacks a genitive case and has possessive suffixes that attach to the possessed noun and agree in number and person with the possessor, e.g. *omki: hute–**nin*** [flying.squirrel child–POSS.3SG] 'the flying squirrel's child' (JaN5: 113, 114, Cincius 1982: 62). The Predicate Possessive suffix *–ŋi:* is attached to the possessor in predicate or independent clauses, e.g. *tay del moyin–**ŋi:*** [that head horse–PRDPOSS] 'that head is the horse's' (Cincius 1982: 20). This suffix also attaches to the oblique stem of personal pronouns when they express possession in predicate position, while in attributive position either the Nominative case of the personal pronouns is used or the bare oblique stem (Cincius 1982: 22).

3.2.1.3.2 Possessive constructions in Amur Tungusic languages and Manchu

The Amur Tungusic language Nanay lacks a genitive case, marking all kinds of adnominal relations (possession, part-whole relations, attributes, etc.) with the use of possessive suffixes (AvrI: 141). In Nanay the unmarked possessor obligatorily precedes the possessum, which takes obligatory possessive suffixes agreeing in person and number with the possessor (AvrI: 143f), e.g. *učenik daŋsa–**ni*** [pupil book–POSS.3SG] 'the pupil's book' (AvrI: 141). In attributive position pronominal possessors are in the unmarked (Nominative case) form (AvrI: 250). Both nominal and pronominal possessors in predicative position take the Predicate Possessive suffix *–ŋi*. It attaches to nouns designating humans and animals, as well as to human

collectives such as organizations and institutions, or settlements. When used with pronouns it attaches to the Nominative stem of the personal pronouns (AvrI: 186f). In addition to being used predicatively, this form can also be used in independent clauses and for postposed possessors, e.g. *ey daŋsa–va, mi:–ŋi–ve, min–du bu:–gu–ru* [this book–ACC 1SG–PRDPOSS–ACC 1SG.OBL–DAT give–ITER–PRXIMP.2SG] 'this book, mine, give it to me' (Avrorin 1956: 94).

Like Nanay, Udihe expresses possession with noun phrases in which the head noun agrees in person and number with the possessor (N/T: 106f), while the nominal possessor is unmarked, e.g. *təŋku bugdi–ni* [stool leg–POSS.3SG] 'a leg of the stool' (N/T: 482). As in Nanay, the word order in possessive noun phrases is fixed, with the dependent preceding the head noun (N/T: 480). When the possessor is used not in attributive function, but predicatively, it is marked by the suffix *–ŋi*. Nikolaeva & Tolskaya suggest that this is the same as the Alienable Possessive suffix; however, as they themselves point out, the Alienable Possessive suffix attaches to the possessum, while the suffix marking possessors in predicative function attaches to the noun denoting the possessor (N/T: 141, 634); as stated above, in my opinion this syntactic difference is sufficient to warrant distinguishing the two suffixes. As in Èven, in Udihe both oblique pronominal stems and nouns take the Predicate Possessive suffix *–ŋi* when used either as the predicate of a copula construction (12a) or when they are used independently, as an answer to the question 'whose' (12b). The possessive pronouns derived with the Predicate Possessive suffix *–ŋi* are not used attributively (N/T: 336, 634); for this, personal pronouns in the Nominative case are used, e.g. **bi** *anda–i* [1SG friend–POSS.3SG] 'my friend' (N/T: 481).

(12a)

  *əi ugda nuati–**ŋi***
  this boat   3PL–PRDPOSS
  'This boat is theirs.'

(12b)

  *ni   kusigə–ni – b'ata–**ŋi***
  who knife–3SG   boy–PRDPOSS
  'Whose is the knife? – The boy's.'

<div align="right">[N/T: 635, 634]</div>

In Oroč, too, possession is expressed by possessive suffixes on the possessum, with the possessor remaining unmarked (A/B: 89ff), e.g. *boloŋgo gida–ńi* [hunter spear–POSS.3SG] 'the hunter's spear'. Pronominal possession is expressed by personal pronouns in the Nominative (A/B: 188), e.g. **bi:** *ugda–i* [1SG boat–POSS.1SG] 'my boat' (A/B: 188). There is no mention in either Avrorin & Boldyrev

(A/B: 88ff) or in Avrorin & Lebedeva (JaN5: 195f, 200f) of the Predicate Possessive suffix; however, since both the closely related language Udihe as well as less closely related Nanay have it, as do the Northern Tungusic languages, this might just be an oversight on the part of the authors.

Manchu is the only Tungusic language that has a Genitive case in regular adnominal use and that does not have a category of morphologically marked possession on the head noun. The Genitive is marked by the postposed marker[6] *(n)i* and expresses possessors, part of a whole, material, and other kinds of attribution. The possessum appears in the unmarked (Nominative) case, e.g. *biya **i** elden* [moon GEN light] 'the light of the moon', *buhu: **i** weihe* [deer GEN horn] 'horns of a Manchurian deer' (Gor: 175ff). Genitive-marked personal pronouns function as possessive pronouns, e.g. ***min–i** ama* [1SG.OBL–GEN father] 'my father' (Gor: 216). When used independently (as answers to the question 'whose'), a possessive suffix/particle *–ŋge/niŋge* attaches to the Genitive case form of the personal pronouns (13). These forms never occur as modifiers of head nominals (Gor: 219).

(13)
   *ere uthai we–i   yaka?   –mini–**ŋge***
   this then who–GEN thing   1SG.GEN–POSS
   'Whose thing is this? – Mine.'

                                                                                   [Gor: 219]

Although the possessive suffix takes its origin as a possession marker, it now appears to primarily derive nominals having an attributive function, with a shift towards denominal adjectives, e.g. *beye–iniŋge* [self–POSS] 'one's own', *eyen niŋge* [ruler POSS] 'belonging to a ruler', *niyalma–iŋge* [person–POSS] 'human' (Gor: 153). The Manchu Genitive construction may have developed under Chinese influence, since in Mandarin Chinese possessive NPs are characterized by a possessive particle linking the dependent and the head noun, e.g. *bàba de lǐngdài* 'father's tie' *xuéxiào de yùndòngchǎng* 'the school's sportsfield', *wǒde xié* 'my shoes' *tāde wàzi* 'his/her socks/stockings' (Yip & Rimmington 1997: 26).

3.2.1.3.3 A genitive case in Tungusic?

Table 3.3 presents an overview of nominal and pronominal possessive marking in the Tungusic languages. As can be seen, all the languages for which I was able to find relevant information except Manchu use the Predicate Possessive

---

[6] The genitive marker *(n)i* is mostly written separately from the noun; only occasionally are the noun and the particle written as one word (Gor: 175).

suffix –ŋi to mark both nominal and pronominal possessors in predicate position. The Manchu suffix –ŋge shows such similarity in form and use that it may be considered cognate to the Predicate Possessive suffix in other languages; this view is supported by Avrorin (1956: 96f). The Amur Tungusic languages do not have a genitive case; both attributive possessive nouns and pronouns are used in the unmarked Nominative form. The Northern Tungusic languages are somewhat mixed, with (Siberian) Evenki, Khamnigan Evenki and the Ola dialect of Ėven occasionally using the Predicate Possessive suffix to mark nominal possessors in attributive position, while Standard Ėven and Negidal use the Nominative case. Attributive possessive pronouns are distinct from personal pronouns in all Northern Tungusic languages; the Evenki dialects use the Predicate Possessive suffix to derive possessive pronouns, while the Ėven dialects and Negidal use the oblique stem of the personal pronouns as possessive pronouns in attributive position, restricting the use of the Predicate Possessive suffix to pronominal possessors in predicate position.

Table 3.3: Overview over possessor marking in Tungusic languages (the 1SG pronoun is used for illustrative purposes)

	attr poss noun	pred poss noun	free pers pron	attr poss pron	pred poss pron
Evenki	NOM –ŋi: (=GEN)	–ŋi	bi	minŋi	minŋi
Khamnigan Evenki	–ŋi: (=GEN) NOM	???	bi(:)	minni:	???
Ėven Standard	NOM	–ŋi	bi:	min (OBL)	minŋi
Ėven Ola	NOM –ŋi (=GEN)?	–ŋi	bi:	min (OBL) minŋi	minŋi
Negidal	NOM	–ŋi	bi	bi min (OBL)	minŋi
Nanay	NOM	–ŋi	mi	mi	miŋi
Udihe	NOM	–ŋi	bi	bi	miniŋi
Oroč	NOM	???	bi:	bi:	???
Manchu	(n)i (=GEN)	???	bi	min–i (GEN)	miniŋge

Abbreviations: attr: attributive; poss: possessive; pred: predicative; pers: personal pron: pronoun

As can be seen from Table 3.3 and the preceding discussion, in nearly all the Tungusic languages with the exception of Manchu the nominal possessor in attributive possessive constructions remains in the unmarked (Nominative) form. Evenki stands out amongst its sister languages in that the 'Predicate Possessive' suffix –ŋi can attach to attributive possessors, which is a use reminiscent of a

genitive case; this suffix has therefore been called a Genitive case in previous descriptions (cf. Sunik 1948: 285; B/G: 14; Ned: 158). However, this 'genitive-like' construction appears to be much more rarely used than the unmarked noun in possessive constructions. With regard to pronominal possessors in attributive position, the Tungusic languages show two basic patterns: In one (found partly in Negidal, in Udihe and in Oroč), the pronominal possessor is identical to the free pronoun in subject position, thus showing a structure parallel to possessive constructions with nominal possessors. In the other pattern, either just the oblique stem or the oblique stem plus Predicate Possessive suffix are used to mark attributive possession; thus, possessive pronouns are distinct in these languages from free personal pronouns. The latter pattern with the oblique stem used for pronominal possessors in attributive function is found in languages belonging to both the Northern Tungusic group (Evenki, Standard Ėven and optionally Negidal) and the Amur Tungusic group (Ulča and optionally Orok). One might possibly analyze Nanay as having this construction as well, with loss of the oblique –*n* (Frederik Kortlandt, pers. comm.): Nanay has *mi* for the 1SG attributive possessor and *mi–ŋi* for the 1SG possessive pronoun in predicative position. However, this analysis is problematic, since in Nanay the 1SG pronoun in subject position is *mi*, identical to the pronoun in possessive function; thus, in Nanay there is no formal distinction between pronouns in subject position and pronouns in possessive constructions, which is similar to what is found in Udihe and Oroč. Since pronouns are often more archaic and retain distinctions that nouns have lost (Bernard Comrie, pers. comm.), one could argue that the Tungusic languages used to have a separate genitive-like category to mark possessors (cf. Kortlandt [2004] 2006: 2) which has been lost in the nominal system of all languages. Thus, Boldyrev (1976: 17) argues that the use of the oblique stem in attributive pronominal possession in Negidal, Ulča, Orok and Standard Ėven goes back to a former Genitive case suffix –*i* that was later lost in these languages when the development of possessive suffixes marking the possessum made the additional marking of the possessor redundant. Similarly, he argues that the Evenki suffix –*ŋi* is a relic of the previous Genitive case suffix which lost its possessive-marking function after the possessive suffixes developed (Boldyrev 1976: 25); however, he does not address the issue of similarity between the Predicate Possessive suffix –*ŋi* found in nearly all Tungusic languages and the Evenki 'Genitive' relic. In Manchu the Genitive case that marks possessors and the loss of the possessive suffixes to mark possessees may well have developed under Chinese influence. The use of the Predicate Possessive suffix to mark possessors in attributive position in Evenki is probably a recent innovation (cf. Avrorin 1956: 98), since it is restricted to this single language. Sunik (1948: 287), however, suggests

that the Tungusic languages originally did not have a genitive case, that this development is unique to Manchu, and that Evenki, though on the way towards developing a genitive case, has not fully realized this development yet.

It has been argued that some Tungusic languages spoken in China have developed a Genitive case under contact influence with neighbouring languages. Thus, Kilen Nanay, the main Nanay dialect spoken on the Chinese side of the Amur river, has developed a Genitive case, in contrast to the Nanay dialects spoken on the Russian side of the Amur river, which lack it. In this Chinese Nanay dialect, possession can be expressed either by the Genitive case, or by possessive suffixes on the possessed noun, or both. That the development of a Genitive case in Kilen Nanay is due to Manchu influence is suggested by the fact that in both Manchu and Kilen Nanay the Genitive case suffix can also fulfil instrumental functions (Tsumagari 1997: 179). Furthermore, Evenki dialects spoken in China have also developed a Genitive case, probably under contact influence with Manchu and/or Mongolian. Thus, Tsumagari (1997: 181) argues that Khamnigan Evenki, which has been in close contact with Khamnigan Mongol, has a Genitive case. However, it is not clear in my opinion whether this can truly be attributed to contact influence, since the 'Genitive case' suffix is *–ŋi:* after vowel stems and *–ni:* after consonant stems (Janhunen 1991: 62), thus resembling the use of the Predicate Possessive suffix in attributive position also found in other Evenki dialects, as described above. Genitive case-marked personal pronouns function as possessive pronouns (Janhunen 1991: 68). Possession can be expressed either by the Genitive case alone, or by a possessive suffix on the possessum, or by a combination of both, e.g. *min–**ni:** girki– l(–**bi**)* [1SG.OBL–GEN friend–PL(–POSS.1SG)] 'my friends' (Janhunen 1991: 71).

3.2.1.4 The genitive in other Siberian languages

Yukaghir has three ways of expressing adnominal relations between heads and their modifiers: 1) through simple juxtaposition of the two nouns (14a), 2) by marking the possessum with a suffix that expresses the grammatical relation between the modifier and its head, but does not agree in person or number with the possessor (14b), and 3) by the use of attributive marking on the modifier (14c) (Mas: 289). This latter form has been interpreted as a genitive case (Krejnovič 1958: 63f; Fortescue 1998: 10f, 65). However, as will be argued here, the function of this attributive marker in Yukaghir differs substantially from the usual functions of genitive cases, and this suffix should therefore not be considered a genitive case.

(14a)

  qa:qa: šötkuri: key–k
  grandfather ski   give[IMP]–2SG

  'Give me grandfather's ski'.

                                               [Mas: 290]

(14b)

  *met Alandin u:yçi:–**gi**   ŋo:–je*
  1SG A.   worker–POSS COP–INTR.1SG

  'I am Alandin's servant.'

                                               [Mas: 298]

(14c)

  *ubuy mošolupka: yuku–yoŋža:–**n**+martl'uö–k min–te–mle*
  true   owl   small.goose–ATTR+girl–PRD   take–FUT–OF.3SG

  'Is that true that the owl is going to marry a small goose girl[7]?'

                                               [Mas: 307]

The difference between simple juxtaposition and possessive marking on the head lies in that head-marked constructions assign greater prominence and separate reference to the possessor than juxtaposed constructions, while juxtaposition is used in apposition as well and can sometimes lead to ambiguity between a possessed reading and an appositional reading (Mas: 296). Thus, (14b) without the possessive suffix could mean either 'I am Alandin's servant' or 'I am servant Alandin' (Mas: 298). The head-marked construction might have developed through contact with Ėven or Sakha (cf. Fortescue 1998: 65). The construction with attributive marking on the modifier, on the other hand, signifies that the modifier is non-referential (Mas: 289, 304). It is used "to signal that the conceptualization provided by the modifier is not salient enough to constitute an appropriate referential expression for the relevant entity in the given situation. […] the head noun provides a clearly more salient part of the description than the modifier noun." (Mas: 305). In this, the attributive construction has more in common with derived adjectives, which also provide additional information on the head noun only, and it is in my opinion not appropriate to classify the attributive marker as a synchronic genitive case suffix[8].

---

[7] There is a mismatch between the number of morphemes separated by syllables between the Yukaghir line and the gloss in the phrase *yuku–yoŋža:–n+martl'uö–k*. I suspect that the gloss should be [small–goose…], but I lack the necessary information on the meaning of the individual morphemes to say this for certain. Unfortunately, the root *yoŋža:* is not found in the Yukaghir-English vocabulary list appended to Maslova's grammar; nor was I able to find it in Kurilov (2001).

[8] However, historically it might have developed from a genitive case (Frederik Kortlandt, pers. comm.).

Chukchi has two different suffixes to mark adnominal relations in noun phrases. They both attach to the possessor, while the possessum remains unmarked (Dnn: 148f). One of these suffixes is called the possessive suffix, the other is called the relational suffix. The possessive suffix expresses possession, kinship relationships, part of a whole, and others (Koptevskaja-Tamm 1995: 304), while the relational suffix expresses meanings of source, origin or purpose[9] (Dnn: 148). Dunn does not count these possessor-marking suffixes among the case suffixes, because they can combine with other case and derivational suffixes (Dnn 149). However, although multiple case marking is cross-linguistically rare, combinations of an adnominal case with an adverbal case do occur (Blake 2001: 107; cf. the volume edited by Frans Plank 1995). Thus, the divergent behaviour of the possessive and relational suffixes with respect to additional case-marking (as well as the fact that they take predication suffixes) alone is not sufficient reason to exclude these forms from the inflectional paradigm (Koptevskaja-Tamm 1995: 313). However, an argument in favour of regarding these suffixes as derivational rather than inflectional is their behaviour in incorporation: In Chukchi, nouns can normally only modify other nouns when they are incorporated in the latter; in incorporation, the modifying nouns lose their case and number categories, although they can retain their derivational suffixes. The possessive and relational suffixes, however, are retained in incorporation (Koptevskaja-Tamm 1995: 314), making them more similar to derivational than to inflectional suffixes. Thus, these suffixes are of a somewhat ambiguous nature. However, regardless of the actual classification of the possessive and relational suffixes as inflectional or derivational, the possessive construction in Chukchi is clearly very different from that found in Sakha, Tungusic, and Yukaghir: In Chukchi, it is the possessor that is marked, while the head noun does not agree with the possessor in person or number. Furthermore, the possessor agrees in number with the possessum when the latter is not overtly present, and in person and number when the possessum is not third person, e.g. ***turɣ–ine–ɣəm qora–ɣənret–ə– lʔ–eɣəm*** [2PL–POSS–1SG.ABS reindeer–herder–E–NR–1SG] 'I am your.PL herdsman.' (Dnn: 150). In Sakha, Tungusic and Yukaghir, the possessor remains in the unmarked (Nominative) case, while in Sakha and Tungusic the noun denoting the possessum takes personal suffixes agreeing with the possessor in person and number, as discussed in section 3.2.1.1 and 3.2.1.3.

In Itelmen, too, adnominal relations (possession, material, part-whole relationship, and place of origin or habitation) are expressed by noun phrases in

---

[9] This difference is comparable to the distinction between the Italian prepositions *di* and *da*, with *da* expressing, among others, purpose (Frederik Kortlandt, pers. comm.; Maiden & Robustelli 2000: 184, 185).

which the possessor is marked by one of three possible suffixes, while the possessum remains unmarked (G/V: 92). Functionally, the possessor-marking suffixes are equivalent to genitive case marking; however, Georg & Volodin classify them as belonging neither to the inflectional nor to the derivational suffixes, because the possessor suffixes can be followed by a further case suffix, namely the Instrumental (G/V: 93). However, as pointed out in the discussion of the Chukchi possessive suffix, although double or multiple case-marking is cross-linguistically rare, it is not impossible (Blake 2001: 107; see also the volume edited by Frans Plank 1995). Thus, the fact that the Itelmen possessives can take further case marking (restricted, in any case, to the Instrumental – G/V: 98) does not mean that they cannot be classified as *bona fide* Genitive case suffixes.

In Eskimo languages, the possessor in a possessive noun phrase stands in the Relative case, which is the case that also marks the A of transitive verbs (i.e. the A of ergative constructions), while the possessum carries a suffix agreeing with the possessor in person and number (Men: 48, 56f; de Reuse 1994: 30; Fortescue 1984: 216). In Siberian Inupik, first and second person possessors are generally omitted, while third person nominal possessors are an obligatory part of the possessive construction (Men: 56). In a chain of possessors, the first possessor carries the bare Relative case suffix, the second element carries a possessive-marked Relative case suffix agreeing in person and number with the first possessor, and the last possessed noun agrees in person and number with the second element. Only the possessum can take case marking according to the function of the possessive NP in the clause, e.g. *navřa–m ïmγ–**a–nun*** [lake–REL water–POSS.3SG–DAT] 'into the water of the lake' (Men: 57).

In Nivkh adnominal cases are lacking; possessive or attributive noun phrases usually consist of two juxtaposed nouns in the Absolutive case (PanI: 156; Mat: 8), e.g. *ïmïk zus* [mother meat] 'mother's meat/flesh' (PanI: 129). However, very rarely the noun referring to the possessor can be marked by the suffix *–eo*, which is also used to derive possessive pronouns for attributive functions, e.g. *kinz–**eo** ranr–š* [devil–POSS sister–?] 'the devil's sister' (PanI: 156, footnote 93), *ïtïk–**eo** ranr* [father–POSS sister] 'father's sister' (PanI: 129); compare: *ń–**eo** ranr* [1SG–POSS sister] 'my sister', *v–**eo** ranr* [3SG–POSS sister] 'his sister' (PanI: 252). Predicatively used possessive pronouns are derived from the personal pronouns with the suffix *–nï*, which has grammaticalized from the noun *nï* 'thing', e.g.

(15)

> *tï bitγï či–nï–la – hï ńi–**nï**–ra*
> this book 2SG–PRDPOSS–Q this 1SG–PRDPOSS–PRD
>
> 'Is this book yours? – Yes, (it's) mine.'   [PanI: 252]

However, usually plural and dual pronouns as well as nouns in attributive constructions modify their head noun without taking any further marking themselves, e.g. *ńïŋ* 'we', *ńïŋ pila dïf* [1PL big house] 'our big house' (PanI: 156). Singular personal pronouns, on the other hand, are most often reduced in form[10] and prefixed to the possessed noun, e.g. *či* 'you', *č–rïf* [2SG–house] 'your house (PanI: 157). The possessively marked pronouns with the suffix *–eo* are nowadays used very rarely (PanI: 252), and Gruzdeva (Grz: 28) doesn't mention this suffix at all.

In Ket, possession can be expressed by either suffixes on the possessor noun, or by prefixes on the possessum; the two kinds of possessive marking do not occur in the same constructions. The possessor-marking suffixes have been classified as Genitive case suffixes (Wer: 112, Dul: 74f); they express adnominal relations such as possession, kinship, a characteristic or a part of a whole. A prefixally marked possessum generally occurs without a modifying noun; however, occasionally a construction consisting of a modifier plus prefixed possessum can occur when the modifier is focussed (Andrej Nefedov, pers. comm.). The possessive prefixes of the third person are formally identical to the Genitive case suffixes (Dul: 62, 74); this can make the distinction between a possessor-marking suffix (Genitive case) and possessum-marking prefix difficult; compare *qus'* 'teepee' *de–qus'* [POSS.3SG.M–teepee] 'his teepee', *ob–de–qus'* [father–POSS–teepee] 'father's teepee'; this latter can be analyzed as *ob–de qus'* [father–GEN teepee] (which is the way Dul'zon 1968: 74 analyzes such constructions) or *ob de–qus'* [father POSS.3SG.M–teepee], which is how Krejnovič (JaN5: 457) would analyze it. Thus, Krejnovič (JaN5: 457) claims that there is no genitive case in Ket, but only possessive prefixes: "Since in the speech flow possessive prefixes that do not carry stress can move to the modifying noun, a Genitive case has erroneously been distinguished in Ket which in reality does not exist." (Krejnovič, JaN5: 457, translation mine[11]). Dul'zon, too, points out that the 'Genitive suffixes' are frequently used separately from a previously mentioned noun and thus turn into prefixes (Dul: 75). In my opinion, it might be best to analyze these morphemes as very mobile possessive clitics that can either encliticize to the possessor or procliticize to the possessum. This analysis is strengthened by the fact that in text counts the majority usage (~ 60%) is as a possessive prefix, with only about 30% of occurrences being as possessor-marking suffixes; occasionally (~ 10% of the time) the possessive marker stands as a separate word between the possessor and the possessum (Andrej Nefedov pers. comm.). However, Werner (Wer: 118f)

---

[10] An exception is made when the possessed noun begins with two consonants (PanI: 157).

[11] Original: "Так как в потоке речи притяжательные префиксы, не несущие на себе ударения, могут отходить к определению, в кетском языке ошибочно выделялся родительный падеж существительных, которого в действительности в нем нет."

rejects this analysis, claiming that the prefixes and suffixes serve different functions and occur in different position in the noun phrase – this latter point, of course, does not contradict the analysis as mobile clitics.

The Northern dialect of the Ob-Ugric language Khanty (also called Ostyak) has only three cases, Nominative, Locative, and Translative. It distinguishes two different kinds of possessive constructions, referred to by Nikolaeva as the neutral and the head-marked construction (Nik: 52). In neither of the constructions is the possessor marked in any way; the difference lies in the marking of the head noun. Thus, head-marked possessive constructions are used when the possessor is referred to by a personal pronoun; in this case, the possessum takes suffixes agreeing with the possessor in person and number (16a), while the possessor can be omitted. In the neutral construction the modifying possessor noun is only juxtaposed to the possessed noun, which does not take any possessive suffixes (16b). The syntactic function of the possessive NP is marked only on the head noun (Nik: 52).

(16a)

   (*ma*) *xo:t–**e:m**–na*
   1SG house–POSS.1SG–LOC

   'in my house'

(16b)

   Juwan xo:t–na
   John   house–LOC

   'in John's house'

                                              [Nik: 52]

The same distinction between head-marked possession with pronominal possessors and neutral possession with nominal possessors is made in the Vakh dialect of Khanty, e.g. *me rït–**əm*** [1SG boat–POSS.1SG] 'my boat' (Trš: 37), *ət'ï–m rït* [older.brother–POSS boat] 'my older brother's boat' (Trš: 44). In the northern dialects, possessive pronouns can occur only in headless noun phrases or predicatively (Nik: 16; Steinitz [1937] 1980: 36), e.g. in the Kazym dialect *tam laŋki muy **naŋen** muy **luvel*** [that squirrel or 2SG.POSS or 3SG.POSS] 'Is that squirrel yours or his?' (Steinitz [1937] 1980: 36f). These appear to be lacking in the Vakh dialect, where one only finds emphatic forms of the personal pronouns with a meaning of 'I myself', 'you yourself' etc; these appear to be occasionally used in attributive position (Trš: 66f; Gulya [1966] 1997: 76).

In the Sosva dialect of Mansi (also called Vogul) the unmarked (Nominative) case expresses possessive attributes (Rie: 25; Rmb: 43). The head noun can be marked with a possessive suffix agreeing in number and person with the possessor,

e.g. *χum sa:γrap–e* [man axe–POSS.3SG] 'the man's axe', *am a:mp–əm* [1SG dog–POSS.1SG] 'my dog' (Rie: 25, 28). The use of the free personal pronoun in such possessive constructions is not obligatory; however, it occurs frequently enough that it does not have an emphatic function (Rie: 28). When the modifying noun designates the material out of which something is made, then the attributive construction consists of two juxtaposed nouns, e.g. *sa:s kol* [birchbark house] 'house covered with birchbark' (Rmb: 44). Furthermore, the possessive suffixes can be omitted in cases when there is no emphasis ("логически не акцентируется") that specifically this item belongs to some other entity, e.g. *a:t'–am **wi:l't** se:mliγ wo:t–wes* [father–POSS.1SG face ?? wind(blow?)–PASS.PST.3SG[12]] 'my father's face darkened from the wind' (Rmb: 64).

The Samoyedic languages (which, like the Ob-Ugric languages, belong to the Uralic language family) have a Genitive case to express possessive attribution which can be reconstructed for Proto-Samoyedic[13] (Janhunen 1998a: 469; MSJ: 117ff). The functions of the Selkup Genitive case are to express adnominal modification, predominantly possession, but also material and part-whole relations (MSJ: 128ff). In the Central and Southern dialects, the possessum may agree with the possessor in constructions with nominal possessors, but this is not obligatory (MSJ: 82f, 130; Helimski 1998: 575). In possessive constructions with a pronominal possessor the possessum agrees in number and person with a first and second-person possessor, but not with a third-person possessor (MSJ: 78; OSJ: 187). Furthermore, older people speaking the Tym dialect, occasionally use possessive constructions consisting of a noun in the unmarked (Nominative) case followed by the head noun agreeing with the possessor, e.g. **era** *ala–gut menemba*[14] 'the old man's boat is old' (MSJ: 113); compare **era–n** *i:–də* [old.man–GEN son–POSS.3SG] 'the old man's son' (Helimski 1998a: 575)[15].

---

[12] In the only Mansi dictionary I was able to consult (Kuzakova 1994), *wo:t* is translated as 'wind', while the verb 'to blow (of wind)' is given as *wo:tx*. I assume that the final *–x* is dropped before the verbal suffix, but do not have any direct information on this.

[13] The issue of whether Proto-Uralic had a genitive case is debated, most especially since the Permic and Ugric languages lack it (Raun 1988: 558).

[14] It was unfortunately not possible to properly segment and gloss this example; however, *era* is 'old.man', *ala* is 'boat', and the *–t* in *–gut* could be the possessive suffix for the 3SG. *mene* might be a root with a meaning of 'old' [in the Selkupisches Wörterverzeichnis (Erdélyi [1970] 1997: 133), however, I was only able to find *mentäl* meaning 'old'], and the element *–mba* might be a verbal TAM marker with a meaning of '(past) narrative 3SG'.

[15] According to Ol'ga Khanina (pers.comm.) possessive constructions with an unmarked modifier are quite frequent in Samoyedic languages, so that this construction may not be restricted just to the Tym dialect of Selkup.

The Nganasan Genitive suffix that was observed by Castrén in the mid-19th
century has been lost in modern Nganasan; in the early 20th century it was already
going out of use, as evinced by a grammar published in 1937 (Ter: 76). In possessive
constructions, word order is strictly possessor – possessum (Ter: 94). The possessor
is suffixless; for some words, the form is identical to the unmarked (Nominative)
case, for others, an oblique stem is used, e.g. *nï ŋoybuo* 'woman's head' but *tundï
ŋoybuo* 'fox's head' (cf. *tuntï* 'fox.NOM'). The oblique stem is used for different
cases with different nouns, depending on whether the Nominative stem is vowel-
final or consonant-final (Ter: 61). In the plural all words have differing Nominative
and oblique stems, and in the possessive declension the possessive suffixes differ
between Nominative and Genitive, so that here a distinction between the Nominative
and Genitive is readily made (Ter: 76f). Notwithstanding the loss of the actual case
suffix, due to the distinctions in stems and possessive suffixes, Tereščenko still
defines a Genitive case for Nganasan (Ter: 61, 73, 78). This Genitive case is used to
express adnominal relations such as actual possession, kinship, body parts and
spatial relations; it is also used to express the person/animal for which something is
designated (Ter: 78). In cases of attribution in which the modifying noun expresses
the material from which the modified noun is made, the two nouns in Nominative
case are simply juxtaposed (Ter: 74). In possessive constructions with a nominal
possessor, the head noun is generally unmarked; however, there is a stylistic variant
that emphasizes the possessive relationship in which a nominal possessor in the
Nominative modifies a possessively-marked possessum, compare e.g. **dedi–tə lu:**
[father.OBL–GEN.2SG clothes] *t'etua xekutïə* and **desï–rə lu:–jï** [father.NOM–
POSS.2SG clothes–POSS.3SG] *t'etua xekutïə* 'your father's clothes are very warm'
(Ter: 98f). Free personal pronouns as possessors can modify possessively-marked
head nouns; however, this is optional (Ter: 95).

Table 3.4: Possessor marking in Eurasian languages

Possessor is unmarked	Possessum agrees with possessor	Possessor is marked	Possessum agrees with possessor
Sakha	Yes	Turkic	Yes
Tungusic	Yes	Eskimo	Yes
Evenki	Yes	(Evenki)	Yes
Yukaghir	No*	Mongolic	No
Nivkh	No	Chukchi	No
Khanty	Yes (pron.poss) No (nom. poss)	Itelmen	No

Table 3.4: Possessor marking in Eurasian languages, cont.

Possessor is unmarked	Possessum agrees with possessor	Possessor is marked	Possessum agrees with possessor
Mansi	Yes	Selkup	Yes (pron.poss.1+2) Yes/No (nom.poss)
		Nganasan	Yes (pron.poss) No (nom. poss)
Ket	Yes	Ket	No

*Although the possessum can optionally take a possessive suffix, this only indicates the possessive relation and does not agree with the possessor in person or number.

3.2.1.5 The loss of the Sakha Genitive in the light of Eurasian possessive constructions

From the above discussion and Table 3.4 it appears that the Sakha possessive construction, consisting of an unmarked modifier and a possessively-marked head, has much in common with possessive constructions in Tungusic languages and in Mansi, since all of these languages use an unmarked modifier, marking the head with a suffix that agrees in person and number with the possessor. Ket, too, has a similar construction, in which an unmarked possessor modifies a possessively-marked head, although the latter is marked by possessive prefixes, not suffixes. Although Yukaghir, too, can use a possessively-marked head in possessive constructions, this does not cross-reference the possessor and thereby differs from the construction found in Sakha. And in Nivkh, which also lacks a genitive case, possessive constructions mainly consist of unmarked juxtaposed nouns. On the other hand, Sakha clearly differs from the Turkic, Mongolic, Chukotko-Kamchatkan, and Samoyedic languages, since these all mark the possessive relationship on the modifier, either cross-referencing the possessor on the head in addition, or leaving the possessum unmarked.

Thus, one could argue for contact influence in the loss of the Sakha Genitive case, as done by Schönig ([1988] 1990: 50, footnote 43) and Stefan Georg (pers. comm.). Since historically speakers of Sakha have been in closer contact with speakers of Tungusic languages than with speakers of Mansi or Ket, a Tungusic source for the loss of the Sakha Genitive is the most plausible. However, given the fact that Evenki, the Tungusic language Sakha has been in most contact with, has an alternative possessive construction with a marked modifier in addition to the marked head (i.e. a construction very similar to that found in Turkic languages), this conclusion is not as straightforward as it might seem. The question of whether

Evenki may have retained a remnant of a formerly widespread Tungusic case, which was lost in most of the other languages, or whether Evenki independently innovated this marked possessor, is very difficult to resolve. Since Evenki uses the suffix –ŋi to mark the possessor in such constructions, which is formally identical to the Predicate Possessive suffix found in all Tungusic languages, one could argue that Evenki independently extended the use of this suffix to attributive position as well. However, even in this case it is not clear when this may have happened, before or after contact with Sakha. If this extension has happened only recently, then the loss of the Turkic Genitive in Sakha might be due to Evenki contact; if it happened at an earlier stage, this is not quite so plausible. An argument in favour of a fairly recent innovation of –ŋi in attributive position in Evenki is the fact that Negidal, a very closely related language, lacks it, as does Èven.

A possible explanation for the loss of the Sakha Genitive in Sakha is that 'frequential copying' (Johanson 1999: 52) or 'enhancement' (Aikhenvald 2002: 238) has taken place, whereby a construction that occurs only rarely in the recipient language becomes more frequent under contact influence. All Turkic languages have so-called izafet constructions in which the modifier stands in the unmarked Nominative case, while the head noun carries possessive suffixes. Even if this was copied into Turkic from Persian, the copying must have taken place at a very early stage, since even Old Turkic has such constructions; thus, the Turkic ancestor of Sakha may have already known such constructions as well. In Khalaj this construction is very frequent, and it has been suggested that this was frequent in Old Turkic, too (Drf: 76). Furthermore, in Old Turkic the izafet construction was not restricted to cases of generic possession (Erd: 381f), as it is in modern-day Turkic languages, further supporting the assumption of this construction as the origin of the Sakha possessive construction. One might therefore postulate a scenario in which the izafet construction was a relatively rare option in the Sakha ancestral language; under contact with Evenki, in which the possessor in possessive noun phrases was always or at least most often unmarked, it rose in frequency until this construction entirely replaced the erstwhile common construction with the Genitive case-marked possessor.

3.2.2 Extension of the Sakha Dative to include locative and allative functions

Sakha differs from the other Turkic languages (with the exception of Dolgan) in that it has lost the Common Turkic Locative case, the functions of which have been taken over by the Dative case[1]. This functional extension of the Dative case in Sakha has been attributed to either Mongolian or Evenki influence (Poppe 1959: 680; Nevskaya 2001: 299). In the following, I will present an analysis of these case functions in some Eurasian languages which will permit a more precise tracing of the contact influence.

3.2.2.1 The Dative case in Sakha

In Sakha, the Dative expresses various functions, among others recipients (17a) and benefactives (17b):

(17a)

ol	*kihi*	*mieχe*	*üčügeydik*	*ülele:bitim*	*ihin*
ol	kihi	mieχe	üčügey–LIk	üle–LA:–BIt–(I)m	ihin
that	person	1SG.DAT	good–ADVR	work–VR–PSTPT–POSS.1SG	for

paχvalnay.gramota	*bierbite*
paχvalnay.gramota	bier-BIt-(t)A
certificate.of.merit	give–PSTPT–POSS.3SG

'That man gave me a certificate of merit because I had worked well.'

[MalA138]

(17b)

oččoγo	*bihigi*	*eỹieχe*	*itinnik*	*kepse:n*
oččoγo	bihigi	eỹieχe	itin–LIk	kepse:–An
in.that.case	1PL	2SG.DAT	this.OBL–ADVR	tell–PF.CVB

bierebit	*duo*
bier–A–BIt	duo
BEN–IPF.CVB–1PL	Q

'So shall we tell you (our story) in this way?'

[Any28]

---

[1] The Common Turkic locative suffix has acquired a new meaning of partitive and indefinite accusative in Sakha, cf. section 3.2.3.

It furthermore fulfills locative functions, both spatially (18a, b) and temporally (18c):

(18a)

ol	*hir**ge***	*hïtammïn*		*ottu:bun*
ol	sir–GA	sït–An–BIn		ot–LA:–A–BIn
that	place–DAT	lie–PF.CVB–PRED.1SG		hay–VR–IPF.CVB–PRED.1SG

'I live in that place and make hay.'

[PotP108]

(18b)

ol	*ala:s**ka***	*törö:bütüm*
ol	ala:s–GA	törö:–BIt–(I)m
that	ala:s–DAT	be.born–PSTPT–POSS.1SG

'In that alaas (circular meadow in the forest) I was born.'

[MalA17]

(18c)

*kü**ŋŋe***	*tüörduoŋŋa*	*ti:ye*	*ölörör*
kün–GA	tüört uon–GA	ti:y–A	öl–(I)Ar–Ar
day–DAT	four ten–DAT	reach–IPF.CVB	die–CAUS–PRSPT

etim
e–TI–(I)m
AUX–PST–POSS.1SG

'I used to kill up to forty (ducks) in a day.'

[Chir181]

Finally, it covers allative and illative functions too, expressing motion towards (19a) and into a place (19b):

(19a)

illere:	*hïl*	*buollaɣïna*	*kuorak**ka** bara*		*hïjjïbïtïm*
illere:	sïl	buollaɣïna	kuorat–GA bar-A		sïrït–BIt–(I)m
before.last	year	however	town–DAT go–IPF.CVB		IPFV–PSTPT–POSS.1SG

'The year before last I went to town (Yakutsk).'

[Afny59]

(19b)

*umuhaχ**χa***	*ki:renner*	*...*	*arï:nï*	*ijje*	*bardïlar*
umuhaχ–GA	ki:r–An–LAr		arï:–(n)I	ilt–A	bar–TI–LAr
milk.cellar–DAT	enter–PF.CVB–PL		butter–ACC	carry–IPF.CVB	go–PST–PL

'They entered the milk-cellar and took away the butter […].'

[RaxA27]

3.2.2.2 The dative, locative and allative cases in Turkic languages

Old Turkic distinguishes three cases, Dative, Locative, and Allative (called 'directive', Erd: 171, 173 and 177). The Dative marks the goal or beneficiary of an action, and the object of emotions, e.g. **saŋa** *amran–mak–ïn … öl–ür men* [2SG.DAT love–INF–INS die–AOR 1SG] 'I die from love for you' (Erd: 368[2]). There are also some instances of the Dative marking direction, e.g. *teŋri–ler … tužit ordo–**ka** yïg–ïl–ur–lar* [god–PL … Tusita palace–DAT gather–PASS–AOR–PL] 'the gods assemble at the Tusita (one of the heavens) palace' (Erd: 366). The difference between the directive function of the Dative and the Allative case is that the latter marks movement towards a goal, while the Dative is used "when the goal is reached (or is meant to be reached)" (Erd: 366ff). The Allative case marks the goal of motion, e.g. *ben ev–**gerü** tüš–eyin* [1SG house–ALL descend–VOL] 'let me (go) home and dismount' (Erd: 370). It is frequent in Orkhon Turkic, and also fairly frequent in Manichaean Uyghur texts, but is rare in later stages and absent from the 11th century onwards (Erd: 370f). The Locative expresses stative location, e.g *ol ev–**de*** [that house–LOC] 'in that house'. Furthermore, it can be used with verbs of motion if the result of the action is a state (e.g. in expressions such as 'entering into the divine country', i.e. 'dying'), and it also has ablatival functions (Erd: 371f), e.g. *ay teŋri ordo–sïn–**ta** ėn–ipen*[3] [moon god palace–POSS.3SG–LOC descend–CVB] 'coming down from the palace of the Moon God' (Erd: 372). The distinction between the ablatival function of the Locative case marker *–dA* and the Ablative case marker *–dAn* is that *–dA* is used to refer to sources "from which the subject merely separates or keeps apart" while *–dAn* is used to express actual "physical movement away from a source" (Erd: 375).

Turkish has lost the Allative case, the functions of which have been taken over by the Dative case, e.g. *Hasan Ankara–**ya** git–ti* [Hasan Ankara–DAT go–PST] 'Hasan went to Ankara' (Krn: 242, 243). The Dative also expresses recipients, e.g. *Hasan kitab–ı Ali–**ye** ver–di* [Hasan book–ACC Ali–DAT give–PST] 'Hasan gave the book to Ali' (Krn: 220) and can mark benefactives (Krn: 226). The Locative case marks location, e.g. *kitap masa–**da*** [book table–LOC] 'The book is on the table' (Krn: 242). Similarly, in Turkmen (which like Turkish is an Oghuz language), the Dative case has extended its function to encompass those of the Old Turkic Allative case,

---

[2] Note that Erdal gives this example on both p. 368 (in the section on the functions of the Dative case) and on p. 378 (in the section on the functions of the Instrumental), cf. section 3.2.4.2.

[3] The vowel here transcribed as *ė* is transcribed by Erdal as *e*, and the vowel I transcribe as *e* is transcribed by him as *ä*. The former developed out of a long *e/ä* (Erd: 50), while the latter is the front counterpart of *a*.

the Dative expressing both motion towards and into an entity as well as indirect objects (Clk: 126ff).

This pattern of case-marking can also be found in the Kypchak languages Kazakh and Kirghiz, where the Dative case expresses allative functions in addition to marking indirect objects (Somfai Kara 2002: 19; SKJ: 163f; Somfai Kara 2003: 18f). The Dative in Kazakh expresses both recipients and benefactives, as well as the goal of motion, e.g. *bul kishi–ge ber* [this person–DAT give[IMP.2SG]] 'give it to this person', *üy–ge bar–a–mïn* [house–DAT go–IPF.CVB–PRED.1SG] 'I go home' (Somfai Kara 2002: 19; SKJ: 163f). The Locative case marks stative location at, on, or in, e.g. *men üy–de–min* [1SG house–LOC–PRED.1SG] 'I am at home' (SKJ: 168f).

In Uzbek, a language belonging to the Southeastern Turkic group, the Dative case has also taken over allative functions, expressing amongst others an indirect object and also a spatial goal, e.g. *toγa–m ötin ol–gali örmon–ga ket–di* [uncle–POSS.1SG firewood take–PURP forest–DAT go–PST] 'my uncle went to the forest to bring some firewood' (Bdr: 113). The Locative case expresses spatial and temporal location, e.g. *opa–m Farγona–da tur–a–di* [older.sister–POSS.1SG Farghona–LOC stand–IPF.CVB–3SG] 'my elder sister resides in Farghona' (Bdr: 127).

The South Siberian Turkic languages Khakas and Tuvan (which, together with Sakha, are classified as the Northeastern Turkic group – Johanson 1998b: 83) differ from the languages spoken to the west in that they have retained the Old Turkic case distinction between Dative, Allative and Locative. Amongst other functions, the Khakas Dative case marks recipients, e.g. *ol kniga–nï maγa: pir–dI* [3SG book–ACC 1SG.DAT give–PST] 'he gave me the book' (And: 8). The Khakas Locative marks simple locations, e.g. *pIs Xakasiya–da čurta–pča–bïs* [1PL Khakasia–LOC live–PRS.I–1PL] 'we live in Khakasia' (And: 11), while the Allative case has only one function, which is to mark direction towards a point, e.g. *min Asχïs–sar par–i–m* [1SG Askiz–ALL go–PRS.III–1SG] 'I am going to Askiz' (And: 13).

Tuvan also distinguishes between a Dative, Locative and Allative case (A/H: 14). As in the other Turkic languages, the Dative marks the recipient or beneficiary of an action, while the Locative expresses location in, at, or on a place or object (A/H: 18 and 20, respectively). However, interestingly enough, the Locative case in Tuvan is temporally restricted, expressing location only in the present tense, e.g. *men Kïzïl–da čurtta–p tur men* [1SG Kyzyl–LOC live–CVB AUX 1SG] 'I live/am living in Kyzyl' (A/H: 20). In the past tense, location is expressed by the Dative case, e.g. *men (...) Kïzïl xo:ray–ga törtön–gen men* [1SG … Kyzyl city–DAT be.born–PSTPT 1SG] 'I was born in Kyzyl…' (A/H: 21). The Allative case marks direction towards a location, e.g. *da:rta men Kïzïl–je azï Abakan–če čor–u:r men* [tomorrow

1SG Kyzyl–ALL or Abakan–ALL go–AOR 1SG] 'Tomorrow I'm going to Kyzyl or Abakan' (A/H: 17).

Tofa, however, another Turkic language spoken in Southern Siberia, resembles most Turkic languages in that it has lost a separate Allative case, marking both recipients and direction with the Dative case, e.g. *bis Moskva:–ga čor–a:n bis* [1PL Moscow–DAT go–PSTPT 1PL] 'we went to Moscow', *oŋ hine:k–tï o:l–ga ber–di* [3SG book–ACC boy–DAT give–PST.3SG] 'he gave the book to the boy' (Ras: 38). The Locative case expresses location in both space and time, e.g. *ol hire–de ög–de kum ta yok bol–gan* [that time–LOC house-LOC who PTL non-existence AUX–PSTPT] 'at that time nobody was home' (Ras: 39).

As in Sakha, the Dative in Dolgan expresses recipients, spatial and temporal location, as well as the goal of motion (Ubr: 119). Furthermore, Ubrjatova (Ubr: 124f) suggests that Dolgan may be developing a new Allative out of the postposition *diegi* 'in the direction of' (20a, 20b). At the time of Ubrjatova's studies (the material presented in the grammar was collected in the 1930s), this was still a clitic[4] *dek ~ det*, but she suggests that it might develop into a full-fledged suffix.

(20a)

ha:	*tïah–ïn=**det***	*bar–an*	*is–pit*
gun	sound–ACC.3SG=ALL	go–PF.CVB	walk–PSTPT

'(S)he went in the direction of the gunshot.'

(20b)

bo	*ogo*	*giniler=**dek***	*kör–böt*
this	child	3PL=ALL	look–PRSPT.NEG

'This child does not look in their direction.'

[Ubr: 125]

However, in recently published Dolgan short stories (Popov 2000) the Allative postposition *diek* is still written as a separate word, implying that the grammaticalization process to a case suffix has not yet been completed. It is not clear whether this orthography reflects actual usage (i.e. that the postposition or clitic actually lacks vowel harmony, is not as reduced as indicated by Ubrjatova, and does not form part of the phonological word), or whether it is due to a prescriptive tendency in language policy makers, who might try to make Dolgan look more similar to Sakha, while the actual spoken usage may be different. According to Marek Stachowski (pers. comm.), the Allative postposition has most definitely not grammaticalized to a suffix yet, as shown by its lack of vowel harmony, and by the

---

[4] I consider this a clitic, since it cannot be used independently and is pronounced as one unit with the head noun (Ubr: 125), even though it is not subject to vowel harmony.

fact that *dieki/diegi* still functions as a noun, e.g. *bu diegi* 'this side', not just 'in this direction'. The above examples from Dolgan do not differ markedly from their Sakha counterparts; (20a) would be *kini sa: tïah–a:–bït sir–in **dieki** bar–da* [3SG gun sound–VR–PSTPT place–ACC.3SG in.the.direction go–PST.3SG] '(S)he went in the direction of the gunshot', while (20b) would be *bu oγo kini–ler **dieki** kör–böt* [this child 3SG–PL in.the.direction see–NEG.PRSPT] 'This child does not look in their direction' (both translated, Ver).

Khalaj has retained the Old Turkic Dative suffix *–KA*, which marks indirect objects, recipients, addressees, and also goals and directions, i.e. the Dative fulfills allative functions in addition to marking indirect objects (Drf: 80, 82). The Khalaj Locative is marked by the suffix *–čA*, which corresponds to the Old Turkic Equative case marker. The Locative expresses stative location, but also a goal where one plans to spend a longer time (Drf: 88f).

There are discrepancies between different descriptions of the Chuvash case system regarding the Dative and Allative cases. Thus, Benzing ([1942] 1993: 52ff) distinguishes between an Accusative and Dative case in Chuvash, even though the case suffixes are identical. Other authors (Krueger [1961] 1997: 102; Clark 1998b: 438; Fedotov 1996: 195f) collapse these into one case, called the 'Objective' (Krueger [1961] 1997: 102) or the 'Dative-Accusative' (Clark 1998b: 438). This case suffix marks recipients and benefactees, as well as definite direct objects (Krueger [1961] 1997: 103). Furthermore, Benzing and Clark do not include an Allative case in the Chuvash system, while Krueger ([1961] 1997: 111) and Fedotov (1996: 194f) do. There is consensus on the presence of a separate Locative case in Chuvash, which expresses both spatial and temporal stative location (Benzing [1942] 1993: 56). See Table 3.5 for a summary of the case markers and their functions in Turkic languages.

3.2.2.3 The dative, allative and locative cases in Tungusic languages

The Tungusic languages have a number of local cases that generally express stative location, motion towards, motion along, and motion from. The traditional names for the local cases (judging from the short sketches in *Jazyki Narodov SSSR 5*) are: Locative, Directive, Prolative, Ablative, Elative, Directive-Locative and Directive-Prolative (although the latter two seem to be an innovation in Evenki and Ėven). For Evenki, Bulatova & Grenoble (B/G: 8) follow the traditional terminology with the exception that they substitute the term Allative for the term Directive. Nedjalkov (Ned: 142) calls the traditional Tungusic Locative case the Allative and the traditional Directive the Locative-Allative; the Directive-Locative he calls

Locative-Directive and the Directive-Prolative he calls Allative-Prolative. Given the fact that the Tungusic Locative case markers are cognate amongst all the languages (cf. Appendix 4), I will continue to use the term 'Locative' to designate the cases expressed by this marker in Tungusic languages, notwithstanding the fact that in Evenki the functions have shifted to an allative case marker (see below). The case termed 'Directive' in *Jazyki Narodov SSSR 5* I will term Allative to fit in with the general terminology chosen in this section. The Directive-Locative and Directive-Prolative cases appear to be a fusion of the Allative with the Locative and Prolative cases, respectively; they are rarely used in either Evenki or Ėven (B/G: 12; Ned: 171; Ben: 61f; Mal: 11) and are absent in other Tungusic languages and will therefore not be discussed further.

In Evenki, the Dative expresses recipients and benefactives (Ned: 148, 153), e.g. *purta–va–s min–du bu:–kel* [knife–DEF.ACC–POSS.2SG 1SG–DAT give–PRXIMP.2SG] 'give me your knife'. It also fulfills general locative functions, expressing spatial location, regardless of the size of the object or orientation of the location, e.g. *Turu–du* [Tura–DAT] 'in Tura', *oron–du* [reindeer–DAT] 'on a reindeer' (Ned: 169); this is one of the most frequent uses of the Dative (B/G: 9). The Dative can also express temporal location, e.g. *jur–du chas–tu* [two–DAT hour–DAT] "at two o'clock" (Ned: 181), and is used to designate the adressee of verbs of speech (although the Allative can fulfill this function, too – B/G: 10). The Dative can also designate the direction of certain verbs, e.g. *ne:–* 'put (down)', *iniv–* 'load', *tu–* 'step on' (Ned: 170), e.g.

(21)

beyetken	*gara–**du***	*tu–re–n*
boy	twig–DAT	step–NFUT–3SG

'The boy stepped on a twig.'

[Ned: 170]

Furthermore, as mentioned above, the old Locative case in Evenki has shifted to an allative meaning; therefore Evenki now has two cases that express 'motion towards'. The Locative case expresses motion directed up to a point (B/G: 10; Kon: 52), while the Allative expresses the goal of a motion event that need not reach its destination (Ned: 170; B/G: 10; Kon: 51); compare:

(22a)

atïrka:n	*ju–**la**–vi*	*ŋene–jere–n*
old.woman	house–LOC–PREFL	go–PRS–3SG

'The old woman is going to her house.'

(22b)

> *atïrka:n   ju–**tki**–vi   ŋene–jere–n*
> old.woman   house–ALL–PREFL   go–PRS–3SG

'The old woman is going in the direction of/towards her house.'

<p align="right">[Ned: 170]</p>

Ėven distinguishes Dative, Locative, and three allative cases: 'Allative', 'Directive-Locative' and 'Directive-Prolative' (Ben: 58ff). The Dative expresses recipients and beneficiaries, e.g. *Iβan Anna–**du** olra–β əmu–n* [Ivan Anna–DAT fish–DEF.ACC bring[NFUT]–3SG] 'Ivan brought fish to Anna', *Semen artel–**du** momi–β ayŋur–ra–n* [Semen cooperative–DAT boat–ACC clean–NFUT–3SG] 'Semen cleans the boat for the cooperative' (Ben: 58). However, the Dative case also has some locative functions, e.g. *bi ju–**du** bi–se–m* [1SG house–DAT be–NFUT–1SG] 'I am in the house' (Ben: 58). According to Malchukov (Mal: 10), these are restricted to inanimate nouns with animate agent to indicate stative location, e.g. *bey aβlan–**du** ilat–ta–n* [man glade–DAT stand–NFUT–3SG] 'the man stands on the glade'. However, Novikova (NovI: 201) gives examples where the Locative case attaches to animate nouns as well, e.g. *atïkan–**du**–s anŋa–t–ti–βun* [old.woman–DAT–2SG spend.the.night–RES–PST–1PL.INCL] 'we spent the night at your wife's (мы ночевали у твоей жены)', *ečin bi eŋeje–**du** bi–si–β* [thus 1SG rich–DAT be–PST–1SG] 'thus I lived with the rich man/at the rich man's place' (NovI: 201). Possibly this is due to the fact that the nouns designating people are here used metonymically. The Locative expresses spatial and temporal stative location, e.g. *uliki hekita–**la** bukče:ŋči–n* [squirrel tree–LOC sit.hunched.up[NFUT]–3SG] 'the squirrel sits on the tree' (Ben: 60). When used with telic motion verbs, it can also express movement towards a goal, similar to the Allative, e.g. *bi ju–**la** em–re–m* [1SG house–LOC come–NFUT–1SG] 'I came into the house' (Ben: 60; NovI: 211; Mal: 10). The Allative expresses the direction or goal of an action, e.g. *bi umen–**teki** bey–**teki** ŋen–ri–β* [1SG one–ALL man–ALL go–PST–1SG] 'I went to one man' (Ben: 62), *nam–**teki** nulge–d–de–p* [sea–ALL move–PROG–NFUT–1PL.EXCL] 'we are migrating to the sea' (NovI: 216).

In Udihe, the Dative case marks the benefactive/malefactive, recipient, goal, and direction (N/T: 123), e.g. *bi sin–**du** xələba–wa bu–o:–mi* [1SG 2SG–DAT bread–ACC give–PST–1SG] 'I gave you (some) bread' (N/T: 524). It also "commonly expresses a local meaning. In this function it may be interchangeable with the Locative" (N/T: 123), e.g. *bi ənimi kluba–**du** ətətə–ini* [1SG mum club–DAT work–3SG] 'my mum works in the club' (N/T: 522). The difference in use between the Dative in its locative function and the Locative is that the Dative seems to be preferred for a meaning 'in', while the Locative is preferred to express 'at' or 'on', e.g. *uli–**du*** [river–DAT] 'in the river' vs. *uli–**lə*** [river–LOC] 'at the river' (N/T: 125).

The Locative case, in addition to expressing stative location, e.g. *Udiə namuŋka–da namu–lə bagdi:–ti* [Udihe Oroch–FOC sea–LOC live–3PL] 'Udihe and Oroch people live on the sea' (N/T: 521), can also have a directional meaning, in which it overlaps with the Allative case (called 'Lative' by Nikolaeva & Tolskaya). The difference in directional meaning between the Allative and the Locative is that the Allative expresses direction towards the object, while the Locative expresses movement into something (N/T: 124), e.g. *zaŋe–ziga Moskwa–**tigi** ŋənə–zəŋə–ti* [boss–PL Moscow–ALL go–FUT–3PL] 'the bosses will go to Moscow', *zugdi–**lə** i:ŋ–ki–ni* [house–LOC come–PST–3SG] 'he entered the house' (N/T: 517, 522).

In Nanay, the different locative cases have various functions, the most important of which will be briefly mentioned here. The Dative case expresses stative location, temporal location, and recipients, as well as the logical subject of verbs of emotion, e.g. *daŋsa–sal xorgo–**du** bi–l* [book–PL cupboard–DAT be–3PL] 'the books are in the cupboard', *basa ayŋani–**du** mi xoton–či ene–de–mbi* [next year–DAT 1SG town–ALL go–FUT–1SG] 'next year I will go to town', *ey daŋsa–va učitel'–**du** bu:–xeri* [this book–ACC teacher–DAT give–DSTIMP.2SG] 'give this book to the teacher' (AvrI: 180f, AvrIII: 167ff). The Locative case expresses the final point of motion, location (predominantly in conjunction with postpositions), a limit, source, and movement into as well as out of something, e.g. *nu:čike:njuen duente–**le** tutu–xe–či* [children forest–LOC run–PST–3PL] 'The children ran (up) to the forest', *jo: jakpa–**la**–ni mo baldi*[5] [house near–LOC–3SG tree grow] 'near the house a tree is growing', *tey mapa gogakta–ni omol–**dola**–ni baldi–xa–ni* [that old.man beard–3SG belt–LOC–3SG grow–PST–3SG] 'the beard of that old man has grown up to his belt', *mi xorgo–**la** daŋsa–va japa–xa–mbi* [1SG cupboard–LOC book–ACC take–PST–1SG] 'I took the book out of the cupboard' (AvrI: 181f; AvrIII: 173ff), *tey mo–**la** i:gu–xe–n* [that tree–LOC go.into–PST–3SG] '(s/he) went into that tree' (AvrIII: 180). The use of the Nanay Locative to express stative location is rare (AvrIII: 178). The Allative case expresses the endpoint or goal of movement, motion into something, and the addressee of verbs of speech, e.g. *sikse bue klube–**či** ene–y–pu* [evening 1PL club–ALL go–NPST–1PL] 'In the evening we'll go to the club', *alo:simji nu:čike:njuen–**či** gisure–y–ni* [teacher children–ALL tell–NPST–3SG] 'The teacher tells the children' (AvrI: 182; AvrIII: 182ff).

The Oroč Dative case expresses spatial as well as temporal location, e.g. *siŋe me:ne ju:g–**du**–i ba:gdi–xa–ńi* [mouse self house.OBL–DAT–3SG live–PST–3SG] 'the mouse lived in her own house', *sa:gdańi–**du** Be:či–du kiaka bi–či–ti* [old.times–DAT Beči–DAT Udihe be–PST–3PL] 'in times long past (в давние времена), Udihes lived

---

[5] This is identical to the root; it is not clear whether the lack of further tense and person marking is due to a typo in the original (AvrI: 182), or whether this is the correct form.

on the river Beči' (A/B: 124, 125). The Dative case also marks recipients, e.g. *amin–du–ń=da: eńin–du–ń=de: xite–du–ń=de: bu:–xe–ti* [father.OBL–DAT–3SG=EMPH mother.OBL–DAT–3SG=EMPH child–DAT–3SG=EMPH give–PST–3PL] '(they) gave to his father, and to his mother, and to his children' (A/B: 126). The Locative case expresses the direction of motion, e.g. *tada–ńi uli–la ti:–xe–ńi* [arrow–3SG river–LOC fall–PST–3SG] 'the arrow fell into the river', *Asiŋkan–dula bua–la–xa–ti Asiŋkan–du bi–či–ti* [Asinkan–LOC place–VR–PST–3PL Asinkan–DAT be–PST–3PL] 'they moved to Asinkan and lived in Asinkan' (A/B: 128). It can also express location, e.g. *Kepin–dule mu:de ana maŋga e–či–ńi bi–ye* [Koppi–LOC flood NEG strong NEG–PST–3SG be–CVB] 'the flood in Koppi was not very heavy', *asa–muna–ka bi–či–ti Xutu–va soli–la O:čoŋko–du* [woman–DER–?? be–PST–3PL Xutu–ACC upper.reaches?–LOC Očonko–DAT] 'the spouses lived in the upper reaches of the river Xutu, in Očonko' (A/B: 129). The Allative case expresses the direction of movement, e.g. *bi: ju:k–ti–mu jima–ya* [1SG house.OBL–ALL–1SG visit–IMP.2SG] 'come into my house!', *bi: a:kin–ti–si xuli–xe–mi* [1SG older.brother.OBL–ALL–2SG go–PST–1SG] 'I went to your older brother' (A/B: 134). It also marks the adressee of verbs of speech and emotions, e.g. *edi–ti–vi guŋ–ki–ńi* [husband.OBL–ALL–REFL say–PST–3SG] 'she said to her husband', *ta:du edi–ńi asan–ti–vi tagda–xa–ńi* [there husband.OBL–3SG wife.OBL–ALL–REFL get.angry–PST–3SG] 'then the husband got angry with his wife' (A/B: 135).

In contrast to the other Tungusic languages, Manchu has a very restricted case system consisting only of Nominative, Genitive, Accusative, Dative, and Ablative (Gor: 163). Thus, it is not surprising that the Manchu Dative case encompasses a large number of functions, among them locative and allative. It expresses recipients, addressees of verbs of speech, instruments in conjunction with past tense verbs, destination of a movement, as well as location in space and time. The marker is not subject to vowel harmony and is mostly written separately from the noun it marks (Gor: 182). Examples of the main functions of the Manchu Dative are: *ere niyalma de bu–mbi* [this man DAT give–IPFV] '(someone) gives to this man' (Gor: 183), *beye–i gala–de jafa–habi* [self–GEN hand–DAT catch–PST] '(someone) caught with his hand', *alin bujan–de tomo–mbi* [mountain forest–DAT live–IPFV] '(they) live in mountains and forests' (Gor: 185), *tere nergin–de* [that time–DAT] 'at that time' (Gor: 186), *tugi de sucuna–me deye–mbi* [cloud DAT soar–CVB fly–IPFV] '(they) fly soaring towards the clouds', *ce boo–de isinji–ha* [3PL house–DAT reach–PTCP] 'they reached the house' (Gor: 186).

As can be seen from the above descriptions, the Dative case in Tungusic languages commonly expresses not only recipients, but also location, while the Locative case expresses not only location, but also direction towards an object.

Evenki differs slightly from the other Tungusic languages in that here the Locative case has completely lost its function of expressing location; this has been entirely taken over by the Dative case. This shift in case functions in Evenki may well be due to contact influence from Mongolian or Sakha, as both languages mark both recipients and locations with the Dative case, as described in section 3.2.2.1 and 3.2.2.4. However, an internally-motivated extension of case functions that were present already may also have taken place.

3.2.2.4 The dative, locative and allative cases in Mongolic languages

In Proto-Mongolic, there appear to have been two case forms *–dU/r* (the Dative case) and *–A* (the Locative case) which both fulfilled identical functions of expressing location in space and time, direction, as well as recipients. The Locative case is attested in Written Mongol and Middle Mongol, but in modern-day languages it is restricted to adverbs and postpositions (Janhunen 2003d: 15). Thus, already at a very early time the Mongolic languages did not formally differentiate between dative, locative and allative functions, using the same case marker(s) to express these.

In Khalkha Mongolian the Dative fulfills the standard dative functions of expressing recipients and benefactives (23a). It also marks simple location in space (23b) and time, as well as expressing the goal of motion (23c) (K/Ts: 84f), e.g.

(23a)

  Dulma:–d cay   ög
  D.–DAT   tea   give[IMP]

  'Give Dulmaa some tea!'

(23b)

  ted ger–t   suu–dag
  3PL yurt–DAT   sit–HAB

  'They live in a yurt.'

(23c)

  ter   Mongol–d   ir–sen
  3SG   Mongolia–DAT   come–PSTPT

  'He came to Mongolia.'

[K/Ts: 84]

In Buryat, too, the Dative case marks recipients and location in space and time, as well as expressing the goal of motion (PopB: 36; Skr: 109). In Dagur the Dative also expresses a location or direction of an action in space and time, as well

as recipients and benefactives (Tsum: 138, Todaeva 1986: 42). Interestingly, Todaeva also describes a separate Limitive case which expresses the limit up to which an action is performed, e.g. *ende so:–ča:r guen os bei* [here armpit–LIM deep water be] 'here the water reaches to the armpits'. This case, however, is rare (Todaeva 1986: 45), which might explain why neither Tsumagari (Tsum: 136ff) nor Martin ([1961] 1997: 23ff) describe it.

For Khamnigan Mongol Janhunen (1990: 51ff) mentions five suffixally marked cases in addition to the unmarked Nominative; these are the Accusative, Genitive, Dative, Ablative and Instrumental. Unfortunately, there is no discussion of the functions of these cases. Since for Khamnigan Evenki Janhunen (1991: 60ff) mentions ten cases (Nominative, Partitive, Accusative, Genitive, Dative, Ablative, Locative, Elative, Directive, and Instrumental) with two 'local cases' (Dative and Ablative; Janhunen 1991: 62) and three 'spatial cases' (Locative, Elative, and Directive; Janhunen 1991: 64), it may be assumed that the influence of the two languages on each other has not greatly interfered with the case system.

Table 3.5 summarises the uses of the dative, allative and locative cases in the languages discussed above.

Table 3.5: The functions of the dative, allative, and locative cases in Turkic, Tungusic, and Mongolic

	Dative	Allative	Locative
Old Turkic	Goal, beneficiary	Direction	Stative location
Khakas	Recipient	Direction	Stative location
Tuvan	Recipient, stative location[*]	Direction	Stative location[*]
Tofa	Recipient, direction		Stative location
Turkish	Recipient, beneficiary, direction		Stative location
Kazakh	Recipient, beneficiary, direction		Stative location
Uzbek	Recipient, beneficiary, direction		Stative location
Khalaj	Recipient, beneficiary, direction		Stative location
Chuvash	Recipient, direction	Direction	Stative location
Sakha	Recipient, direction, stative location		
Dolgan	Recipient, direction, stative location		
Evenki	Recipient, stative location	Direction	Direction
Èven	Recipient, (stative location)	Direction	Stative location
Udihe	Recipient, (stative location)	Direction	Stative location, motion into
Mongolic	Recipient, direction, stative location		

[*]In Tuvan, the Dative has locative functions only in the past tense. In the present tense, these functions are fulfilled by the Locative.

3.2.2.5 The dative, allative and locative cases in other Siberian languages

In Yukaghir, which lacks a separate allative case, both the Dative and the Locative can express direction towards an object (Mas: 96f, 105f). The Dative case marks recipients, benefactives and malefactives, as well as addressees of verbs of speech (Mas: 96ff, 352f), e.g. *tintaŋ pulut mon–i šaqale–ŋin* [that old.man say–INTR.3SG fox–DAT] 'that old man said to the fox' (Mas: 97); and it can express the purpose (24a) or the direction of movement (24b, c) (Mas: 98, 356), e.g.

(24a)

taskan	*laŋi*	*kebe–y–te–yl'i*	*čay–**ŋin***	*tabaq–**ŋin***
T.	towards	go–PFV–FUT–INTR.1PL	tea–DAT	tobacco–DAT

'We'll go to Taskan for some tea and tocacco.'

(24b)

mon–delle	*pud–o:–l*	*lebie–**ŋin***	*qoy–pe–**ŋin***	*qon–i*
say–PF.CVB.SS	top–VR–ANR	earth–DAT	god–PL–DAT	go–INTR.3SG

'Having said (it), he went to Upper Earth, to the gods.'

(24c)

*šoromo–**ŋin***	*čuöte*	*me+eyre–y*
person–DAT	always	AFF+walk–INTR.3SG

'He would always come to that person.'

[Mas: 98]

The Locative case expresses stative location (both 'in' and 'at', 25a) as well as direction towads a location (25b) (Mas: 105, 356); it is common when the noun referring to the location or destination designates people (25c) (Mas: 106).

(25a)

taŋ	*yalhil–**ge***	*irk–in*	*šu:ke–die*	*eyre–y*
that	lake–LOC	one–ATTR	pike–DIM	walk–INTR.3SG

'In that lake lived one pike.'

(25b)

taŋ	*nume–**ge***	*šöy–l'el*
that	house–LOC	enter–INFR.3SG

'He entered that house.'

[Mas: 105, 106]

(25c)

*taŋ omni:–pe–**ge** yoq–to–lle tude–gele ta:*
that family–PL–LOC approach–CAUS–PF.CVB.SS 3SG–ACC there

pöń–i:–l'el–ŋa:
remain–CAUS–INFR–3PL.TR

'…(they) brought him to those people and left him there.'

[Mas: 107]

Verbs of movement can take directional arguments marked with either the Dative or the Locative case or by a postposition; the choice between the two cases may reflect some semantic nuances inherent in the cases rather than being grammaticalized in any way (Mas: 356f).

A look at other Siberian languages shows that the Turkic languages follow a common pattern in having only one case to express both dative and allative functions, and having a separate case to express stative location (cf. Table 3.6).

In Itelmen, the Locative case predominantly marks stative location, both spatial and temporal (G/V: 74). The precise spatial location can be further defined through a postposition; in this case, both the noun and the postposition take Locative case marking, e.g. *iy–**enk** çəs–**k** əńç–e ʔn suńḷ–qzu–z–e ʔn* [water–LOC inside–LOC fish–PL live–IPFV–PRS–3PL] 'fish live in the water' (G/V: 74). The main function of the Itelmen Dative case (called 'Dative-Directive' by Georg & Volodin) is to mark directions (26a), both spatially and temporally (expressing the point in time up to which an action is performed). It can furthermore mark the purpose of an action, similar to Yukaghir, and it expresses recipients (26b) and addressees of verbs of speech (G/V: 76f).

(26a)

*t–ç'e–kiçen kest–**anke***
1SG>enter<1SG house–DAT

'I went into the house.'

(26b)

*Ememqut k–zil–ʔin ᵒplah ᵒksalx ənnan–**ke***
E. INF.III>give<INF.III big bladder 3SG–DAT

'Ememqut gave him a big bladder.'

[G/V: 76, 77]

In Chukchi, the Locative case is used to mark general spatial location without further semantic specification, but it is also used when an adverb or postposition specifies the location (Dnn: 116, 274). In addition, the Inessive case marks location inside an object, and the Sublative case marks location underneath an object (Dnn:

279). The Allative case marks movement towards an object (Dnn: 274). It can also mark third person recipients and addressees of verbs of speech, which is why Skorik (1961: 164) calls it the 'Dative-Directive' case; however, the dative function is restricted to animate nouns. According to Dunn (Dnn: 275), this case does not mark first or second person pronominal recipients of the verb 'give', which are cross-referenced on the verb in object position, with the object given appearing as a separate unmarked noun. For the pronouns, Skorik (1961: 406) analyzes two separate cases, a Dative and an Allative, possibly based on the fact that there exist two suffixes expressing a directional and dative meaning. Dunn, on the other hand, analyzes the two suffixes as being in free variation and belonging to one Allative case (Dnn: 102f), an analysis which coincides with Skorik's observation that the Dative and Allative cases are not clearly delimited ("Направительный и дательный падежи разграничены недостаточно четко." Skorik 1961: 408, footnote 298).

In Siberian Inupik Eskimo the predominant function of the Lative ('Dative-Allative' in Menovščikov's description) case is to mark the direction of an action. It also marks the point in time up to which an action is performed, the purpose of an action as well as addressees of verbs of speech and emotions (Men: 52). In West Greenlandic, indirect objects can be expressed by the Allative case; however, with certain verbal forms this function can also be fulfilled by the Absolutive (Fortescue 1984: 88f). The Locative case in Siberian Inupik expresses stative location in space and time (Men: 54).

The Ket Dative case marks indirect objects and the goal of motion (Wer: 113), while the Benefactive marks benefactees and also the object of verbs of speech and mental processes, e.g. 'sing about', 'read about' (Wer: 114). Dul'zon (Dul: 76) calls the Dative case 'Dative-Directive' and claims that its predominant function (in approximately 90% of the instances he analyzed) is to mark the direction of an action, especially if the marked noun is inanimate. With animate nouns it marks recipients and the addressee of verbs of speech (Dul: 77). Occasionally, it can have a locative function, and with nouns denoting periods of time it expresses the limit in time up to which an action is performed. The Locative case is restricted to inanimate nouns; it expresses a stative location, and in the northern Ket dialects it can express temporal location as well (Wer: 114f, Dul: 79f). Occasionally, the Locative can have an allative function, so that its function overlaps with that of the Dative (Dul: 81). In its function of expressing spatial location the Locative can often be replaced by the Adessive case; this, however, shows no restriction as to the animacy of the nouns it occurs with; Dul'zon calls this case the 'Personal Locative' (Wer: 114, Dul: 81). In addition to its locative function the Adessive case can also express 'have' possession (Wer: 103).

Nivkh differs from the other Siberian languages in that case functions are distributed amongst several cases in a rather unusual way. In this discussion I follow the terminology of Mattissen (Mat: 9) rather than that of Gruzdeva (Grz: 18ff). The (unmarked) Nominative case can express an indirect object, e.g. *ïtïk p^h–meot'u ń–ïkïn k^him–j* [father REFL–gun 1SG–elder.brother give–FIN] 'father gave his gun to my elder brother' (PanI: 127), *ńi čo hï ńivx ar–j* [1SG fish this man feed–FIN] 'I fed this man with the fish' (Grz: 18). The Allative (called 'Dative/Additive' by Gruzdeva: 18, 20) marks both the direction of an action or movement, an addressee of verbs of speech, as well as the final point of movement, e.g. *tï ńivɣ–a t^ha ïkïn–doχ t^haxta–ya* [this man–VOC not elder.brother–ALL be.angry–IMP] 'this man, don't be angry at (your) elder brother', *ńin–doχ p^h–vo–x t^hamdid xer–ya* [1PL–ALL REFL–village–ABL what tell–IMP] 'tell us, what (is going on) in your village!' (Grz: 20). Some verbs can take an allative argument standing either in the unmarked (Nominative) case or in the Allative case, e.g. *ïtïk p^h–ray–rox tamx si–j* [father REFL–pipe–ALL tobacco put–FIN] and *ïtïk tamx p^h–ray si–j* [father tobacco REFL–pipe put–FIN] 'father put tobacco into his pipe' (PanI: 127). Note the difference in word order, since in a sentence with several arguments in the unmarked (Nominative) case the order is S–DO–IO–V. The meaning of spatial and temporal limit is also expressed by the Terminative case (called 'Limitative' by Gruzdeva: 18, 21); this has a narrower range of meaning than the Allative, however, expressing only the endpoint of time or motion or the timespan within which an event takes place (PanI: 142, 144), e.g. *ńi eri–rχa vi–nï–j–ra* [1SG river–TRM go–FUT–FIN–PRD] 'I shall go up to the river', *č–ïtk haimŋaf–toɣo hunv–nd–ra* [2SG–father old.age–TRM live–FIN–PRD] 'your father lived up to an old age' (Grz: 21). The Ablative case (called 'Locative/Ablative' by Gruzdeva: 18, 20) can mark not only the source of an action or movement as well as the source of information, but also a stative location, e.g. *t'ivlaŋ čaɣ–ux ŋat'x vezla–j* [cold water–ABL foot cramp–FIN] '(I) have a cramp in (my) foot in the cold water', *umgu ńo–x p^hu–j* [woman barn–ABL come.out–FIN] 'A woman came out from the barn' (Grz: 20). It can also express a starting point in time as well as movement through space (PanI: 136). Furthermore, the Amur dialect of Nivkh has a separate Locative case expressing stative location in space and time as well as 'have' possession, e.g. *ńïŋ hï dïv–uin q^ho–j–ra* [1PL this house–LOC sleep–FIN–PRD] 'we slept in this house' (Grz: 19). The Locative case covers most instances of stative location in this dialect, with the Ablative expressing movement in a location; only rarely does the Ablative have a parallel function to the Locative, and then generally in conjunction with postpositions (PanI: 134f).

In Nganasan, the Dative (called 'Dative-Directive' by Tereščenko: 82) expresses direction towards a location, a spatial or temporal limit, as well as marking

recipients and addressees. There is a beginning tendency in Nganasan to distinguish between an illative and an allative meaning with the use of different suffixes, e.g. *desïmə ŋenduy–t'e t'iⱬeje* 'my father sat (down) in the boat', where *ŋenduyt'e* takes the suffix with an illative meaning vs. *tïmïńa desïmə ŋentu–d'a konïⱬïtə* 'now my father is going to the boat', where *ŋentud'a* carries the suffix with an allative meaning. Similarly, verbs of speech and emotions predominantly govern indirect objects carrying the 'illative' suffix, while verbs like 'bring', 'give', 'show oneself to' take indirect objects marked with the 'allative' suffix (Ter: 83). However, this tendency has not yet been grammaticalized to an obligatory distinction (Ter: 84). The Nganasan Locative case (called 'Locative-Instrumental' by Tereščenko: 87) has two main functions, to express a stative location and period of time as well as an instrument or (less frequently) accompaniment. According to Tereščenko, the locative uses are the primary ones, with the instrumental/comitative meanings being a later development (Ter: 87).

Similarly in Selkup, the Dative case marks recipients, benefactees, and addressees of verbs of speech as well as the direction of movement (OSJ: 178). There are two suffixal variants, one of which according to MSJ (197, 199) attaches to inanimate nouns, the other to animate nouns. However, this point of view is refuted by the OSJ (179), who show that both suffix variants attach to inanimate and animate nouns. There is an additional Illative case which expresses the endpoint of movement when attached to inanimate nouns; it can also have a broader function of expressing recipients and addressees and is then not distinct in function from the Dative case (OSJ: 179f). The Locative case expresses the location in time or space of an action (OSJ: 180f). An additional 'Personal-Locative' case in the Southern Selkup dialects attaches only to animate nouns and expresses 'have' possession[6] (MSJ: 224ff).

In the Sos'va dialect of Mansi, the Lative case combines allative, illative and dative functions, expressing the direction of movement as well as its endpoint, e.g. *kol–n* [house–LAT] 'into/onto/towards the house', as well as marking recipients, benefactees, addressees of verbs of speech, e.g. *ma:n ńa:wram–n a:kań miɣuw* [1PL child–LAT doll give.1PL] 'we give the child a doll' (Rse: 25, Rmb: 45ff). With certain nouns it can express the purpose of movement, e.g. *wit–n yal–s–əm* [water–LAT go–PST–SBJ.1SG] 'I went to get water' (Rmb: 46). There are differences between subdialects in the use of the Lative case, with some subdialects predominantly using directional postpositions instead of the Lative case (Rmb: 48). The Locative case

---

[6] Note the parallel category of a 'Personal Locative/Adessive' case in the neighbouring language Ket, as described above.

marks stative location in space (expressing both adessive and inessive meanings) and time (Rse: 25; Rmb: 49).

The northern Khanty dialect of Obdorsk differs very much from the other Uralic (and Siberian) languages in having only three cases: a Nominative, a Locative, and a Translative (Nik: 13). A number of locative and other meanings are expressed by postpositions which agree with pronominal dependents in person and number. Postpositions going back to noun stems take Locative case marking, while postpositions going back to adverbial stems are unmarked for case, e.g *u:rŋ–e:m–na* [for–1SG–LOC] 'for me', *pul'an–e:m* [in.front–1SG] 'in front of me', *muŋ xo:t–e:w xośa* [1PL house–1PL to] 'to our house' (Nik: 37). Amongst other meanings, the Nominative case marks recipients and benefactees, with the direct object taking Locative case suffixes and the recipient or benefactee triggering object agreement on the verb (27a). Alternatively, the recipient can be marked by a lative postposition; in this case, the direct object can, but need not, trigger object agreement on the verb, (27b):

(27a)

(ma)	*juwan*	*a:n–**na***	*ma–s–e:m*
1SG	John	cup–LOC	give–PST–SG.OBJ.1SG

'I gave John a cup.'

(27b)

(ma)	*juwan*	*e:lti*	*a:n*	*ma–s–ə–m*
1SG	John	to	cup	give–PST–E–SBJ.1SG

'I gave John a cup.'

[Nik: 40]

However, the case system in the Northern Khanty dialects differs for the personal pronouns, which lack the Lative and Locative case, but distinguish an Accusative and Dative case; the Dative expresses mainly addressees and benefactees (Koškareva 2001: 236, 250). The Locative case in Obdorsk Khanty expresses both stative location in space and time as well as the direction of movement, e.g. *we:t–na* [five–LOC] 'at five o'clock' (Nik: 13), *luw xo:t–ə–l–na o:məs–l* [3SG house–E–3SG–LOC sit–PRS.SBJ.3SG] 'he is sitting in his house' (Nik: 14), *pa wo:s–na man–ə–s* [another city–LOC go–E–PST.SBJ.3SG] 'he went to another city' (Nik: 38). In the Eastern Vakh dialect of Khanty, which has a large number of cases (ten according to Abondolo 1998a: 361, eight according to Tereškin: 42), the Allative case expresses the direction of movement and marks addressees of verbs of speech (Trš: 44f), while the 'Allative-Purposive' case expresses the goal of movement, the location of an event or action, the purpose of an action (e.g. to get water) as well as addressees of

verbs of speech and benefactees (Trš: 46ff). The difference in meaning between the Allative and the Allative-Purposive case as markers of addressees does not become clear from Tereškin's description. The Locative case expresses spatial and temporal location (Trš: 48f). As in the Northern Khanty dialects, in the Eastern dialects (e.g. Surgut) the personal pronouns have an additional distinction between an Accusative and a Dative case (Koškareva 2001: 236, 250).

Table 3.6: The functions of the dative, allative, and locative cases in Siberian languages

	Dative	**Allative**	**Locative**
Mansi	Direction, recipients, addressee, purpose		Stative location
Selkup	Recipient, addressee, direction		Stative location
Nganasan	Direction, recipient, addressee		Stative location, instrument
Ket	Direction, recipient, addressee		Stative location, (direction)
Siberian Inupik	Direction, purpose, addressee		Stative location
Itelmen	Direction, purpose, recipient, addressee		Stative location
Yukaghir	Recipient, addressee, purpose, direction		Stative location, direction
Chukchi		Direction, recipient, addressee	General spatial location
Obdorsk Khanty*			Stative location, direction
Vakh Khanty		Direction, addressee	Stative location
Nivkh**		Direction, addressee	Stative location, motion from; AD Locative: stative location

*In Obdorsk Khanty, the case system is rather different than that of other Siberian languages, because there are only three cases and thus there is widespread case syncretism. Recipients and benefactees are expressed by the Nominative case.
**In Nivkh, case functions differ from those in other Siberian languages, making a proper representation in the table impossible. Indirect objects can be expressed by the Nominative case. AD = Amur Dialect.

From the descriptions of the case systems outlined in the preceding sections (cf. 3.2.2.2, 3.2.2.3, 3.2.2.4, 3.2.2.5) it is clear that the merging of dative and locative functions in one case found in Mongolian, Sakha/Dolgan, and Evenki is not a Siberian areal feature, but is quite distinctive.

3.2.2.6 The functional extension of the Sakha Dative case in the light of Eurasian case-marking patterns

As can be seen from Table 3.1, the Turkic Dative and Locative suffixes are clearly cognate in all the languages that have retained them, with the exception of the Khalaj Locative. The extension of the Dative to include allative functions which is found in some languages seems to have happened very early, since it is present in almost all branches of Turkic. Furthermore, there are already some instances of the Dative taking on allative functions in Old Turkic, and the allative is lost in Uyghur later than the 11th century (Erd: 370f).

It has been suggested that the retention of the Allative case in the South Siberian Turkic languages Tuvan and Khakas may have been due to areal influence: "The formal opposition of a DAT[ive] and an ALL[ative] case is found in both Tyvan and Xakas, and is an areal feature of Siberian native languages, lacking only in the languages of the extreme western and northeastern peripheries (e.g. W. Siberian Turkic, NW Samoyed, W Ob-Ugrian dialects, Yakut, Chukchi, and Siberian Eskimo)." (And: 13). However, as can be seen from the above discussion, it is only the Tungusic languages that distinguish between a Dative case to express indirect objects and an Allative case to express direction. According to Dolgix (1960: map; cf. Figure 1.2), Tungusic tribes were settled somewhat to the southwest of Lake Baykal in the 17th century and may therefore have been in contact with South Siberian Turkic groups. This is further corroborated by genetic data, which shows that South Siberian Turkic groups share part of their mtDNA genepool with Tungusic-speaking peoples; these mtDNA lineages are shared by the Sakha as well, indicating that the period of close contact between South Siberian Turkic speakers and Tungusic speakers may go back a fairly long time (Pakendorf et al. 2006). However, the main substratal influence expected in the Tuvans and Khakas is from Samoyedic speakers (Dolgix 1960, map; cf. Figure 1.2). Unfortunately, the only grammar sketches available to me of the single Sayan Samoyedic language that was described in some detail, Kamass, do not state how indirect objects are marked in this language; however, they do not list a dative case in their case paradigms (Künnap 1999: 14; Simoncsics 1998: 585). Kamass has a Lative case expressing illative and allative meanings and a Locative case with inessive and adessive

meanings (Künnap 1999: 16f). Judging from the text published in Künnap's sketch (39ff), addressees of verbs of speech are marked by the Lative or Locative case [the two cases take the same suffix in the possessive declension (Künnap 1999: 14f) and it is therefore not possible to distinguish between them in the two examples found], e.g. *di nüke ńi:–**gəndə** mo–l'e* [this old.woman son–LAT/LOC.POSS.3SG say–PRS.3SG] 'this old woman says to her son' (Künnap 1999, text line 23), *də ńi šo–bi:ze iya:– **ndə** ne:rbə–lie* [this son come–CVB mother–LAT/LOC.POSS.3SG retell–PRS.3SG] 'the son, after he returned, retells his mother' (Künnap 1999: text line 34). Although it is impossible to judge from the lack of description, there thus does not seem to be much evidence for a distinction between a dative and an allative case in Kamass. It is therefore rather difficult to understand where the supposed areal influence that led to the retention of the Dative-Allative distinction in Tuvan and Khakas would have come from, especially since Tofa, spoken in the vicinity of Tuvan and Khakas, has lost the distinction.

The use of only one case to mark indirect objects and goals of motion in the Siberian languages is not a striking feature. The lack of a separate allative case is cross-linguistically quite common, the goal of movement being frequently expressed by the accusative, the dative, or the locative case (Blake 2001: 160). Furthermore, the use of only one case in Sakha to mark both the goal of motion and stative location is also not particularly rare: according to DeLancey (2003: 61), probably more languages do not make a distinction on the noun phrase between goals and locations than languages that do make such a distinction (see also Newman 1998: 8-17). Thus, one might wonder whether the extension of the Sakha Dative case to express the semantic roles of goal and location may not have been due to an independent internal innovation rather than being attributable to contact influence. However, there are two arguments against such an assumption. One is the fact that a cross-linguistic hierarchy of cases can be established, according to which languages that have an ablative and/or instrumental case tend also to have a locative case (Blake 2001: 156). In particular, a separate locative case is very common in languages with fairly large case systems: "Systems of six or more cases almost always have a locative." (Blake 2001: 157). Thus, the fact that Sakha, which has an Ablative, Instrumental and Comitative (not to mention a Comparative) case – all of which are lower on the case hierarchy than the locative – lacks a separate locative case, is cross-linguistically quite rare. A much weightier argument, however, is provided by the survey of the Siberian languages, which clearly shows that although the use of one case to mark both recipients/addressees as well as the goal of direction is widespread in Eurasia, most languages of this region have a separate locative case to express stative location. Thus, the extension of the Dative case in

Sakha to express locative and allative functions as well as indirect objects distinguishes it from both the western Turkic languages such as Turkish or Kazakh and Kirghiz, where the Dative took over allative functions, but which retained the Locative case, and also from its nearest linguistic neighbours, the South Siberian Turkic languages, which retained the distinction in all three cases. It also distinguishes it from most of the other Siberian languages, in which commonly the dative and allative are combined, but which have a separate locative case. The fact that Sakha shares the rare feature of having lost its Locative case with Mongolic and Evenki points towards contact influence rather than independent innovation.

The fact that the Sakha Dative covers the functions of the dative, locative and allative cases points towards Mongolic, rather than Evenki, as the source of influence[7]. Although the Dative case carries locative functions in both Mongolic and Evenki, in Mongolic, as in Sakha, it additionally has allative functions, while Evenki has two separate allative cases for the expression of the goal of motion. Had Evenki been the source of influence, one would expect Sakha to have retained its Allative case. That such a retention is possible is shown by Tuvan and Khakas. This analysis is further strengthened by two observations: firstly, in Tuvan, which has been in long-standing contact with Mongolian[8], we find a partial conflation of the Dative and Locative, such as we find in Sakha. However, in Tuvan this is restricted to the past and future tenses, with the Locative case still being used in the present tense (A/H: 18). Furthermore, with the exception of Evenki, the Tungusic languages distinguish between a Dative and a Locative case in addition to marking direction with an Allative case. The borders between the case functions are not clear-cut, with the Dative case expressing both dative and locative meanings, and the Locative case expressing both locative and allative meanings; nevertheless, the dative and locative functions haven't collapsed completely in Èven and the Amur Tungusic languages, as they have in Evenki. In Evenki we clearly have an extension of a minor use pattern (using the Dative case to express locative meanings) to a major use pattern, where the Dative is the only case possible to express location. Such a shift from a minor use pattern to a major use pattern can well be an internal development; however, it is plausible that contact influence from Mongolic and/or Sakha strengthened this.

---

[7] This was also suggested by Gregory Anderson (pers. comm.).
[8] For instance, a number of Tuvans from the southeast of the Republic are bilingual in Mongolian and Tuvan (Leighton & Bicheldei 1994).

3.2.3 The development of the Partitive case in Sakha

Yet another case in which Sakha differs from the remaining Turkic languages is the presence of the Partitive case [initially called Indefinite Accusative by Böhtlingk ([1851] 1964)]. The Partitive in Sakha developed out of the Common Turkic Locative case in *–DA* (Stachowski & Menz 1998, Tenišev 2002; compare the case suffixes for Turkic languages presented in Table 3.2 in section 3.2). This had both locative and ablative meanings in earlier Old Turkic (Erd: 173, 174f), and it is suggested that the Partitive developed from such an ablative meaning of the Locative, e.g. 'from the water' > 'some water' (Poppe 1959: 681). However, it is believed that the shift in meaning in Sakha was enabled by Evenki influence (Poppe 1959: 681; Ubr: 118; Artem'ev 1999b: 105, 106).

3.2.3.1 Sakha case-marking on direct objects

The differences in case-marking on direct objects (or related differences in subject vs. object agreement on the verb) that are made in a number of Eurasian languages (most often a distinction between an unmarked (Nominative) and a marked Accusative case) have variously been explained as a distinction between definiteness and indefiniteness (e.g. Imr: 876, 882; Ras: 36, 37), specificity and non-specificity (e.g. Skr: 109), 'relevance of referent identification' (Comrie 1978: 12), or topicality and rhematicity of the direct object (Nik: 73, 76; Skribnik 2001: 350ff). It should be further noted here that often it is not clear what exactly the authors of language descriptions mean by '(non-)specificity' or '(in)definiteness', the principles of classification being kept rather vague. An (admittedly cursory) examination of the Sakha data appears to me to be most compatible with a distinction between specific (Accusative-marked) and non-specific (Nominative-marked) direct objects, although there are a number of instances of use where this categorization cannot explain the particular use of the Accusative or Nominative case to mark a direct object. However, 'relevance of referent identification' or topicality/rhematicity of the object do not appear to provide an explanation, either. The precise functions of the differential direct object marking in Sakha clearly require a more detailed study; that, however, will need to be done at a different time and place than in the current thesis.

In Sakha direct objects in the Indicative and Conditional mood are marked with either the Accusative case or the (suffixless) Nominative case. Based on a brief and preliminary examination of the data, it appears that the Accusative is used to mark definite direct objects (28a), objects modified by a quantifier or a possessor

(28b, c), as well as specific indefinite direct objects[1] (28d), while the Nominative case marks generic indefinite direct objects (28e). However, this rule is not hard and fast, with the case-marking being different than expected in a number of instances.

(28a)

gosudarstvennay	*interiehi*	*körböt*	*buolbuttar*
gosudarstvennay	interies–(n)I	kör–BAt	buol–BIt–LAr
state	interest–ACC	see–PRSPT.NEG	AUX–PSTPT–PL

'They don't look to the state's interests anymore.'

[MatX2_110]

(28b)

bi:r	*ku:l*	*haχarï*	*ïllaχχïna*	*baltara*
bi:r	ku:l	saχar–(n)I	ïl–TAχ–χInA	baltara
one	sack	sugar–ACC	take–MDL–COND.2SG	one.and.half

tïhï:nčaŋ	*barar*
tïhï:nča–(I)ŋ	bar–Ar
thousand–POSS.2SG	go–PRSPT

'If you take (i.e. buy) one sack of sugar, your one and a half thousand go (i.e. you have to pay 1,500 rubles).'

[LukP135]

(28c)

iti	*hoγotuopkaγa*	*tutar*	*buolbuttara*	*kuobaχ*
iti	soγotuopka–GA	tut–Ar	buol–BIt–LArA	kuobaχ
this	stockpile–DAT	hold–PRSPT	AUX–PSTPT–POSS.3PL	rabbit

etin	*tiri:tin*
et–(t)In	tiri:–(t)In
meat–ACC.3SG	leather–ACC.3SG

'They have started to take rabbit meat and leather for stockpiling.'

[Efmy265]

(28d)

aγa	*doydunu*	*kömüskü:r*	*seri:*	*sïllarïgar*	*kilbienne:χ*
aγa	doydu–(n)I	kömüske:–Ar	seri:	sïl–LArIgAr	kilbien–LA:χ
father	country–ACC	defend–PRSPT	war	year–DAT.3PL	glow–PROP

ületin	*ihin*	*dien*	*mede:li*	*ïlbïtïm*
üle–(t)In	ihin	die–An	mede:l–(n)I	ïl–BIt–(I)m
work–GEN	for	say–PF.CVB	medal–ACC	take–PSTPT–POSS.1SG

'I received a medal for my outstanding work during the war for the defense of the fatherland.'   [RaxA191]

---

[1] In my classification of direct objects I mostly follow Givón (1978: 296), choosing the term 'specific' over 'referential' following Lyons (1999: 57f).

(28e)

onu	*me:kkele:n*	*leppieske*	*oŋoron*	*hi:r*
on–(n)I	me:kkele:–An	leppieske	oŋor–An	sie–Ar
that.OBL–ACC	grind–PF.CVB	flatbread	make–PF.CVB	eat–PRSPT

etibit
e–TI–BIt
AUX–PST–1PL

'We ground that and made flatbread and ate that.'

[BesP65]

In the affirmative Imperative mood, however, partially affected mass nouns in direct object position take the Partitive case (29a, b), so called precisely because of this function. Furthermore, the Partitive is also used to mark generic indefinite and non-referential direct objects (30a-d), even when they are modified by a quantifying noun (30c) or an adjective (30d), while definite and specific indefinite direct objects take Accusative case marking (31a, b). Because of this function, Böhtlingk (Btl: 256, 320) initially named this the 'Indefinite Accusative' case. In the negative Imperative mood, direct objects generally take Accusative case marking[2] (32a, b).

(29a)

emie	*eji:y*	*K.*	*χaččïta*	*ï:ttïn*	*dien*
emie	eji:y	K.	χarčï–TA	ï:t–TIn	die–An
also	older.sister	K.	money–PART	send–IMP.3SG	say–PF.CVB

Joku:skayttan	*huruyar*
Joku:skay–(t)tAn	suruy–Ar
Y.–ABL	write–PRSPT

'She also wrote from Yakutsk "Let Aunt K. send (some) money".'

[Efmy473]

(29b)

halamatta	*huoratta*	*amsayïŋ*
salamat–TA	suorat–TA	amsay–(I)ŋ
salamat–PART	yoghurt–PART	taste[PRXIMP]–2PL

'Try some salamat (Yakut dish), some yoghurt.'

[Efmy543]

---

[2] This is reminiscent of Russian, where the Accusative case is preferred over the Genitive in negated imperatives as well (Timberlake 1975: 129). Since I elicited negative imperative sentences not only by asking for translations from Russian (in the field), but also from German and English (from Katja Potapova in Bonn and Njurgujana Petrova in Buffalo by e-mail), a direct influence of Russian as contact language for the elicitation can be excluded.

(30a)

kihi	*ere*	*buollargïn*	*ütüöte*	*oŋor*
kihi	ere	buol–TAr–GIn	ütüö–TA	oŋor
person	PTL	AUX–COND–PRED.2SG	good–PART	make[PRXIMP.2SG]

'… if you are a man do good…'

[Afny154]

(30b)

mieχe	*atta*	*tutan*	*bier*
mieχe	at–TA	tut–An	bier
1SG.DAT	horse–PART	hold–PF.CVB	BEN[PRXIMP.2SG]

'Catch me a horse.'

[translation, Oln, Sun, Tat, Ver]

(30c)

ikki buočuka	*u:ta*	*belemne:riŋ*
ikki buočuka	u:–TA	belem–LA:–A:r–(I)ŋ
two barrel	water–PART	ready–VR–DSTIMP–2PL

'Prepare two barrels of water!'

[GSJa: 134]

(30d)

Mašaγa	*üčügey*	*kinigete*	*ïl*
Maša–GA	üčügey	kinige–TA	ïl
M.–DAT	good	book–PART	take[PRXIMP.2SG]

'Buy a good book for Masha.'

[translation, Ver]

(31a)

bïra:kkïn	*ïskuolaγa*	*aγal*	*ere,*
bïra:t–GIn	oskuola–GA	aγal	ere
younger.brother–ACC.2SG	school–DAT	bring[PRXIMP.2SG]	PTL

di:ller
die–Ar–LAr
say–PRSPT–PL

' "Bring your nephew to school", they said.'

[PotP9]

(31b)

uoppuskaγa	*bara:ččïnï*	*eŋin*	*holbuy*	*diebitim*
uoppuska–GA	bar–A:ččI–(n)I	eŋin	solbuy	die–BIt–(I)m
leave–DAT	go–HAB–ACC	etc.	replace[PRXIMP.2SG]	say–PSTPT–POSS.1SG

' "Replace someone who left on holiday", I said.'

[IvaP374]

(32a)

*oɣolor**u***	*χa:yan*	*hïljïma*	*di:r*
oɣo–LAr–(n)I	χa:y–An	sïrït–(I)mA	die–Ar
child–PL–ACC	lock.up–PF.CVB	go[PRXIMP.2SG]–NEG	say–PRSPT

' "Don't leave (your) children alone!", he said.'

[RaxA271]

(32b)

parkaɣa	*sibekki**ni***	*ü:rge:me*
parka–GA	sibekki–(n)I	ü:rge:–(I)mA
park–DAT	flower(s)–ACC	tear.out[PRXIMP.2SG]–NEG

'Don't pick flowers in the park!'

[translation, Ver]

The use of the Nominative for non-specific direct objects with affirmative Imperative verbs is acceptable (e.g. 29b, 30a and 30b could all be used with the object in the unmarked Nominative case rather than the Partitive case); however, Partitive case-marked objects are more common and are given spontaneously. The Partitive is occasionally also acceptable with verbs in the Indicative mood, e.g. *burduk–**ta** tard–ï–bap–pït daɣanï* [grain–PART grind–E–NEG.PRSPT–1PL PTL] 'we didn't even grind grain' (back translation, Ver).

3.2.3.2 Case-marking on direct objects in other Turkic languages

In Old Turkic, unmarked nouns can serve as direct objects even when these are specific or definite, e.g. *kara kum aš–miš* [kara kum cross–PSTPT] 'they had crossed the Kara Kum (desert) (Erd: 362), *bo biti–dökte* [this write–TEMP] 'when I wrote this' (Erd: 363). The Accusative case appears to mark only specific direct objects in Orkhon Turkic. An explanation for the difference in use between the unmarked Nominative and the Accusative case is hard to find (Erd: 366). Interestingly, in an Uygur example, the Accusative case is used with a NP modified by a numeral: *altï kïz–lar–**ig** bulun al–ïp* [six girl–PL–ACC prisoner take–CVB] 'taking six girls as prisoners'. As Erdal states, "[t]he girls were obviously not mentioned earlier in the story and should not have the accusative suffix if that were a mark of definiteness." (Erd: 366). This parallels the use of the Accusative case with NPs that are modified by a quantifier in Sakha, cf. (28b).

In Turkish, indefinite direct objects take the indefinite determiner *bir* 'a'. In general, they remain unmarked, i.e. in the Nominative case (G/K: 174). Definite direct objects obligatorily take Accusative marking, which is also required in certain syntactic positions (G/K: 175, 176). In certain contexts the use of Accusative case

marking can disambiguate between a specific and a non-specific use of the indefinite direct object, e.g. *Gürcistan folklor–uyla ilgili bir kitap ar–ıyor–um* [Georgia folklore–?? concerning one book look.for–AOR–1SG] 'I'm looking for a book on Georgian folklore' (G/K: 375). Here, the unmarked NP indicates that any book on Georgian folklore would do, while marking it with the Accusative case (*bir kitab–ı*) would indicate that the speaker is looking for a specific book on Georgian folklore. In addition, the Accusative case can be used to mark an indefinite noun phrase in Turkish if the referent will feature in further discourse, while the case-marking can be omitted when the identification of the referent is irrelevant in the current context (Comrie 1978: 12). Since Turkish does not have a separate partitive or indefinite accusative case, indefinite noun phrases or partially affected mass nouns remain unmarked, even in the Imperative mood, while definite noun phrases take Accusative case marking, e.g. *patates al* [potato take[IMP.2SG]] 'buy potatoes!', *yemek–ten önce şekerleme ye–me* [dinner–ABL before crystallized.fruit eat–NEG.IMP.2SG]] 'don't eat sweets before dinner!', *bana tuz–u uzat lütfen* [1SG.DAT salt–ACC hand[IMP.2SG] please] 'pass me the salt, please!' (Mehmet Somel, pers. comm.).

In Kirghiz, the unmarked Nominative case marks indefinite direct objects (Imr: 876), while definite direct objects are obligatorily marked with the Accusative case (Imr: 882). In Kazakh, too, direct objects can be unmarked (called 'unmarked Accusative' in SKJ: 160, 168) when they are indefinite or non-specific ('when expressing abstractness or indefiniteness, when there is no obvious need to be more concrete' SKJ: 168), while the Accusative case marks definite direct objects (Somfai Kara 2002: 19).

In Uzbek, definite direct objects take the Accusative case (Bdr: 71ff), while indefinite direct objects ("nouns without lexical, contextual, or morphological definition") remain unmarked (Bdr: 74). Note that Bodrogligeti calls such unmarked direct objects 'Nonspecific Accusative case forms' rather than Nominative case forms, similar to the terminology of the *Sovremennyj Kazaxskij Jazyk* (SKJ: 160, 168).

In Tuvan, indefinite direct objects remain in the unmarked (Nominative) case, while definite or specific direct objects take Accusative case marking (A/H: 15, 17). However, judging from the following example [given by Anderson & Harrison (A/H: 15) to demonstrate the use of the Nominative case with indefinite direct objects], specificity does not necessarily entail Accusative case marking, since here an indefinite specific noun is unmarked: *men oon čagaa al–d–ïm* [1SG 3SG.ABL letter take–PST–1SG] 'I got a letter from him' (A/H: 15). In addition, the Accusative case marks direct objects that are at a distance from the verb (Isxakov & Pal'mbax

1961: 132). In Khakas, too, the Accusative case marks definite direct objects, while indefinite, non-specific, or unquantified noun phrases stand in the Nominative case (And: 7, 10).

As in the other Turkic languages, in Tofa indefinite direct objects appear in the unmarked Nominative case, e.g. *men balïk tut–adïr–men* [1SG fish catch–PRS–1SG] 'I catch fish', *ool–gïs akkaš čul–gan–nar* [boy–girl flower tear–PSTPT–PL] 'the children picked flowers' (Ras: 36). However, contrary to what we find in Turkic languages in general, in Tofa definite direct objects do not always take Accusative case marking. When the direct object is a possessed noun phrase with a first or second person possessor, the possessive marking is sufficient to indicate definiteness, e.g. *men at–ïm baɣla–d–ïm* [1SG horse–POSS.1SG tie.up–PST–1SG] 'I tied up my horse', cf. *at–ïn baɣla–d–ï* [horse–ACC.3SG tie.up–PST–3SG] 'he tied up his horse' (Ras: 36), although Accusative case-marking can also co-occur with the first or second person possessive suffixes (Ras: 37). The Accusative case is used when the direct object is definite, e.g. *bo at–tï mun–ub al* [this horse–ACC ride–CVB take] 'ride this horse' (Ras: 37). Furthermore, like Sakha Tofa has a Partitive case that is used to mark partially affected mass nouns in direct object position in the Imperative mood, e.g. *sug–da hal* [water–PART bring[IMP.2SG]] 'bring (some) water!', *šey–da iši–vit* [tea–PART drink–RES[IMP.2SG]] 'drink some tea!', cf. *šey–ni iši–vit* [tea–ACC drink–RES[IMP.2SG]] 'drink (all) the tea!' (Ras: 40).

3.2.3.3 The origin of the Sakha and Tofa Partitive case

As can be seen from the examples, the Partitive case endings in Sakha and Tofa are cognate, with the archiphonemic form *–TA/–DA*. This is cognate to the Common Turkic Locative case (cf. Table 3.2 in section 3.2). Tofa still has the Common Turkic Locative case, marked by the suffix *–DA*; this expresses location in place or time, e.g. *oŋ taš–ta oluru* [3SG stone–LOC sit.PRS] 'he is sitting on a stone'; *ol hire–de ög–de kum ta yok bol–gan* [that time–LOC house–LOC who PTL non-existence AUX–PSTPT] 'at that time nobody was home' (Ras: 39). In Sakha, however, the Common Turkic Locative case has been lost, with the Dative taking over locative functions (cf. section 3.2.2.1).

In Old Turkic, the Locative case had both locative and ablative functions, e.g *ol ev–de* [that house–LOC] 'in that house' (Erd: 371); *Tabgač xagan–ta bediz–či kel–ür–tü–m* [China khagan–LOC ornament–NR come–CAUS–PST–1SG] 'I brought decorators from the Chinese emperor'; *ay teŋri ordo–sïn–ta ėn–ipen* [moon god palace–POSS.3SG–LOC descend–CVB] 'coming down from the palace of the Moon God' (Erd: 372). It is assumed that the Tofa and Sakha Partitive case (with a

partitive meaning) developed from the ablatival function of the Old Turkic Locative (Poppe 1959: 681, Ras: 41). The development of a partitive function out of ablatives is a cross-linguistically common grammaticalization path, found for example in French, Bulgarian, Lezgian, Finnish, and Basque (Heine & Kuteva 2002: 32f). Thus, part of the functions of the Sakha Partitive case can be explained through language-internal development. However, there is a difference between the Tofa and the Sakha Partitive cases in that the former can express only partitive meanings, that is, it marks only partially affected mass nouns in direct object positions. The function of marking indefinite direct objects, which the Partitive additionally has in Sakha, is not fulfilled by the Tofa Partitive (Gregory Anderson, pers. comm.). Thus, the 'indefinite accusative' function of the Sakha Partitive distinguishes this case from its Tofa counterpart, and it is clearly not a common feature of Turkic languages. Contact influence may well have played a role in its development, as will be examined below.

3.2.3.4 Case-marking of direct objects in Mongolic and in Evenki

In Written Mongolian and Khalkha Mongolian, indefinite direct objects take zero case-marking (i.e. they are identical to the Nominative case – 33a), while definite direct objects take Accusative case-marking (33b). From the examples given by Kullmann & Tserenpil (K/Ts: 88) it is clear that the unmarked direct objects can be both specific and generic indefinite as well as non-referential, cf. 33a and also: *bi mor' unax dur–tay* [1SG horse[NOM] ride desire–PROP] 'I love to ride horses' (K/Ts: 89). Partially affected mass nouns in direct object position take Ablative case-marking (33c) (PopWM: 147ff, 151; K/Ts: 87f, 91). In Dagur, too, the distinction between the unmarked (Nominative) direct object and a direct object in the so-called Connective case (which has both genitive and accusative functions) is analyzed as being due to the difference between indefinite and definite direct objects (Tsu: 138).

(33a)
   bi   nom   av–la:
   1SG   book   take–PST

   'I bought a book.'

(33b)
   *bi   ene   nom–**ig**   av–la:*
   1SG   this   book–ACC   take–PST

   'I took this book.'

                                                                    [K/Ts: 87, 88]

(33c)

> *ene cayn–**a:s** u:–ž*   *xo:ln–**o:s** id–e:rey*
> this tea–ABL   drink–IPF.CVB   food–ABL eat–PRESCR
>
> 'Please drink some of this tea and eat some of this food.'

<div align="right">[K/Ts: 91]</div>

In Buryat, unspecific or indefinite direct objects stand in the suffixless oblique stem, while the Accusative case marks specific or definite direct objects (Skr: 109; PopB: 114f). Partially affected mass nouns remain unmarked, while objects that are completely affected by the action designated by the verb take Accusative case-marking; compare: *uha asar–ii–t* [water[NOM] bring–PREC–2PL] 'bring water!' vs. *bi uh–**iye** uu–gaa–b* [1SG water–ACC drink–IPFV–1SG] 'I have drunk the water (i.e. all the water mentioned before)' (PopB: 115). As pointed out by Skribnik (2001: 353), the Accusative case is used not so much as a marker of definiteness, but rather to mark the direct object as topical, since a topical, but indefinite, direct object takes Accusative marking: *ene tere lama böö–ner–**iye** šüte–že huu–dag bai–gaa* [this that lama shaman–PL–ACC believe–IPF.CVB AUX–HAB be–PRF.3] '(People) used to believe in lamas, shamans and such.'

In Evenki, direct objects can be marked in three ways: with the Definite Accusative case (suffix *–vA/–mA*), with the Indefinite Accusative case (suffix *–(y)A*), or with the unmarked Nominative case plus reflexive-possessive suffixes (Ned: 147). The Definite Accusative case is the predominant way to mark direct objects (34a), while the Indefinite Accusative is used only to mark clearly indefinite direct objects (34b, c), objects that have not been made yet, and partially affected mass nouns (34d) (Ned: 147, 192f, 194; Kon: 49). The unmarked Nominative case plus reflexive-possessive suffixes is used to mark definite (possessed) direct objects when the possessor is coreferential with the subject (34e); in this case, the Definite Accusative case marker cannot be used (Ned: 144, 192). Although the use of the Definite Accusative case is obligatory when the direct object "has clear referential status either for the speaker or for any participant of the situation" (34a; Ned: 192), it can also be used to mark indefinite direct objects (34f); accordingly, example (34a) can also have the meaning of 'catch an/any reindeer' (Ned: 193). The Indefinite Accusative can only be used to mark indefinite direct objects or parts of whole; however, as mentioned, this is optional. There are further restrictions on the use of the Indefinite Accusative case, which is used only with the Future Indicative and Imperative mood (34b, d), as well as with habitual verbs (34c) (Ned: 194). The Definite Accusative, on the other hand, is used with all the past tenses (Ned: 194). It thus appears that the Definite Accusative case is the default case for marking direct objects, while Indefinite Accusative case forms are used only when the speaker

wants to emphasize the non-referential or partitive nature of the direct object, although this needs further investigation (Ned: 192).

(34a)
*oron–**mo***   *java–kal*
reindeer–DEF.ACC take–PRXIMP.2SG

'Catch that (definite) reindeer.'

(34b)
*oron–**o***   *java–kal*
reindeer–INDF.ACC   take–PRXIMP.2SG

'Catch yourself a/any reindeer.'

(34c)
beye   *mo:ka–r–**e***   *genne:–vki*
man   stick–PL–INDF.ACC   bring–HAB

'The man usually brings firewood.'

                                                  [Ned: 193]

(34d)
min–du   *ulle–**ye***   *kolobo–**yo***   *bu:–kel*
1SG–DAT   meat–INDF.ACC   bread–INDF.ACC give–PRXIMP.2SG

'Give me (some) meat and (some) bread.'

                                                  [Ned: 194]

(34e)
bi   *oro–r–**vi***   *eteyet–če–m*
1SG   reindeer–PL–PREFL   guard–PST–1SG

'I guarded my reindeer[3].'

                                                  [Ned: 144]

(34f)
tar   *asi*   *kniga–**va***   *taŋ–jara–n*
that   woman   book–DEF.ACC   read–PRS–3SG

'That woman is reading a/the book.'

                                                  [Ned: 193]

---

[3] In Nedjalkov's description (Ned: 144, ex. 555a) the suffix *–čA* is glossed as present tense, and the translation is 'I guard my reindeer'. This contradicts the information on tense (Ned: 235, 238ff), in which the suffix *–čA* is described only as a past tense marker. I have here opted for the presumably correct gloss and translation, which therefore differs from the original.

3.2.3.5 Case-marking of direct objects in other Siberian languages

In Kolyma Yukaghir, there is a distinction in the case-marking of direct objects depending on the person of the agent. If the agent is first or second person, then a third person direct object remains unmarked (35a), while first or second person direct objects take Accusative case marking (with a special pronominal Accusative suffix different from the form that attaches to nouns – Mas: 95) (35b). If the agent is third person, definite direct objects, possessive noun phrases and proper nouns take Accusative case marking (35c), while indefinite direct objects take Instrumental case marking (35d) (Mas: 89, 95). Partially affected direct objects can take Ablative case marking (36a) (Mas: 112f), although this is relatively rare (Elena Maslova, pers. comm.), or they remain in the Nominative (unmarked) case (36b).

(35a)

met	*me:me:*	*iŋi:*
1SG	bear	be.afraid[TR.1SG]

‘I am afraid of the bear.’ [Mas: 89]

(35b)

met	*tet–**ul***	*kudede–t*
1SG	2SG–ACC	kill–FUT[TR.1SG]

‘I will kill you.’

[Mas: 95]

(35c)

titte	*ču:l–**gele***	*min–ŋa:*
3PL.POSS	meat–ACC	take–TR.3PL

‘They took their meat.’

[Mas: 93]

(35d)

*ńumuji:–**le***	*mij–u–m*
axe–INS	take–E–TR.3SG

‘He took an axe.’

[Mas: 95]

(36a)

met–in	*tet*	*čolhoro–**get***	*qarte–k*
1SG–DAT	2SG	hare–ABL	share[IMP]–2SG

‘Share some of your hare with me.’ [Mas: 113]

(36b)

tet	*čolhoro*	*kudeje*	*lek–telle*	*yaqte–ge–k*
2SG	hare	liver	eat–PF.CVB.SS	sing–DSTIMP–2SG

‘Eat some hare liver and then sing!’ [Mas: 177]

In Itelmen, the Absolutive case marks the subject of intransitive verbs, the agent of transitive verbs, as well as the patient of transitive verbs (G/V: 71; Volodin 1976: 147). No distinction is made between definite and indefinite direct objects. In the related language Chukchi, however, the Absolutive is used to mark only the subject of intransitive verbs and the patient of transitive verbs, while agents of transitive verbs are marked by the Ergative case (Dnn: 104f). However, in antipassive constructions the direct object is demoted to an oblique object, e.g. an Instrumental (which is formally identical to the Ergative). In such a construction, the direct object encoded as an oblique is less individuated than the direct object marked by the Absolutive case in an ergative construction; it can be used contrastively, and also to avoid the assertion that changes in the state of the direct object are pragmatically relevant (Polinskaja & Nedjalkov 1987: 240, 244ff). However, there does not appear to be a correlation between the type of object encoding and the definiteness of the object (cf. Polinskaja & Nedjalkov 1987: 240, ex. 1a-e).

In Siberian Inupik Eskimo direct objects can be marked either by the Absolutive case, with the verb being transitive and the agent being in the Relative case, or by the Instrumental case, with the verb being intransitive and the agent being in the Absolutive case (Men: 48, 51). Other Eskimo languages follow the same pattern, and from de Reuse's description (de Reuse 1994: 30f) it becomes clear that the difference in case-marking makes a distinction between definite and indefinite direct objects, e.g.

(37a)

qikmimaneʁaa		*kayu*
qikmiʁ–ma	neʁ–aa	kayu
dog–REL	eat–IND[TR.3SG]	fish[ABS]

'My dog ate the fish.'

(37b)

qikmiq	*neʁtuq*	*kayumeŋ*
qikmiʁ	neʁ–tuq	kayu–meŋ
dog[ABS]	eat–IND[INTR.3SG]	fish–MDS[4]

'The dog ate a fish.'

[de Reuse 1994: 30, 31]

In Ket, both subjects of transitive verbs as well as direct objects of transitive verbs do not take any case-marking, being expressed by the Absolutive case. The difference in function is marked by SOV word order as well as by the order of subject and object agreement markers on the verb, with the subject agreement affix

---

[4] The Siberian Yupik Modalis case is cognate to the Siberian Inupik Instrumental case.

preceding the object agreement affix (Wer: 111f; Dul: 73f). The definiteness or indefiniteness of the direct object does not appear to be marked.

According to Mattissen's analysis (Mat: 140ff, 137), in Nivkh the primary object, i.e. the patient of monotransitive verbs and the addressee/goal/recipient of ditransitive verbs forms a complex with the verb, "while the other undergoer and the subject are external and non-case-marked participants" (Mat: 137). No distinction is made between definite and indefinite direct objects, compare for example: *hï–zadača ń–ïskïm–ya* [that–exercise[NOM] 1SG–explain–IMP.SG] 'explain that exercise to me' and *ńi čo hï–ńivx–ar–j* [1SG fish[NOM] that–person–feed–FIN] 'I am feeding fish to him/her' (Mat: 142). Interestingly, Gruzdeva (Grz: 18) gives this latter example as well, without the synthesis of the recipient and the verb (this being Mattissen's analysis), and translating it as 'I fed this man with the fish' (emphasis mine). This appears to be further evidence for the fact that the definite/indefinite distinction is not marked on direct objects.

In Mansi, the suffixless Nominative case expresses both subjects and nominal direct objects (Rse: 25). However, the definiteness or indefiniteness of the direct object is not left unmarked, as in Itelmen, Ket, or Nivkh, but rather it is expressed through differential suffixes on the verb. The suffixes of the indefinite (subjective) conjugation are used with intransitive verbs and with transitive verbs if the direct object is indefinite, while the definite (objective) conjugation is used with transitive verbs and definite objects; compare: *am ti pu:maś kńiga lowińt–**iləm*** [1SG this interesting book[NOM] read.PRS–SG.OBJ.1SG] 'I am reading this interesting book' vs. *am pu:maś kńiga lowińt–**eɣəm*** [1SG interesting book[NOM] read.PRS–SBJ.1SG] 'I am reading an interesting book' (Rse: 41). The definite conjugation can also be used without the object being overtly present (Murphy 1968: 107). According to Rombandeeva (1979: 105), a definite direct object can also be used with a verb in the indefinite conjugation, with the direct object marked by the Instrumental case. This occurs in 'beneficiary' constructions, e.g. 'feeding grass to the goat' or 'braiding hair for one's daughter'. Murphy (1968: 120) specifically states that this construction concerns "the personal object of a verb of 'giving' as the direct object, with the thing given in the instrumental, parallel to English 'present with' and 'treat to'. In many cases the personal pronoun of the recipient is present in the accusative." An Accusative case exists to mark pronominal direct objects (Murphy 1968: 113, Rse: 30).

Although in Khanty, too, there is a distinction between an indefinite (subjective) and definite (objective) verbal conjugation, definite direct objects do not always trigger the objective conjugation on the verb (Abondolo 1998a: 379; Nik: 64ff). Thus, in the following examples the direct objects are identical, yet in one

instance the subjective conjugation is used and in the other the objective conjugation is used: *ma tam kalaŋ we:l–s–ə–m* [1SG this reindeer[NOM] kill–PST–E–SBJ.1SG] vs. *ma tam kalaŋ we:l–s–e:m* [1SG this reindeer kill–PST–SG.OBJ.1SG] 'I killed this reindeer' (Nik: 64). The distinction between the object that triggers object agreement on the verb and the one that doesn't appears to lie in pragmatic principles: the former is a secondary topic within the clause, while the latter has focus status, introducing a new referent into the discourse (Nik: 76; Skribnik 2001: 357, 359).

In Selkup, direct objects can remain in the unmarked Nominative case or they can take Accusative case marking. The distinction in use between the two cases is not entirely straightforward, but it does not primarily serve to mark the distinction between indefinite and definite direct objects. According to the authors of the *Očerki sel'kupskogo jazyka* (OSJ: 172, 174), the Accusative case is the most frequent, typical way to mark direct objects. It is used when the distinction between the subject and the direct object is not clear from the context, i.e. when marking of the direct object is necessary to disambiguate the meaning of the sentence. When such disambiguation is not necessary, the Nominative and Accusative case appear to be in free variation (OSJ: 382, 384). In the second person Imperative, however, nominal direct objects are obligatorily marked by the suffixless Nominative case, while pronominal objects take Accusative case-marking (OSJ: 383). The definiteness of direct objects in Selkup is achieved mainly via possessive marking, even if from the context no possession is actually present (OSJ: 385); on the other hand, the indefiniteness of the direct object can be emphasized by the use of the indefinite (subjective) conjugation, which is, however, rare (OSJ: 234).

In Nganasan, definite direct objects in the singular are marked by the Accusative case, with indefinite direct objects remaining unmarked or, for those verbs that have it, appearing with the suffixless oblique stem (Ter: 80). As in Selkup, direct objects are most frequently marked by the Accusative case (Ter: 80); this implies that in the dual and plural number Accusative case-marking does not express the definiteness of the direct object. The verbal conjugation does not serve to express the definiteness or indefiniteness of the direct object, unlike what is found in Mansi (Ter: 190).

Table 3.7 presents a summary of the direct object case-marking found in the languages examined so far. I have simplistically assumed that differences in case-marking make a distinction between definite and indefinite direct objects; although this is not fully accurate, it appears to be a somewhat acceptable common denominator, which accounts for the majority of instances.

Table 3.7: Direct object marking in some languages of Eurasia

Language(s)	Agent	Patient		
		Definite	**Indefinite**	**Partitive***
Turkic**	NOM	ACC	NOM	NOM
Tofa	NOM	ACC	NOM	**PART**
Sakha	NOM	ACC	NOM/**PART**	**PART**
Written Mongolian	NOM	ACC	NOM	ABL
Khalkha	NOM	ACC	NOM	ABL
Buryat	NOM	ACC	NOM	NOM
Evenki	NOM	DEF.ACC	**INDF.ACC**	**INDF.ACC**
Yukaghir	NOM$_{1/2}$	ACC$_{1/2}$	—	—
	NOM$_{1/2}$	NOM$_3$	NOM$_3$	NOM, ABL
	NOM$_3$	ACC	INS	
Itel'men	ABS			?
Chukchi	ERG	ABS		?
	ABS	OBL		?
Siberian Inupik	REL	ABS	—	?
Eskimo	ABS	—	INS	?
Ket	NOM			?
Nivkh	NOM			?
Mansi	NOM	NOM + Verb$_{OBJ}$	NOM + Verb$_{SBJ}$	?
Khanty	NOM	(NOM + Verb$_{OBJ}$)	(NOM + Verb$_{SBJ}$)	?
Selkup	NOM	(POSS)	(Verb$_{SBJ}$)	?
Nganasan	NOM	ACC	NOM	?

*this refers to partially affected mass nouns in direct object position
**excluding Sakha and Tofa

As can be seen from Table 3.7, the majority of languages in Eurasia/Siberia use differential case-marking to make a distinction between definite and indefinite direct objects (or possibly specific/unspecific, or topical/rhematic direct objects). However, for the indefinite direct object most of them use the unmarked or nominative case that is also used to express the subject, in opposition to a marked accusative case to express the definite direct object. Only a few languages use a case other than the nominative to mark indefinite direct objects. These are: Sakha (in the affirmative Imperative mood), Evenki, Yukaghir (with third person subjects), and Eskimo. The situation in Siberian Inupik Eskimo is different from that in Sakha,

Evenki, and Yukaghir, however, in that here the construction used to express indefinite direct objects is in effect an intransitive one, with the subject in the Absolutive case, the direct object in an oblique (Instrumental) case, and the verb taking only subject agreement. The definite direct object construction, on the other hand, is a transitive one, with the subject being in the Relative case and the direct object in the Absolutive; the clearly transitive verb takes both subject and object agreement marking. Yukaghir differs from both Sakha and Evenki in that the distinction is restricted to third person subjects. It is notable that in both Sakha and Evenki the case marking of indefinite direct objects is also used to mark partially affected direct objects. In this, Sakha and Evenki differ from all the other languages examined. Although Tofa, too, has a separate Partitive case to mark partially affected direct objects, this does not mark indefinite direct objects. The use of the Partitive case (shared in its partitive meaning between Tofa and Sakha) as a marker of indefinite direct objects in Sakha alone among the Turkic languages thus appears to be a relatively clear-cut case of language contact influence from Evenki on Sakha.

The development of the partitive meaning of the Turkic Locative case in Sakha and Tofa may, on the other hand, be due to contact influence from Mongolic. In Written Mongolian the Ablative case has partitive functions, and the Old Turkic Locative had ablative functions in addition to its expression of location. Thus, the development of the meaning of ablative to partitive in Sakha and Tofa may be due to Mongolic influence. On the other hand, since the grammaticalization of ablatives to partitives is cross-linguistically common, and since it did not take place in other South Siberian Turkic languages also in contact with Mongolian (e.g. Tuvan), it may also be due to an internal innovation in Sakha and Tofa. This may have happened prior to the migration of the Sakha to the north, at a time when the lects ancestral to Sakha and Tofa may well still have been spoken in close proximity, i.e. it could be indicative of contact between the ancestors of the Tofa and the Sakha.

However, although the explanation of Evenki influence on the development of the indefinite accusative meaning of the Sakha Partitive case seems straightforward, there is a serious problem with it: an indefinite accusative case is not commonly found in the Tungusic language family, making Evenki nearly as big a genealogical outlier as Sakha. This will be discussed in detail in the following sections.

3.2.3.6 Case-marking of direct objects in the Tungusic languages

As can be seen from Appendix 4, the (Definite) Accusative case is found in all Tungusic languages, including Manchu; the suffix is cognate in all the languages, consisting of a labial consonant (mostly the labial fricative or glide *–v/–β/–w*) and a low vowel. In Èven, definite (38a), indefinite (38b) as well as partially affected direct objects (38c) are expressed by the Accusative case. Furthermore, when the subject is coreferential with the possessor of the direct object, the direct object stands in the unmarked Nominative case with the reflexive possessive suffix (Mal'čukov 1999: 92 – 38d); on its own the Nominative case does not serve as a marker of direct objects (Ben: 56). A third means of expressing direct objects in Èven is the so-called Designative case; this will be described in more detail in section 3.2.3.8 below.

(38a)

tara–β   ora–m   xepke–li
that–ACC   reindeer–ACC catch–PRXIMP.2SG

'Catch that (definite, specific) reindeer!'

[Elena Nesterova, pers. comm.]

(38b)

kuma: olra–β   jeb–bo:t–te–n
freshwater.seal fish–ACC   eat–HAB–NFUT–3SG

'The freshwater seal eats fish (freshwater seals eat fish).'

[Rišes & Cincius 1952: 770]

(38c)

min–du   (abal–u)   ulre–β   nya:n (abal–u)   kileb–u
1SG–DAT   (a.bit–ACC) meat–ACC again (a.bit–ACC) bread–ACC

bu:–li
give–PRXIMP.2SG

'Give me (some) meat and (some) bread.'

[Elena Nesterova, pers. comm.]

(38d)

bi   ayaβ–ri–y   ike:–y   ike:–βe:t–te–m
1SG   love–PTCP–PREFL.SG   song–PREFL.SG   sing–HAB–NFUT–1SG

'I used to sing my favourite song.'

[Ben: 82]

In Udihe, the unmarked Nominative case can serve to express non-specific or newly mentioned direct objects (N/T: 119, 120f, 122 – 39a), while the Accusative

case appears to mark specific direct objects[5] (39b). If the subject of the transitive verb is coreferential with the possessor of the direct object, the Nominative case is used together with reflexive possessive suffixes (N/T: 119). The direct object also remains unmarked under certain phonetic conditions, e.g. if the stem ends in *–wA* or *–fA* (N/T: 120). Furthermore, direct objects can be marked with the so-called Destinative case; this is discussed in more detail below (section 3.2.3.8). In the Imperative mood, the direct object is "likely not to have the Accusative marker" (N/T: 122 –39c). However, the two examples that follow on p. 122 of Nikolaeva & Tolskaya are both with indefinite direct objects, while in their discussion of the function of imperatives (N/T: 264), the two transitive examples have Accusative-marked direct objects, even though the translation of one of these examples is partitive (39d). Definiteness (not only of the direct object, but of a noun in general) can moreover be marked on the noun by the use of the third person singular possessive suffix (N/T: 131 – 39e).

(39a)

   uti sita–ni   bui   magi–ə–ni
   that son–3SG   animal   kill–PST–3SG

   'Her son was killing animals.'

                                           [N/T: 121]

(39b)

   *bi   coŋku–**wə**   ńiəntilə:–mi*
   1SG   window–ACC open.PST–1SG

   'I opened the window.'

                                           [N/T: 512]

(39c)

   zəkpu–nə–i   zəu
   eat–DIST[6]–IMP.2SG   food

   'Come to eat some food.'

                                           [N/T: 122]

---

[5] Nikolaeva & Tolskaya (N/T: 119ff) do not specifically say this; however, since they discuss the instances when the direct object remains unmarked, one can assume that the direct object is marked in the remainder of instances, i.e. when the object is specific.

[6] Nikolaeva & Tolskaya (N/T: 122) gloss this as DEST, which is their abbreviation for the Destinative case. Since it does not make sense to have case marking on a verb, I feel that this must have been a typo, and that the suffix is meant to be the Distributive Aktionsart, which is marked by the suffix *–nA* and is glossed by them as DIST. Although Nikolaeva & Tolskaya (N/T: 313) say that "… the Distributive suffix *–nA–* is restricted to intransitive verbs", they also say that "on rare occasions" it "has an imperfective meaning: *oŋmo–no–* 'often forget' (*oŋmo–* 'forget')" (N/T: 314).

(39d)

   min–du   *ulə:–**wə***   *xauliə*   *bu–yə*
   1SG–DAT   meat–ACC   please   give–IMP.2SG

   'Please, give me some meat.'

                                                            [N/T: 264]

(39e)

   əi   *mo:–wa–**ni***   *kusigə–zi*   *ə–zi*   *tinda*
   this tree–ACC–3SG   knife–INS   NEG–IMP.2SG   cut

   'Do not cut these trees with knives.'

                                                            [N/T: 131]

With regard to the example (39e), Nikolaeva & Tolskaya say that "[…] the potential possessor cannot be reconstructed at all, and the only function of the suffix *–ni* seems to be to encode situational definiteness." (N/T: 131).

In Nanay, the Accusative case is the default case for expressing direct objects, although the unmarked Nominative case used to be quite common as well. However, this function of the Nominative has slowly but surely ceded to the Accusative, especially in the standard language, where it is hardly used anymore to mark direct objects (AvrIII: 155f). The distinction between the Nominative and Accusative to mark direct objects is not quite clear: the Nominative used to be somewhat more frequent with indefinite direct objects, but this is not a rule, and definite direct objects can also stand in the Nominative case. Furthermore, in all instances with indefinite direct objects the Accusative can be used instead. Unmarked direct objects are avoided when they might be confused with subjects, and personal pronouns in direct object position are never used in the Nominative case (AvrIII: 156ff). A third option of marking direct objects is the Designative case, which is discussed in more detail in section 3.2.3.8 below. In the Imperative mood, the Accusative appears to be used with both definite and indefinite direct objects as well as with partially affected mass nouns: *ey daŋsa–**va** učitel'–du bu:–xeri* [this book–ACC teacher-DAT give–DSTIMP.2SG] 'give this book to the teacher' (AvrI: 181); *mi neu–du–ive=de tetue–**ve** aŋosi–ru* [1SG younger.sib–DAT–??=PTL clothes–ACC make–PRXIMP.2] 'and sew clothes for my younger brother'; *si min–du amtaka–**va** ga–du* [2SG 1SG.OBL–DAT berry–ACC buy–PRXIMP.2] 'you buy me berries (ты мне ягод купи)' (AvrIII: 169, 168).

In Oroč as well direct objects can stand in the unmarked (Nominative) case, in the Nominative plus reflexive possessive suffixes (when the direct object is possessed and the subject and possessor are coreferential), in the suffixally marked Accusative case (A/B: 108, 109, 111), or in the Designative case discussed below (section 3.2.3.8). The differences in use between the Nominative and Accusative are

not clear and are not addressed by Avrorin & Boldyrev, but the two cases appear to be in free variation at least occasionally. Thus, pairs of sentences differing in their marking of the direct object are provided as examples, e.g. *eńe: e:ki boggo beyu va:–xa–ńi* [mother[VOC] frog fat elk kill–PST–3SG] 'mama, the/a frog killed the/a fat elk' vs. *e:ki boggo beyu–me va:–xa–ńi* [frog fat elk–ACC kill–PST–3SG] 'the/a frog killed the/a fat elk' (A/B: 108). In the Imperative mood, it appears from some examples that partially affected mass nouns may remain unmarked, while completely affected mass nouns take Accusative case marking; this is in good accordance with the description of the basic function of the Accusative as being the marking of the "direct action on an object which is completely affected by the process" (A/B: 110). Compare for example: *siŋe min–du iŋekte bu:–ɣe* [mouse 1SG.OBL–DAT bird.cherry give–IMP] 'mouse, give me (some) bird cherries (дай мне черёмухи)' vs. *e:ke, iŋekte–ve bu:gi–ɣe* [frog[VOC] bird.cherry–ACC give.back–IMP] 'frog, return the berries!' (A/B: 317, 318). However, this distinction does not always hold, for example: *iŋekte–ve ičene–ve* [bird.cherry–ACC go.and.look–IMP] 'go and look for bird cherries (пойди поищи черёмухи)!' (A/B: 318), i.e. the use of the Accusative case marker with a non-referential direct object.

In Manchu, both the Accusative (marked by the separate case particle *be*; Gor: 166) and the Nominative can mark direct objects. Non-specific (generic) direct objects stand in the unmarked Nominative case (Gor: 163). In the Manchu dialect Sibe, the Accusative case appears to mark specific direct objects; in Classical Manchu, on the other hand, although it appears to be connected with notions of definiteness, the functions of the Accusative are not quite as straightforward and need further study (Gor: 172). It is obligatory, however, when the direct object does not occupy the immediately preverbal position (Gor: 171).

As can be seen from the above description and from the summary in Table 3.8 below, with the exception of Evenki the Tungusic languages do not have a separate Indefinite Accusative case to mark indefinite direct objects. The only other Tungusic language for which such a case is mentioned is Negidal, a Northern Tungusic language closely related to Evenki. Unfortunately, I have no information on the use of the Indefinite Accusative case in Negidal; Kolesnikova & Konstantinova (JaN5: 113) and Cincius (1982: 27) only list the Indefinite Accusative case [with the suffix *–(y)a*] in tables of Negidal cases, next to the (Definite) Accusative case (with the suffix *–va*). Xasanova & Pevnov (2003: 245) say that the Negidal case system is similar to the Evenki system; in their description of discrepancies between the two languages (245-250) they do not mention the Indefinite Accusative case. This might be an indication that the Indefinite Accusative in Negidal has the same functions as the Evenki case.

Table 3.8 Case suffixes used to mark direct objects in Tungusic languages

	Definite direct object		Indefinite direct object	
Evenki	(DEF) ACC	*–vA/–mA*	INDF.ACC	*–(y)A*
Negidal	(DEF) ACC	*–vA*	INDF.ACC	*–(y)A*
Èven	ACC	*–v/–m/–bu/–u*	ACC	
Nanay	ACC	*–vA/–bA*	ACC/NOM	
Oroč	ACC	*–vA*	ACC/NOM	
Udihe	ACC	*–vA/–mA*	NOM	
Manchu	ACC	*be*	ACC/NOM	

From the discussion so far it is quite clear that a separate case functioning as a marker of indefinite or partially affected direct objects is rare in Siberia and Eurasia. Thus, none of the Tungusic languages (except for Negidal, a language closely related to Evenki) has a case comparable to the Evenki Indefinite Accusative – not even Èven, which also belongs to the Northern Tungusic branch of the Tungusic language family. In the Turkic language family, only Tofa and Sakha have a separate Partitive case, but this is restricted to the expression of partially affected mass nouns in Tofa, whereas it fulfills additional functions of an indefinite accusative in Sakha. The fact that Evenki and Sakha both have a case marking indefinite direct objects, which is not a common feature of their respective languages families and is rare in Eurasia, is strongly indicative of contact influence (cf. Heath 1978: 23; Gensler 1993: 33f, 46). However, since this case is a common feature of neither language family, it is not clearly inherited in either Evenki or Sakha; therefore, it is very hard to come to a conclusion concerning the direction of contact influence. An examination of further functions of the Evenki Indefinite Accusative case in comparison with the other Tungusic languages may, however, shed some light on this issue.

3.2.3.7 Additional functions of the Evenki Indefinite Accusative

In addition to marking indefinite and partially affected direct objects, the Indefinite Accusative case in Evenki has a designative function, i.e. it marks a direct object that benefits someone. The beneficiary is encoded in the case-marked direct object through obligatory possessor marking – either through reflexive possessive suffixes, when the the beneficiary is coreferential with the subject (40a), or through personal possessive suffixes, when the subject and beneficiary are not coreferential

(40b, c). Thus, the object simultaneously encodes the direct and the indirect object (Ned: 147):

(40a)

jepïle–ye–ver	*ga–kaim*	*suru–če–tïn*
food–INDF.ACC–PREFL.PL	take–CVB	go.away–PST–3PL

'Taking food for themselves, they went away.'

(40b)

su	*unta–ya–n*	*o:–kallu*
2PL	fur.boots–INDF.ACC–POSS.3SG	make–PRXIMP.2PL

'You[PL] make fur boots for him!'

(40c)

jav–ya–v	*o:–kal*
boat–INDF.ACC–POSS.1SG	make–PRXIMP.2SG

'Make a boat for me.'

[Ned: 147, 148]

In addition, the Evenki Indefinite Accusative is frequently used in privative constructions with the postposed negative noun *a:čin* 'none' (Ned: 159, B/G: 9):

(41a)

jal–ya	*a:čin*	*beye*
intellect–INDF.ACC	NEG	man

'a man without intellect'

[Ned: 159]

(41b)

agi:–du:	*ŋina–ya*	*a:čin*	*e:kun*	*ayan*	*bi–jeŋe:–n*
taiga–DAT	dog–INDF.ACC	NEG	what	good	be–FUT–3SG

'Nothing good will come of being in the taiga without a dog.'

[B/G: 9]

However, it should be noted that Konstantinova (Kon: 78) rejects the analysis of the suffix marking the object of privation as the Indefinite Accusative case marker, claiming this is coincidental homonymy. She compares the suffix participating in the privative constructions with a formally identical suffix marking associative plurals, e.g. *Pačekiya* 'Pačeki and his relatives or friends' as well as participating in coordinated noun phrases in the Comitative case, e.g. *Stepan–a Volodya–nun* [S.–SOC V.–COM] 'Stepan with Volodya' (Kon: 77). In this, Konstantinova argues that the suffix *–(y)A* found in the privative construction and that found in the associative construction are the same in form, function and meaning (Kon: 78). Nedjalkov (Ned: 190f) does not discuss the similarity in form of

the Associative Plural (called Collective by Nedjalkov) and the Indefinite Accusative; however, he analyses the Associative Plural suffix as containing an obligatory glide (–*yA*), i.e. as having a slightly different form than the Indefinite Accusative case marker.

3.2.3.8 Designative case and privative constructions in other Tungusic languages

As was already indicated in section 3.2.3.6, a designative case is found in other Tungusic languages as well, with the exception of Manchu. The functions of this case are the same in all the languages that have it: the Designative [called Indefinite Accusative for Èven by Benzing (Ben: 57f) and Destinative for Udihe by Nikolaeva & Tolskaya (N/T: 126, 600)] simultaneously marks a direct object that benefits somebody and the beneficiary, the latter being marked by possessive suffixes on the object (NovI: 188ff; N/T: 126, 600ff; AvrIII: 149ff; Pet: 51f; A/B: 115ff). The case suffix, however, differs between the languages (cf. Table 3.9): in Udihe, the Destinative case suffix is –*nA*, e.g. *bi zugdi–**nə–i** wo:–iti* [1SG house–DES–1SG make–3PL] 'they are making a house for me' (N/T: 601); in Nanay the suffix is –*goa*; in Orok it is –*do/–du*, e.g. *bi apun–**do–si** ga–tči–mbi* [1SG hat–DES–2SG take–PST–1SG] 'I took a hat for you' (Pet: 52), in Oroč it is –*nA:/–yA:/–lA:*, and in Èven, the Designative case suffix is –*GA*:

(42a)
bi	*etiken*	*ora–ŋ–**ga–n***	*emu–re–m*
1SG	old.man	reindeer–ALN–DES–3SG	bring–NFUT–1SG

'I brought the reindeer for the old man.'

[Mal: 10]

(42b)
mut	*pektere:β–**ŋe:–βur***	*uni–ri–t*
1PL.incl	gun–DES–PREFL.PL	buy–PST–1PL.INCL

'We bought ourselves a gun.'

[Ben: 57]

The privative constructions in the Tungusic languages vary (cf. Table 3.9): in some languages, the noun expressing the entity that is lacking remains in the unmarked (Nominative) case, while in others it is marked by a case suffix –*lA* [called 'Partitive' by Benzing (Ben: 30) and Nikolaeva & Tolskaya (N/T: 825f)]. The negative element that is part of the construction also differs somewhat in form between languages; in general, it is postposed, but in Èven and the Lower dialect of Negidal it is preposed (cf. Table 3.9). Nikolaeva & Tolskaya (N/T: 145) call this the

'negative copula'; however, given that this may take case-marking (N/T: 146), it might be more correct to talk of a negative noun.

Thus, in Ėven we find the preposed negative element *ač* followed by the noun marked by the suffix *–lA*, e.g. *tar–al asa–l **ač** hu–**le**–sel* [that–PL woman–PL NEG child–LA[7]–PL] 'These women don't have children' (Ben: 30). In Negidal there is a dialectal split in the expression of privative constructions: the Upper dialect Privative construction is identical to the Evenki Privative, with the noun expressing the object of privation carrying the Indefinite Accusative suffix *–(y)A* and the negative element *a:čin* postposed, e.g. *xute–**ye** **a:čin** beye* [child–INDF.ACC NEG man] 'a childless man', while the Lower (Amgun') dialect follows the Ėven construction, with the preposed negative particle *a:čin* followed by the noun marked by the suffix *–lA*, e.g. ***a:čin** xute–**le** beye* [NEG child–LA man] 'a childless man' (Xasanova & Pevnov 2003: 244).

In Udihe, the Privative construction consists of the noun expressing the lacking entity marked by the suffix *–lA*, and the postposed negative noun *anči*, e.g. *nuani kəsi–**lə** **anči*** [3SG luck–LA NEG] 'He does not have luck' (N/T: 826), *bi in'əi–wə isə:–mi igi–**lə** **anči**–wə* [1SG dog–ACC see.PST–1SG tail–LA NEG–ACC] 'I saw a dog without a tail' (N/T: 146). There is a distinction whether a proprietive or a circumstance construction is being negated (N/T: 836, 837): in the latter instances, the negative noun *anči* takes either Instrumental or Ablative case-marking, and the object of privation is either marked with *–lA* (in conjunction with *anči* in the Instrumental case) or remains unmarked (with the Ablative case of *anči*). There is no discernible semantic difference between the two constructions[8], e.g. *bi čaya–wa umi–mi sata–**la** **anči**–**zi*** [1SG tea–ACC drink–1SG sugar–LA NEG–INS] 'I drink tea without sugar', *əi suəsə **anči**–**digi**–ni nuani ə–ini xuli* [this axe NEG–ABL–3SG 3SG NEG–3SG walk] 'he does not walk without this axe' (N/T: 836, 837).

There are two varying (very brief) descriptions of the Privative construction in Orok: on the one hand, the construction is described as being similar to the Udihe Privative, with the noun denoting the lacking entity carrying the suffix *–lA* followed by the negating element *ana*, e.g. *asi–**la** **ana** nari* [woman–LA NEG person] 'unmarried person/man' (JaN5: 177). On the other hand, the Orok Privative is described as consisting of the noun in the Nominative followed by the negating

---

[7] Although Benzing (Ben: 30) as well as Nikolaeva & Tolskaya (N/T: 825f) call this suffix the 'Partitive', I prefer to gloss it as *–LA* to avoid confusion with 'real' partitives that mark partially affected mass nouns.

[8] Interestingly, although Nikolaeva & Tolskaya (N/T: 836) say that there is no discernible semantic difference between the two options, all the examples with *anči* in the Instrumental case have affirmative verbs, while the two examples with *anči* in the Ablative case are negated sentences.

element *ana/anaɣa*, e.g. *min–du ula anaɣa* [1SG.OBL–DAT reindeer NEG] 'I don't have a reindeer' (Pet: 46). However, further on in this grammar sketch, the Privative is again described as consisting of the suffixally marked noun followed by *ana*, e.g. *ula–la ana nari* [reindeer–LA NEG person] 'a person without reindeer, a reindeerless person' (Pet: 57). The difference may possibly lie in the attributive vs. predicative use of the construction; however, I cannot say that with certainty based on such a very limited number of examples.

In Oroč, there are two privative constructions. The first is synthetic, with the noun that expresses the lacking entity taking the Privative suffix *–LAči*, e.g. *digga–lači* 'mute' (< *digga* 'voice'), *ugdalači* 'without a boat' (< '*ugd*a' boat) (A/B: 225). As shown by Xasanova & Pevnov (2003: 245), this is the grammaticalized form of a privative construction similar to that found in Udihe, i.e. consisting of the noun carrying the suffix *–lA* followed by the negative element *ači(n)*. The second Oroč Privative consists of an unmarked noun followed by the negative element *ana*, e.g. *bi: ami=da: ana eńi=de: ana* [1SG father=EMPH NEG mother=EMPH NEG] 'I have neither father nor mother' (lit. 'I am without father and without mother') (A/B: 226). Nanay is described as having a privative construction that is similar to the second variant of Oroč, that is, the noun expressing the lacking entity remains unmarked and is followed by the negative element *ana:* (AvrI: 196).

In Manchu, the unmarked noun is followed by the negative particle *aku:*, which also functions as a negative copula, e.g. *gu:nin aku: niyalma* [brains NEG person] 'stupid person' (lit. 'a person without brains') (Gor: 372f). Table 3.9 summarizes the forms of the designative case and the privative constructions in the Tungusic languages.

Table 3.9: Designative case suffixes and privative constructions in Tungusic languages

	Designative case	**Privative construction**
Evenki	*–(y)A* (= INDF.ACC) + POSS	*–(y)A a:čin*
Negidal	??	*–(y)A a:čin* *a:č –lA*
Ėven	*–GA* + POSS	*a:č –lA*
Nanay	*–goa* + POSS	NOM *ana:*
Orok	*–do/–du* + POSS	*–lA ana* (NOM *ana/anaɣa*)
Oroč	*–nA:/–yA:/–lA:* + POSS	*–LAči* (< *–LA* + *ači(n)*) NOM *ana*
Udihe	*–nA* + POSS	*–lA anči*
Manchu	---	NOM *aku:*

The variation in the designative case suffixes is rather big, and it is unlikely that they are cognate (Andrej Malchukov, Igor Nedjalkov, pers. comm.). The suffix –*lA* occurring in privative constructions in combination with a negative element is fairly widespread in the Tungusic languages. Petrova (Pet: 58) suggests that the suffix –*lA* is connected to the Indefinite Accusative case found in Evenki and Negidal, claiming that –*lA* is used with an indefinite accusative meaning in Negidal (Pet: 57). However, this could also be due to an extension of the functions of the suffix –*lA* in parallel to the functions of the Indefinite Accusative suffix. Whether the allomorphs of the Oroč Designative case are an indication that this case is related to the Evenki Indefinite Accusative on the one hand and to the 'privative' suffix –*lA* on the other, is unclear. If there should be such a connection, this would provide some indication that the Designative case and the Privative suffix are historically related, as well as providing a link to the Evenki Indefinite Accusative. However, such a relationship cannot be confirmed at this moment.

3.2.3.9 The origins of the indefinite accusative function of the Sakha Partitive case

Since the function of a designative case and the way of expressing the lack of an entity (with a suffixally-marked noun followed by a negative element) is found in nearly all of the Tungusic languages, these might be ancestral features of the Tungusic language family, although the forms expressing them have changed in the individual languages. The indefinite accusative and partitive function of the Evenki Indefinite Accusative case, however, are restricted to Evenki (and apparently to Negidal), and it is unclear whether this is due to language-internal development, to contact, or whether these functions used to be more widespread in the language family as a whole, but got lost elsewhere. In this connection, it is interesting to note that Benzing calls the Ėven Designative case the Indefinite Accusative (Ben: 57), which implies that he at least believes it to be connected to the Evenki Indefinite Accusative. In a similar vein, the use of the term 'partitive' (N/T: 825f) for the suffix marking the Udihe Designative case is curious, since there is no indication that nowadays this suffix has any truly partitive functions. However, it is not possible at this moment to decide whether all the functions of the Evenki Indefinite Accusative are inherited or whether the case suffix was originally restricted in its functions to the Designative case and Privative construction, with a later extension to include the marking of indefinite and partially affected direct objects. Thus, it is difficult to

judge the direction of contact influence involved in the development of the Sakha Partitive with its indefinite accusative[9] function.

There are two possible scenarios for the development of an indefinite accusative function of the Sakha Partitive: a) language-internal development in Sakha and b) Evenki contact influence. In either case it is clear that both Sakha and Tofa initially extended the meaning of the Common Turkic Locative case (in its ablatival meaning) to a Partitive case following a cross-linguistically common grammaticalization path from ablatives to partitives (cf. section 3.2.3.3). In this, Mongolic contact influence may have played a role, since the Written Mongolian Ablative case has a partitive function, as does the Middle Mongolian Ablative (Rybatzki 2003: 68). However, a purely language-internal development of this function is equally possible.

Following upon this, according to scenario a) Sakha speakers independently extended the functions of the Partitive case to the expression of indefinite direct objects. After coming into contact with speakers of Evenki, the latter copied the indefinite direct object and partitive meaning of the Sakha Partitive case, making use of the suffix used for their Designative case and in privative constructions. Since the Negidals are descendants of Evenks (Black 1988: 25, Forsyth 1992: 207, Janhunen 1996: 67, 72f, 79 *inter alia*) or are descended from a common ancestor with the Evenks (Xasanova & Pevnov 2003: 285), it is probable that the Tungusic speakers the Sakha came into contact with were not yet differentiated into Evenks and Negidals. Under this assumption, the ancestral Evenk-Negidal group copied the functions of the Sakha Partitive case and only later diverged into two ethnic groups speaking closely related languages. The Tungusic groups that later developed into Ėvens must already have been separated geographically from the Evenk/Negidal ancestors and the incoming Sakha, because they did not copy the indefinite accusative and partitive function.

According to scenario b), at the time of initial contact with speakers of Sakha, the Evenki Indefinite Accusative case had all four functions that it carries nowadays. After migrating to the middle Lena, the Sakha came into contact with Evenks and copied the indefinite accusative function of the Evenki Indefinite Accusative case. For this, they extended the function of their Partitive case suffix, since the partitive functions of the Evenki Indefinite Accusative and the Sakha Partitive are very similar.

In either scenario, it remains unclear how the initial development of an indefinite accusative function of the case suffix of the model language came about.

---

[9] As described above (section 3.2.3.1), what I call here an 'indefinite accusative' function is the marking of non-specific direct objects.

It is, of course, tempting to attribute this to earlier contact influence, e.g. to influence from Yukaghir. As discussed above (section 3.2.3.5), Yukaghir uses the Accusative case to mark definite direct objects, and the Instrumental case to mark indefinite direct objects when the subject is third person. Thus, this language makes a formal distinction between definite and indefinite direct objects, albeit a restricted one. Yukaghir influence could, however, be plausibly claimed only for Evenki, since speakers of Tungusic languages migrated to the north and northeast (i.e. to current-day Yakutia) prior to the migration of the Sakha ancestors (cf. section 1.1.2.2); they would thus have been in contact with Yukaghirs before the arrival of the Sakha on the scene. However, in this case it is hard to explain why Èven lacks the Indefinite Accusative case found in Evenki and Negidal: early Yukaghir-Tungusic contacts would have included the ancestors of the Èvens as much as the ancestors of the Evenks and Negidals, and after Russian colonization at least the Èvens were in closer contact with Yukaghirs than the Evenks were. Thus, a Yukaghir source of the separate case to mark indefinite/non-specific direct objects is not very plausible. Since no other candidate can be found, the initial development of this case function in either Sakha or Evenki cannot be explained at this moment.

A factor that may have played a role in facilitating the copying of the indefinite accusative meaning between Sakha and Evenki (irrespective of the direction of influence) is that for nouns ending in vowels, the Sakha Privative construction appears to consist of a Partitive case-marked noun followed by the negative noun *suoχ* (43a), i.e. it appears to be identical to the Evenki Privative construction. For nouns ending in a consonant it is clear that this is not the case (43b), and that the Sakha Privative is formed by the object of privation carrying the 3SG possessive suffix. Nevertheless, the surface similarity in a number of privative expressions may have been enough for speakers of Sakha or Evenki to equate the Sakha way of expressing partially affected mass nouns with their means of expressing lack of an object, and to identify these constructions with the Evenki way of expressing these. This could have led to the extension of the functions of the case identified as being involved in such constructions by copying the indefinite accusative function from the other language.

(43a)

bihieχe	*ulaχan*	*suolta–ta*	*huoχ*
1PL.DAT	big	meaning–POSS.3SG	non.existence

'(It has) no meaning for us.'

[XatR118]

(43b)

kini	törüt	üöreγ–*e*	huoχ	kihi	e–bit
3SG	completely	studies–POSS.3SG	non.existence	person	AUX–PSTPT

'He was a person without any education, it seems.'

[XatR67]

The Partitive case of *suolta* 'meaning' would be *suoltata*, identical to the 3SG-possessive-marked noun form; whereas the Partitive case form of *üöreχ* 'studies' would be *üöreχte*. That it is possible for such a reanalysis to be based on only a subset of the instances involved is shown by Finnish. Here, the subjects of certain participial clauses used to take the object case-marking governed by the matrix verb (i.e. Accusative, Partitive, or Nominative case). Through phonological change, the case endings of the Accusative and Genitive for singular nouns collapsed; however, the distinction between the Accusative and Genitive case remained valid for plural nouns and for pronouns. The collapse of the Accusative and Genitive in the singular paradigm led to a reanalysis of the subjects of such participial clauses as being marked by the Genitive case, not the object case expected from the matrix verb, so that nowadays no object case is permitted in this construction (Harris & Campbell 1995: 77f).

It is very difficult (and perhaps impossible) to judge which of the two scenarios outlined above is correct; however, there is one argument in favour of scenario b) (Evenki contact influence in Sakha) rather than a) (Sakha contact influence in Evenki). This is the fact that the Evenki Indefinite Accusative case occurs in a wider range of contexts than the Sakha Partitive: it can occur in the Imperative mood, Future tense, in negative sentences, and with habitual aspect, while the Sakha Partitive is restricted to the affirmative Imperative mood. In addition, the Evenki Indefinite Accusative functions as a designative case and in privative constructions, while the Sakha Partitive is restricted to marking partially affected and indefinite direct objects. It is more likely that speakers of Sakha copied only one limited function and context of the Evenki case than that Evenki copied the functions of the Sakha case, but later extended the range of contexts in which to use the suffix. This is similar to what Heath (1978: 23, 75) calls 'penetration' in his discussion of internal reconstruction used in contact situations: "Another indication of relative antiquity of a morpheme is what I will call 'penetration'. For example, a morpheme <u>M</u> which originates as a nominal suffix […] may develop secondary functions as a verbal suffix. In other words, the morpheme has penetrated more deeply into the overall morphological structure of the language. […] In the borrowing process it may well be that only one major function of the morpheme is adopted by the borrowing language." (Heath 1978: 75). Note that King (1969: 92)

proposes that the exact opposite conclusion be drawn from differences in the range of functions[10]: "If, for example, we know that the living or attested languages A and B share a rule but that this rule is more general in the grammar of A than of B, and if we know that early contact between the two languages existed, then our assumption would be that the rule was transmitted from B (less general) into A (more general) instead of vice-versa." However, King himself acknowledges that his view contradicts the widely-held assumption that "rules tend to narrow in generality as they spread farther from the point of origin" (King 1969: 92; cf. Hock 1991: 435, 437).

It may not be a coincidence that the Partitive case in Sakha is restricted to the Imperative mood. Cross-linguistically, there is a 'sporadic tendency' of 'unusual object-marking' in the imperative, with some languages that use the accusative case to mark direct objects in the indicative mood leaving them unmarked in the imperative (Sadock & Zwicky 1996: 174f; Koptjevskaja-Tamm 2006: 189). In particular, this is typical of the Samoyedic languages Nenets, Enets, and Selkup (Tereščenko 1973: 177f). At first glance one might suspect contact influence from Samoyedic in Sakha, whereby Sakha speakers copied only the fact of marking the direct object of imperatives differently, without copying the actual means of doing this. This is somewhat doubtful, however, since in the Samoyedic languages even definite direct objects remain unmarked in the imperative mood, e.g. the translation of one of the Selkup examples given by Tereščenko (1973: 178) is 'Lock away **this water**' (Эту воду заприте[11]), with *üt* 'water' being unmarked for case. In Sakha, however, the differential object marking is restricted to generic indefinite, non-referential and partially affected direct objects. In addition, the Samoyedic languages make a distinction between the Accusative and Nominative case in the indicative mood, and use the Nominative in the imperative mood, while Sakha has developed a third case to mark a subset of direct objects in the imperative mood. Furthermore, the overlap in function between the Sakha Partitive and the Evenki Indefinite Accusative, as well as the seeming overlap between the Sakha and Evenki Privative constructions (at least for Sakha nouns ending in a vowel) do point towards Evenki influence. However, the slight cross-linguistic tendency for direct objects of imperatives to take differential case-marking may have facilitated the development of the Common Turkic Locative case to a Partitive in Sakha and Tofa.

A further indication that the direction of contact influence may have been from Evenki to Sakha is the fact that Dolgan, which is assumed to have undergone more Evenki influence than Sakha, has copied more of the functions of the Evenki

---

[10] I thank Bernard Comrie for bringing this publication to my attention.
[11] I thank my Russian colleagues for help in interpreting this sentence.

Indefinite Accusative case[12]. Thus, in Dolgan direct objects with a Partitive case suffix in the possessive declension express an object intended for some beneficiary (44a), i.e. this function parallels the use of the Evenki Indefinite Accusative case as a designative case marker. However, it should be noted that the Dative case-marked beneficiary is retained here, in contrast to similar constructions in Evenki. This is a function not found in Sakha, where such constructions involve a Dative-marked indirect object (44b). Furthermore, the range of contexts in which the Dolgan Partitive is used is wider, since it is frequently used with verbs in the Future tense (44c) and Conditional mood (44d) as well (Ubr: 117; Artem'ev 1999b: 112), whereas it is restricted by and large to the affirmative Imperative mood in Sakha.

(44a) Dolgan:

h–anï–ka:n   *mińieke*   *bolop–**puna***   *oŋor*
EMPH–now–EMPH   1SG.DAT sword–PART.1SG   make[PRXIMP.2SG]

'Make a sword for me right now!'

[Ubr: 118]

(44b) Sakha:

untu:–buna***   *oŋor*
*unty–PART.1SG   make[PRXIMP.2SG]

mieχe   *untu:–**ta***   *oŋor*
1SG.DAT   unty–PART   make[PRXIMP.2SG]

'Make unty (reindeer fur boots) for me!'

[translation, Ver]

(44c) Dolgan:

*moro:sko–**to***   *komuy–uom*   *baltï–bar*   *bier–iem*
cloudberry–PART   gather–FUT.1SG   younger.sister–DAT.1SG   give–FUT.1SG

'I'll gather some cloudberries and give them to my younger sister.'

(44d)

öl–böt   *möŋö*   *u:–**ta***   *bul–lar*
die–PRSPT.NEG   eternal   water–PART   find–COND

'If he found the water of life…'

[Artem'ev 1999b: 112]

If scenario a) (Sakha influence on Evenki) were correct, one would have to assume that Sakha innovated the indefinite accusative function of its Partitive case and then passed the indefinite accusative and partitive functions on to Evenki. Evenki at a later stage passed the designative case function to Dolgan. Under scenario b), on the

---

[12] In this, I disagree with Artem'ev (1999b: 106) who proposes that the Dolgan Partitive case represents an early stage of the development of the Sakha Partitive.

other hand, Evenki influence led to the copying of the indefinite accusative function in Sakha, and more intense Evenki influence led to the copying of further functions of the Evenki Indefinite Accusative into Dolgan, such as the designative function.

Altogether, the shared feature of a separate case to express indefinite/non-specific direct objects, which is further used to mark partially affected direct objects, is highly indicative of contact between speakers of Evenki and Sakha. Unfortunately, given the current state of knowledge of the case suffixes in the Tungusic languages, the direction of this contact cannot be resolved with complete conviction. However, following Heath's proposal of using 'internal reconstruction' to assess the relative antiquity of a specific feature in two languages in contact, it appears more probable that Sakha speakers copied the indefinite accusative function from Evenki than the other way round.

3.2.4 The distinction between an Instrumental and a Comitative case in Sakha

In contrast to the other Turkic languages, Sakha has retained the Old Turkic distinction between an Instrumental and a Comitative case. Like other features that distinguish Sakha from the other Turkic languages, there have been suggestions that the retention of this distinction may be due to Tungusic contact influence (Ubrjatova 1956: 91; 1960: 11; 1966: 49f; 1985a: 46). In the following discussion I will examine the distribution of separate instrumental and comitative cases in Eurasian languages in order to provide an assessment of this claim.

3.2.4.1 The Sakha Instrumental and Comitative

Sakha has two cases expressing some form of accompaniment/association: the Instrumental and the Comitative. The Instrumental in Sakha covers 'prototypical' instrumental functions (Stolz & Stroh 2001: 391), expressing an instrument (45a) or means (45b); it also functions in adverbials of manner (45c), cause, time, and space, e.g.

(45a)

bu	*Pöpügey*	*dieχi*	*kïssanï*	*pa:hïnan*
bu	Pöpügey	dieki	kïrsa–(n)I	pa:s–(I)nAn
this	P.	in.the.direction.of	polar.fox–ACC	trap–INS

ïlallar	*ete*	*pa:ska*	*tüheren*
ïl–Ar–LAr	e–TA	pa:s–GA	tüs–(I)Ar–An
take–PRSPT–PL	AUX–PST.3SG	trap–DAT	fall–CAUS–PF.CVB

'In the direction of Popigaj they used to take polar foxes with traps, making them fall into traps.'

[LukP87]

(45b)

satï:	*hïljaɣïn*	*atïnan*	*sïljabït*	*inogda*
satï:	sïrït–A–GIn	at–(I)nAn	sïrït–A–BIt	inogda
on.foot	go–IPF.CVB–PRED.2SG	horse–INS	go–IPF.CVB–1PL	sometimes

oɣuhunan	*hïjjabït*	*uonna DT*	*dien*	*traktorïnan*
oɣus–(I)nAn	sïrït–A–BIt	uonna DT	die–An	traktor–(I)nAn
ox–INS	go–IPF.CVB–1PL	and DT	say–PF.CVB	tractor–INS

hïjjaɣïn	*belasipedinan*	*hïjjaɣïn*
sïrït–A–GIn	belasiped–(I)nAn	sïrït–A–GIn
go–IPF.CVB–PRED.2SG	bicycle–INS	go–IPF.CVB–PRED.2SG

'You go on foot, we go by horseback, sometimes we go with oxen, and you go with a DT tractor, you go by bike.' [RaxA100]

(45c)

onu_buollaɣïna	*bihigi*	*tïaɣa*	*oloror*	*jonnor*
onu_buollaɣïna	bihigi	tïa–GA	olor–Ar	jon–LAr
CP	1PL	taiga–DAT	sit–PRSPT	people–PL

*üörü:**nen***	*ïlar*	*buollaχpït*
üör–I:–(I)nAn	ïl–Ar	buol–TAχ–BIt
be.glad–NR–INS	take–PRSPT	AUX–MDL–1PL

'Well, we, people living in the woods, took that gladly.' (lit. 'with gladness')

[XatR53]

The Comitative shows 'prototypical' comitative functions (Stolz & Stroh 2001: 391), primarily expressing a joint action between two nouns of equal standing that are most often animate, and even human (46a, b). With nouns denoting animals the Comitative case can only be used when the animals are viewed as being similar to people – compare (46c) and (46d).

(46a)

min	*baltïm*	*podrugalar**ïna:n***
min	balïs–(I)m	podruga–LAr–(t)InA:n
1SG	younger.sister–POSS.1SG	girlfriend–PL–COM.3SG

o:nńu:	*hïjjar*
o:nńo:–A	sïrït–Ar
play–IPF.CVB	IPFV–PRSPT

'My younger sister is playing with her girlfriends.'

[translated, Sun]

(46b)

bu	*oɣonńor**du:n***	*otton*	*tüörduon*		*hettis*	*hïlbïtïgar* …
bu	oɣonńor–LI:n	otton	tüört	uon	sette–(I)s	sïl–BItIgAr
this	old.man–COM	CP	four	ten	seven–ORD	year–DAT.1PL

olorobut
olor–A–BIt
live–IPF.CVB–1PL

'And this old man and I are living together in our forty seventh year….'

[Efmy229]

(46c)

*at	*ïnaχ**tï:n***	*küölge*	*u:*	*iheller*
at	ïnaχ–LI:n	küöl–GA	u:	is–Ar–LAr
horse	cow–COM	lake–DAT	water	drink–PRSPT–PL

*'A/the horse and cow drink water in the lake.'

[back translation, Katja Potapova]

(46d)

kuoska	*ït*	*oγotuna:n*	*naha:*	*üčügeydik*	*o:nńu:llar*
kuoska	ït	oγo–(t)Ina:n	naha:	üčügey-LIk	o:nńo:–Ar–LAr
cat	dog	child–COM.3SG	very	good–ADVR	play–PRSPT–PL

'The cat and puppy play nicely together.'

[translation, Ver]

Here, I was told that horses and cows are not 'relatives or friends' and therefore the Comitative case-marking is not possible, while the sentence about the cat and the puppy was readily accepted, because cats and dogs are practically family members and thus more easily personified (Liza Migalkina, Katja Potapova, pers. comm.).

Furthermore, although the Comitative is mainly used to express an association or joint action between two animate nouns, there are instances of it having not so much a comitative function as a function as an inclusive coordination marker (47a-c). In these cases, the suffix is added to both conjoined nouns and the meaning is one of total affectedness:

(47a)

χorguyu:	*haγana*	*barï*	*ölbüttere*	*daγanï*
χorguy–I:	saγa–nA	barï	öl–BIt–LArA	daγanï
be.hungry–NR	time–LOC	all	die–PSTPT–POSS.3PL	PTL

oγolu:n	*daγanï*	*ulaχannï:n*	*daγanï*
oγo–LI:n	daγanï	ulaχan–LI:n	daγanï
child–COM	PTL	big–COM	PTL

'… in the time of hunger everybody died, children died and adults died.'

[P90_6]

(47b)

biligin	*kïrjaγastï:n*	*ederdi:n*	*barï*	*kopruonovay*
biligin	kïrjaγas–LI:n	eder–LI:n	barï	kopruonovay
now	old–COM	young–COM	all	kapron

habïnan	*i:steneller*
sap–(I)nAn	i:sten–Ar–LAr
thread–INS	sew–PRSPT–PL

'Nowadays young and old all sew with kapron (similar to nylon) threads.'

[MatX2_38]

(47c)

… isti:n	*tastï:n*	*hellik*	*buolan*
is–LI:n	tas–LI:n	sellik	buol–An
inside–COM	outside–COM	tuberculosis	AUX–PF.CVB

'…inside and outside there was tuberculosis (the whole organism was infected with tuberculosis).' [MalA30]

In addition, the *Grammatika sovremennogo jakutskogo literaturnogo jazyka* (GSJa: 132) gives a number of examples with inanimate referents (48a), and even examples of the Comitative functioning as an adverbial of manner (48.b), i.e. in a similar function to the Instrumental case. According to the judgements of my consultant in the Verxojansk district, in these cases, the Comitative case serves to underline the togetherness of the arguments, and the attention of the hearer is directed somewhat towards the case-marked argument. However, judging from the examples found in my text collection as well as in the elicited data, in normal speech the Comitative is primarily used to express a joint action of human referents.

(48a)

teliege	*tahaɣas**ti:n***	*tüŋnestibit*
teliege	tahaɣas–LI:n	tüŋnehin–BIt
cart	load–COM	turn.upside.down–PSTPT

'The cart toppled over together with the load.'

(48b)

oɣo	*taŋas**ti:n***	*utuybut*
oɣo	taŋas–LI:n	utuy–BIt
child	clothes–COM	sleep–PSTPT

'The child fell asleep in its clothes.'

[GSJa: 132]

3.2.4.1.1 Other means of expressing joint actions in Sakha

In Sakha, the Comitative case is not the only means to express joint actions between two nouns of equal standing. One alternative is the postposition *kïtta* 'with', which is the Imperfective Converb of the verb *kïtïn* 'to join'. When used with animate nouns, *kïtta* is practically synonymous with the Comitative case, and in elicitated sentences they were often given interchangeably, or as synonymous alternatives (49a, b). There does appear to be a possible slight preference to use the Comitative case with kin terms, and *kïtta* with non-kin terms, but this is tenuous and may be an attempt by my consultant to be very accurate in her judgements. There also appears to be an occasional nuance of contrastive focus inherent in the use of *kïtta* with kinship terms (49c, d).

(49a)

min	*Dariya*	*eme:χsin*	*kiyi:tin*	***kïtta***
min	Dariya	eme:χsin	kiyi:t–(t)In	kïtïn–A
1SG	D.	old.woman	daughter.in.law–ACC.3SG	join–IPF.CVB

selsovet	*munńaγar*	*bardïbït*
selsovet	munńaχ–Ar	bar–TI–BIt
village.soviet	meeting–DAT.3SG	go–PST–1PL

'Old Dariya's daughter-in-law and I went to the meeting of the village soviet together.'

(49b)

min	*Dariya*	*eme:χsin*	*kiyi:tine:n*	*selsovet*
min	Dariya	eme:χsin	kiyi:t–(t)InA:n	selsovet
1SG	D.	old.woman	daughter.in.law–COM.3SG	village.soviet

munńaγar	*bardïbït*
munńaχ–Ar	bar–TI–BIt
meeting–DAT.3SG	go–PST–1PL

'Old Dariya's daughter-in-law and I went to the meeting of the village soviet together.'

(49c)

iỹe	*kï:hïna:n*	*iher*
iỹe	kï:s–(t)InA:n	is–Ar
mother	daughter–COM.3SG	walk–PRSPT

'The mother is walking with her daughter.'

(49d)

iỹe	*kï:hïn*	***kïtta***	*iher*
iỹe	kï:s–(t)In	kïtïn–A	is–Ar
mother	daughter–ACC.3SG	join–IPF.CVB	walk–PRSPT

'The mother is walking with her daughter (and not with her son).'

<div align="right">[translation, Ver]</div>

Another means of expressing joint actions is the use of the Proprietive suffix –*LA:χ* (50a); this is not restricted to animate nouns, but can be used with inanimate nouns as well (50b). The Proprietive suffix can be used interchangeably with the Comitative case and the postposition *kïtta*, as shown by examples (50c) and (50d) with practically the same meaning as (50a). All the examples were provided as an explanation by my consultant in the Verxojansk district. Lastly, when the coordinated noun phrases are proper names, the joint action can be expressed by plural marking on both names (51).

(50a)

uolla:χ   *kï:s*   *diskotekaγa*   *barallar*
uol–LA:χ   kï:s   diskoteka–GA   bar–Ar–LAr
boy–PROP   girl   disco–DAT   go–PRSPT–PL

'The/a boy and girl are going to the disco.'

(50b)

kinigele:χ   *uru:čka*   *ostuolga*   *hïtallar*
kinige–LA:χ   uru:čka   ostuol–GA   hït–Ar–LAr
book–PROP   pen   table–DAT   lie–PRSPT–PL

'The/a book and pen are lying on the table.'

(50c)

uollu:n   *kï:s*   *kensierge*   *barallar*
uol–LI:n   kï:s   kensier–GA   bar–Ar–LAr
boy–COM   girl   concert–DAT   go–PRSPT–PL

'The/a boy and girl are going to the concert.'

(50d)

uolu   ***kïtta***   *kï:s*   *ki:neγe*   *barallar*
uol–(n)I   kïtïn–A   kï:s   ki:ne–GA   bar–Ar–LAr
boy–ACC   join–IPF.CVB   girl   cinema–DAT   go–PRSPT–PL

'The/a boy and girl are going to the movies.'

(51)

onton   *Valyalar*   *Valerkalar*   *Lenskay*   *Ostuolbï*
onton   Valya-LAr   Valerka-LAr   Lenskay   Ostuolbï
then   V.-PL   V.-PL   L.   O.

dieŋŋe   *barbïttara*
die–An–GA   bar–BIt–LArA
say–PF.CVB–DAT   go–PSTPT–POSS.3PL

'Then Valja and Valerka went to the Lena Pillars.'

[Efmy392]

3.2.4.1.2 The two variants of the Sakha Comitative suffix in the possessive declension

The Comitative suffix takes the form *–LI:n* for unpossessed nouns, and *–(I)nA:n/–(I)nI:n* in the possessive declension (cf. Appendix 3), e.g. *eji:γ–**bine:n*** [older.sister–COM.1SG] 'with my older sister'. According to the *Grammatika sovremennogo jakutskogo literaturnogo jazyka* (GSJa: 142f), the variant with the high vowel in the possessive declension is the standard, literary form (and the

paradigm they give on p. 142 lists only suffixes with high vowels), while the variant with the low vowel is the colloquial form. Oddly enough, however, the sentence examples given on p. 143 contain only nouns carrying the low vowel variant of the Comitative suffix. Xaritonov (1947: 119) also lists examples with the high vowel first, stating however that in spoken Sakha ("в живой речи") the low vowel is often heard. Ubrjatova (Ubr: 123), on the other hand, claims that the variant with the low vowel is used with kinship terms, while the variant with the high vowel is used with general nouns, e.g. *Alasov iŷe–**tine:n** Da:rïya eme:χsin–ni:n kel–li–ler* [A. mother–COM.3SG D. old.woman–COM come–PST–PL] 'Alasov and his mother, granny Darya, came' (Ubr: 123). It is hard to judge the accuracy of this claim using data from the life histories I collected, since nearly all the occurrences of the Comitative in the possessive declension concern kin (and the example given by Ubrjatova also contains only a possessively marked kinship term with the Comitative suffix, not a possessively-marked general noun). However, there is one example in favour of Ubrjatova's hypothesis:

(52)

*oɣonńor**buna:n***	*maŋnay*	*bara*	*hïjjïbïtïm*
oɣonńor–BInA:n	maŋnay	bar–A	sïrït–BIt–(I)m
old.man–COM.1SG	first	go–IPF.CVB	IPFV–PSTPT–POSS.1SG

Jögüöse	*kiyi:**tini:n***	*bara*	*hïjjïbïppït*
Jögüöse	kiyi:t–(t)InI:n	bar–A	sïrït–BIt–BIt
J.	daughter.in.law–COM.3SG	go–IPF.CVB	IPFV–PSTPT–1PL

'At first I went with my husband, (then) Egor's daughter-in-law and I went together[1].'

[P95_138]

Here, the speaker worked first with her husband, who is considered kin by marriage, and she uses the Comitative marker with a low vowel. Later on she worked with an unrelated person – Egor's daughter-in-law – and for this she uses the variant with the high vowel. However, as stated above, since this is the only instance of the variant with the high vowel, it is impossible to draw any reliable conclusions from this.

---

[1] The difference in person agreement of the verb indicates a difference in relative importance of her companions to the speaker: the use of the singular person agreement in the first clause implies that the speaker is focussing the attention of the hearer more on herself, her husband being of little importance to her, while the use of plural person agreement in the second clause implies that Egor's daughter-in-law must have been someone the speaker liked, someone she enjoyed being with; the hearer's attention is directed equally at her and her companion through use of the plural verb agreement (Liza Migalkina, pers. comm.).

It should further be noted that my main consultant in the Verxojansk district consistently rejected forms with the high vowel, saying they were incorrect. However, a second consultant in the same village accepted both the forms with the low and the high vowel, although she thought that the variant with the low vowel was used more commonly in the village. Translations of sentences in which something is done with both possessively-marked kin (e.g. 'my younger sister') and possessively-marked non-kin (e.g. 'my teacher') by Katja Potapova (who originally comes from the Taatta district) and Njurgujana Petrova (who originally comes from the Suntar district) also resulted in only case endings with the low vowel being offered, e.g. *min baltï–**bïna:n** Moskva–ɣa Kreml' kör–ö bar–a sïljï–bïp–pït* [1SG younger.sister–COM.1SG Moscow–DAT Kremlin see–IPF.CVB go–IPF.CVB IPFV–PSTPT–1PL] 'I went to Moscow with my younger sister to see the Kremlin', as well as *min uču:tal–**bïna:n** Ba:taɣay–ga bar–an kel–bit–im* [1SG teacher–COM.1SG Batagay–DAT go–PF.CVB come–PSTPT–POSS.1SG] 'I went to Batagay with my teacher'. However, when asked about the acceptability of the high-vowel variants in these cases, both Katja Potapova and Nyurguyana Petrova said they were acceptable, although they preferred the variants with the low vowel. Katja Potapova's husband, however, who is from the Suntar district, had spontaneously preferred *baltïbïnï:n* over *baltïbïna:n* (Katja Potapova, pers. comm.), which contradicts Ubrjatova's hypothesis about the low-vowel variant marking specifically kinship terms. It thus appears that the variants with the low vowel are replacing the variants with the high vowel; the example above (52) comes from a woman who must have been born at the beginning of the 20th century, prior to any prescriptive tendencies in language use and might thus reflect the older usage. A search (using Word) within all of Uvarovskij's text (which was written in the mid-19th century, cf. section 2.1.1) did not find any possessively marked Comitative suffixes; it is thus unfortunately not possible to judge whether the variant with the high vowel used to be preferred previously.

As can be seen from Table 3.2, with the exception of Old Turkic the other Turkic languages do not have a separate Comitative case. Instead, they use the Instrumental postposition/clitic *bilen/bile/ile/menen* (all of which are derived from the Old Turkic postposition *bi(r)len/bi(r)le* meaning 'with, together with') to express both a comitative relation as well as an instrument or means. It is therefore possible that the retention of the Comitative/Instrumental distinction in Sakha is due to contact influence. This question shall be investigated below.

3.2.4.2 Instrumentality and comitative relations in Turkic languages

Old Turkic has both an Instrumental and a Comitative case. The Instrumental case, which is marked by the suffix *–(X)n/–(I)n* (Erd: 175f), expresses instrumentality, e.g. *ok–**un*** [arrow–INS] 'with an arrow'. It also expresses manner, circumstance, and cause, e.g. *saŋa amran–mak–**ïn** ... öl–ür men* [2SG.DAT love–INF–INS die–AOR 1SG] 'I die from love for you' (Erd: 378[2]). Furthermore, it frequently occurs in temporal adverbials, e.g. *kišïn* 'in winter' (Erd: 379). The Comitative suffix has the shape *–lXgU* in early sources of Old Turkic, while it later seems to have been reinforced by the Instrumental suffix *–(X)n* to give *–lUgUn*. This latter form was probably the source of the Sakha suffix *–LI:n* (Erd: 180). The Old Turkic Comitative is rare, and it expresses the meaning '(together) with', e.g. *teŋri üč teŋri–ler–**lügün** ...kel–ir* [god three god–PL–COM … come–AOR] 'the … god comes with the three gods' (Erd: 379). In one case the Comitative seems to have an instrumental meaning rather than a comitative one: *teŋri–ler sözin–**lügün** yek–ke süŋüš–geli kel–ti* [god–PL word–COM demon–DAT fight–CVB come–PST] 'with the word of the gods he came to fight the demon' (Erd: 379), which could be understood as 'using the words of the gods as a weapon he fought'. In addition to the Comitative case, Old Turkic had a Comitative postposition *bi(r)le* 'with', which appears to have had very similar functions to the Comitative case, e.g. *Xormuzta teŋri beš teŋri **birle** ... kel–ti* [X. god five god with … come–PST] 'the god Ohrmizd came with the fivefold god' (Erd: 394). In later sources, this postposition was occasionally expanded with the Instrumental suffix, giving *bi(r)len* (Erd: 332f). It is from this Old Turkic postposition that the current-day Turkic languages mostly derive their instrumental markers.

In Turkish, both instrumental and comitative functions are expressed by the clitic *=(y)lA* or by the free postposition *ile* (Krn: 227, 228). Thus, we find:

(53a)

kapı–yı	*bir*	*çekiç=**le***	*kır–dı–m*
door–ACC	a	hammer=INS	break–PST–1SG

'I broke the door with a hammer.'

[Krn: 227]

---

[2] Erdal gives this example on both p. 368 (in the section on the functions of the Dative case) and on p. 378 (in the section on the functions of the Instrumental), cf. section 3.2.2.2.

(53b)

(ben)   *konser–e*   *Hasan=**la***   *git–ti–m*
1SG   concert–DAT   H.=INS   go–PST–1SG

'I went to the concert with Hasan.[3]'

[Krn: 228]

This clitic can also function as a coordinating conjunction of two noun phrases or two nominalized clauses (Krn: 114). However, the clitic in its comitative use and in its coordinating use differ syntactically: the Comitative cliticizes to the second constituent and triggers singular verb agreement, while the coordinator cliticizes to the first constituent and triggers plural verb agreement (Krn: 115).

Similarly, the Turkmen postposition *bilen* expresses primarily comitative relations and instruments, as well as being used in adverbials of time, goal and cause (*Grammatika turkmenskogo jazyka* 1970: 401ff; Clk: 407f). It follows nouns in the Nominative case, though personal pronouns followed by *bilen* take Genitive case marking (Clk: 404), e.g. … *on–uŋ **bilen** oyna–š–ma–nï govï gör–erdi* [3SG–GEN with play–REC–NR–ACC good see–HAB.PST] '… and he used to like playing with him…' (Clk: 407), *yašulï pičag–ï **bilen** bir zat–lar kes–išdir–məge otur–dï* [old.man knife–POSS.3SG with one thing–PL cut–ITER–INCP sit–PST.3SG] 'the old man began cutting up some things with his knife' (Clk: 408).

In Kazakh, the postposition *men(en)* is grammaticalizing to an Instrumental case suffix –*men*. Although this does not yet undergo vowel harmony, it has been phonologically reduced, and the initial consonant assimilates to the final consonant of the stem it attaches to, e.g. *kïz–**ben*** [girl–INS] 'with the girl', *at–**pen*** [horse–INS] 'with a horse, on horseback', *at–ï–**men*** [horse–POSS.3SG–INS] 'with his horse'[4]. Along with meanings of instrumentality, means, and manner, e.g. *soyïl–**men** ür–ïp* [club–INS beat?–CVB] 'beating with a club' (Krippes 1996: 16); and also means of transport, e.g. *sol poez–**ben** žür–ip ket–ti* [that train–INS move–CVB go–PST.3SG] 'left by the same train' (SKJ: 177), this suffix also expresses comitative relations (SKJ: 176f), e.g. *Božey–**men** bol–γan tatulïq–tï da ayt–tï* [B.–INS become–PSTPT friendship–ACC also speak–PST.3SG] '(He) also spoke about (his) friendship with Božey.' (SKJ: 177).

---

[3] The Sakha parallel to (53b) with Instrumental case marking on a noun denoting a person has a meaning not of joint action, but of directionality, a perlative nuance. Thus, *min Maša–nan kensierk–ke bar–a–bïn* [1SG M.–INS concert–DAT go–IPF.CVB–PRED.1SG] means 'I go to the concert via Masha's house'. I might possibly take her to the concert with me, but the primary meaning is that of passing by her house on my way to the concert.
[4] Note that Somfai Kara (2002: 33) rejects the analysis of *menen/men* as a case suffix, calling it a 'suffixed postposition'. Unfortunately, he does not adduce any evidence for his analysis.

In Kyrgyz, another Kypchak language, the postposition *menen* expresses a comitative relation and general accompaniment, e.g. *al meni **menen** süylö–š–mök* [3SG 1SG–GEN with speak–REC–NR] 'he has to talk with me', *čay–dï kant **menen** ič–tim* [tea–ACC sugar with drink–PST.1SG] 'I drank my tea with sugar' (Imr: 2294). It is also used to express instruments, e.g. *taš **menen** koy–dum* [stone with hit–PST.1SG] 'I hit (him) with a stone' (Imr: 2295). Similarly, the Bashkir postposition *menen* expresses comitative relations, instruments, manner, means of transport, goal, and other functions (GSBJ: 135f).

In Uzbek, a language of the southeastern Turkic branch, the Locative case expresses some meanings often conveyed by the instrumental case, i.e. a means of an action or a means of transportation, e.g. *men ŭyl–im **bilan** telefon–**da** sŭzla–š–dim* [1SG boy–POSS.1SG with telephone–LOC speak–REC–PST.1SG] 'I spoke with my son on the phone' (Bdr: 128). The postposition *bilan* has comitative, instrumental, temporal and spatial functions, e.g. *χozir kŭča–ga čiq–ib bola–lar **bilan** ŭyna–y–man* [now street–DAT go.out–CVB child–PL with play–IPF.CVB–1SG] 'Now I go out to the street and play with the boys' (Bdr: 267) (also see the previous example).

Khakas, which is a member of the northeastern Turkic branch together with Tuvan and Sakha/Dolgan, differs from the majority of Turkic languages in that it has an Instrumental case marker not derived from the Old Turkic postposition *bi(r)le(n)*. The Instrumental case (marked by the suffix *–nAŋ*, And: 6) expresses the instrument of an action, e.g. *aɣas–tï paltï–**naŋ** o:t–χan–nar* [tree–ACC axe–INS split–PSTPT–PL] 'they split the tree with an axe'. It also has comitative functions, e.g. *uluɣ–lar–**naŋ** toɣïn–ar–ɣa* [big–PL–INS work–AOR–DAT] 'to work with the older ones'. Furthermore, it can mark an overt agent in a passive construction, rather than the Dative case (And: 14).

Similarly to other Turkic languages, Tuvan uses the clitic *=bile* derived from the Old Turkic postposition *bi(r)le(n)* to express instrumentality, accompaniment, and coordination (A/H: 87f), e.g. *kampyuter=**pile** boda–p tur–ar men* [computer=INS think–CVB stand–AOR 1SG] 'I think with a computer', *me:ŋ=**bile** ba:r sen be* [1SG.GEN=INS go.AOR 2SG Q] 'will you go with me?' (A/H: 88); *Goša=**bile** Gena–nï iyildir–zin kïygïr–ïp ekkel* [G=INS G–ACC both–ACC.3SG call–CVB AUX[IMP.2SG]] 'call both Gosha and Gena' (A/H: 21). In Tofa as well, the postposition *bile* expresses both instrumental and comitative relations (Ras: 259).

Like Khakas, Khalaj has a grammaticalized Instrumental case [marked by the suffix *–lA(n)*] expressing means and instruments, accompaniment and coordination. It is furthermore used in manner adverbs (Drf: 92f). The Chuvash Instrumental case has similar functions: expressing instrumentality, comitative relations, and forming temporal adverbials. The case suffix *–vA* is very different from the instrumental

suffix of Old Turkic, Sakha, Khakas, and Khalaj, however (Benzing [1942] 1993: 65f).

The Dolgan Instrumental case has primarily instrumental meaning, e.g. *mu–nu ikki ili:–tinen kus–put* [this–ACC two hand–INS.3SG seize–PSTPT] '(s)he seized it with both hands', although it also serves to form temporal and spatial adverbs (Ubr: 121). Unexpectedly, the Sakha Comitative case (marked by the suffix *–LI:n*) is practically absent in Dolgan (Ubr: 122). The only instances recorded by Ubrjatova are of the Comitative in its use as an inclusive coordinator, with the case marking found on both conjoined nouns, e.g. *oγo–lu:n beye–li:n o:nń–u:r* [child–COM self–COM play–PRS.3SG] 'he and his children, everyone (the whole family) plays (cards)' (Ubr: 122). Instead, accompaniment is expressed by alternative means which are available in Sakha as well: primarily by the postposition *gïtta* 'with' (identical to the Sakha postposition *kïtta*), but also by the suffix *–LA:k* (identical to the Sakha Proprietive suffix *–LA:χ*), or by a verb carrying the Cooperative-Reciprocal suffix *–(I)s* (Ubr: 122f).

3.2.4.3 Accompaniment/comitative relations in Mongolic languages

Proto-Mongolic is reconstructed as having both an Instrumental and a Comitative case. The Instrumental, for which the suffix *–xAr* is reconstructed, expresses the 'means of action ('by what')', while the Comitative (with the reconstructed suffix *–lUxA*) expresses 'the social context of action ('with whom')' (Janhunen 2003d: 15). This distinction is also found in Written Mongolian (an exclusively written lect used throughout the thirteenth to the twentieth centuries), where the Instrumental (marked by the suffixes *–bAr* or *–iyAr*) expresses instruments, manner, as well as accompaniment (PopWM: 76, 152ff). This latter function is called 'instrumentalis sociativus' by Poppe (PopWM: 152, 154), and it "expresses the idea of companionship with someone or connection with something, answering the question 'together with whom?'", e.g. *manu morin tegün–ü morin–iyar belči–müy* [1PL.EXCL.GEN horse.OBL that.OBL–GEN horse–INS graze–PRS] 'our horse grazes together with his horse' (PopWM: 154). The case marked by the suffix *–lUgA*[5] expresses comitative relations, e.g. *eme–lüge* 'with the woman' (PopWM: 76, 155). It can also occur together with the Instrumental case with a meaning very similar to that of the simple Comitative, e.g. *eke–lüge–ber* [mother–COM–INS] 'with

---

[5] This suffix has been compared to the Old Turkic Comitative suffix *–lXgU* (Erd: 180). However, Janhunen (2003d) claims that the Proto-Mongolic Comitative suffix *–lUxA* is derived from a denominal derivational suffix *–lUx* (which formed possessive adjectives) and the early Proto-Mongolic Locative suffix *–A*.

the mother' (PopWM: 155). Furthermore, Written Mongolian has a denominal derivational suffix *–tAy* that expresses possession or some form of association, e.g. *mori–**tay*** [horse–PROP] 'having a horse, horseman', *suryayuli–**tay*** [school–PROP] 'learned, educated [lit. with school(ing)]' (PopWM: 44). This suffix has taken over comitative functions in modern Written Mongolian, probably under influence from spoken lects (PopWM: 76); thus in Classical Written Mongolian we find *vaqe **luqhe*** [elder.brother COM], while in Modern Written Mongolian this has been replaced by *vaqe **tai*** [elder.brother PROP] 'with the elder brother'. In its adverbal use the 'Possessive case' marker (glossed here as Proprietive) is written as a separate particle, while in its adnominal (derivational) use it is written together with the noun, e.g. *vganar **tai*** [woman PROP] 'with a woman' vs. *vganar–**dai*** [woman–PROP] 'having a wife, married' (Janhunen 2003a: 46).

Middle Mongolian (the language spoken in the Mongol Empire from the thirteenth to the fifteenth centuries) continues to have both an Instrumental case (marked by the suffix *–(ni)'Ar*) and the primary Comitative case (marked by the suffix *–lU'A/–lA:*) (Rybatzki 2003: 68f); the later 'Possessive case' (marked by *–tU/–tAy/–tAn*) here still has only derivational functions (Rybatzki 2003: 64f). The Instrumental expresses instruments, materials, cause, manner and other adverbial functions (Rybatzki 2003: 67, 68f), while the Comitative case has both comitative and instrumental functions (Rybatzki 2003: 69).

At a later stage in the development of Mongolic, the Comitative case in *–lUgA* was replaced in some emergent languages by the derivational suffix *–tAy*, which took over its comitative functions, while the original Instrumental retained its function of expressing different circumstances, such as tools, manner, cause, etc. Thus, in Khalkha, Buryat, Khamnigan, and Dagur the suffix *–tAy* (or its respective language-specific variants) functions as a Comitative case (sometimes called Possessive in the Mongolic literature[6]) when used adverbally (54a), but has retained its derivational function in adnominal use (54b, c; K/Ts: 93ff, 97f; Skr: 109; Janhunen 2003c: 90; Tsum: 139), e.g.

(54a)

bi	*a:v–**tay**–ga:*	*xödö:*	*yav–san*
1SG	father–PROP–PREFL	countryside	go–PST

'I went with my father to the countryside.'

[K/Ts: 97]

---

[6] Kullmann & Tserenpil refer to the Comitative case, while the authors in Janhunen (ed. 2003): *The Mongolic Languages* distinguish between the Possessive case (developed from the Written/Middle Mongolian derivational suffix *–tay*) and the Comitative case (retained from the Written/Middle Mongolian Comitative).

(54b)

> *ene oyu:tan   olon   nom–**toy***
> this student   many   book–PROP

'This student has many books.'

(54c)

> *exner–**tey**   xün*
> wife–PROP   man

'a married man (lit. a man with a wife)'

<div align="right">[K/Ts: 98]</div>

The Instrumental in these languages expresses a means of transportation and instruments, e.g. Khalkha *xutg–**a:r** yum ogtol–dog* [knife–INS thing cut.off–HAB] 'you cut things with a knife' (K/Ts: 94) and Dagur *bi: terg–**e:r** ir–sem–by* [1SG car–INS come–PST–1SG] 'I came by car' (Tsum: 138)

Ordos and Oirat are analyzed as having retained the primary Comitative case in addition to acquiring the secondary Possessive, while also retaining the original Instrumental, as did the other Mongolic languages. The Ordos Instrumental expresses instruments or means of action, while the Comitative marks co-subjects. There is no big functional difference between the Ordos Comitative and Possessive cases, except that the Possessive in addition to expressing comitative relations has retained its use as a denominal derivational marker (Georg 2003: 200). In Oirat the Instrumental also marks instruments and material, while the Comitative is the primary case to express comitative relations. It is often used with postpositions such as *adali* 'similar (to)' or *xamtu* 'together (with)' and expresses two agents of equal standing in the clause. The Possessive is rare as an actual adverbal case marker, being used mostly in its function as a derivational suffix (Birtalan 2003: 218).

3.2.4.4 Accompaniment and instrumentality in Tungusic languages

In Evenki, the Instrumental case (marked by the suffix *–(i)t/–di*) expresses primarily an instrument, e.g.

(55)

> *beyumimni   mo:–va   suke–**t**   ive–jere–n*
> hunter   tree–DEF.ACC   axe–INS   hew–PRS–3SG

'The hunter is hewing firewood with an axe.'

<div align="right">[Ned: 154]</div>

Furthermore, Evenki has two Comitative suffixes used only with animate nouns (Ned: 155; B/G: 12; Kon: 70f), *–nun* and *–nAn*, the former being used much more frequently than the latter [in fact, Bulatova & Grenoble (B/G: 8, 12) do not mention *–nAn*]. These are very close in meaning and function; *–nun* marks equal participants of an action, primarily animate ones (Kon: 69), while *–nAn* is used only when the participants are related to each other or are friends (56a) (Ned: 155; Kon: 70, 71). It is very rarely used with possessive suffixes, in contrast to *–nun*, which is frequently attached to possessively marked nouns (Kon: 71). The status of *–nun* as a case-marker is somewhat debated, since it can occasionally combine with another case suffix. Thus, Nedjalkov (Ned: 142) does not list a Comitative case for Evenki, and Konstantinova (Kon: 70) also denies its status as a case suffix, while Bulatova & Grenoble (B/G: 8, 12), do count it among the Evenki cases, arguing that "… in spontaneous speech it rarely co-occurs with any case marking, even when used with a direct object NP." (B/G: 12). As discussed in section 3.1, double or even multiple case marking does occasionally occur cross-linguistically, especially the combination of an adnominal with an adverbal case (Blake 2001: 107), so that in my opinion the occasional co-occurrence of the Evenki Comitative suffix *–nun* with other cases is not enough reason by itself to exclude it from the case system. Although *–nun* most often conjoins two subject NPs (56b), it can occasionally conjoin two nouns in object position (56c) or functioning as spatial adverbials (56d):

(56a)

*beyumimni asi–**nan***   *suru–re–n*
hunter   wife–COM   go.away–NFUT–3SG

'The hunter went away with his wife.'

(56b)

eni   *hunat–**nun**–mi*   *ŋene–jere–n*
mother   daughter–COM–PREFL   go–PRS–3SG

'Mother is going with her daughter.'

(56c)

bi   *kiran–ma*   *munnukan–**nun**–me iče–m*
1SG   crane–DEF.ACC   hare–COM–DEF.ACC see[NFUT]–1SG

'I saw a crane with a hare.'

(56d)

nuŋartïn   *eme–re*   *asatkan–dula ŋinakin–**nun**–dule*
3PL   come–NFUT.3PL   girl–LOC   dog–COM–LOC

'They came to the girl with the dog.'

[Ned: 155]

Four other suffixes combine mainly with inanimate nouns: *–gAli, –či, –tAy* (copied from Mongolic), and *–lAn* (Ned: 155f, 157). These occur very rarely in a comitative function, being mainly used to express circumstance or possession, e.g. *er jahik torga–l–**gali*** [this box cloth–PL–COM] 'this box is with cloth', *beyumimni pektïrevu–**tey*** [hunter gun–PROP] 'a hunter with a gun' (Ned: 157); however, they can occasionally be used with a comitative meaning, e.g.

(57)

eme–kel	*min–dule*	*girki–**geli**–vi*
come–PRXIMP.2SG	1SG–LOC	friend–COM–PREFL

'Come to me with your friend.'

[Ned: 156]

Similar to Evenki, in Èven, the Instrumental case predominantly fulfills 'prototypical' instrumental functions, expressing instruments (58a), means of transport, and circumstance (NovI: 195; Ben: 63), but it can also be used to express a comitative meaning (58b):

(58a)

Iβan	*halka–**č***	*gurgeβči–n*
I.	hammer–INS	work[NFUT]–3SG

'Ivan worked with the/a hammer'

[Ben: 63]

(58b)

tačin	*bi–ji–l–re*		*če:lejur*
thus	be–PROG–INCP–NFUT.3PL		all

Tarčakan	*akïn–**j**–ur*
T.	older.brother–INS–PREFL

'And so they started living all together, Tarchakan and his older brothers.'

[NovI: 196]

There are, furthermore, three suffixes with a comitative function that are very similar both semantically and syntactically. These are *–ńun* (glossed COM1), *–gli* (glossed COM2), and *–čil* (glossed COM3). Of these, *–ńun* has the broadest range of functions, serving to mark equal participants of an action or event in general (59a, b); *–gli* attaches only to personal names, while *–čil* attaches only to animate nouns, predominantly to kinship terms (59c). Thus, *–ńun* can be used interchangeably with the other two suffixes. One difference between *–ńun* and *–gli* is that *–ńun* is used more with transitive verbs, while *–gli* is used more frequently with intransitive verbs. Two different Comitative suffixes can be found in one and the same sentence (59c) (NovI: 178ff).

(59a)

min *ju:–la–β*   *akm–u*
1SG.OBL house–LOC–POSS.1SG   older.brother–POSS.1SG

atïka–ńum–i   *bid–de–n*
woman–COM1–PREFL live–NFUT–3SG

'In my house live my brother with his wife.'

[NovI: 178]

(59b)

min *kini:γe–ruk–le–β* *kini:γe–l* *titra:ji–l–ńun*
1SG.OBL book–DER–LOC–POSS.1SG book–PL notepad–PL–COM1

desči–r
lie–NFUT.3PL

'In my book bag lie books with notebooks.'

[NovI: 179]

(59c)

bi *eken–čil*   *Mariya–γli* *bi–se–m*
1SG older.sister–COM3 M.–COM2 be–NFUT–1SG

'I live with my older sister Maria.'

[NovI: 181]

Benzing (Ben: 64) mentions a fourth Comitative suffix *–(y)e/–ńe*, which he says is used mainly with personal names and pronouns, e.g. *kuŋa–l–a hupkučimŋe em–re* [child–PL–COM4 teacher come–NFUT.3PL] 'the teacher came with the children' (Ben: 64). This suffix is not mentioned by Novikova (NovI: 178ff) in her discussion of the Èven Comitative case. It looks very similar to the Yukaghir Comitative suffix *–ńe* (see section 3.2.4.5 below), and might therefore possibly have been copied from Yukaghir into Èven.

   In Negidal, the Instrumental case suffix is *–ji*, as in Amur Tungusic, not *–t/–di* as in Evenki, and the range of its meanings is comparable to that in Èven rather than to Evenki. Furthermore, the Negidal Instrumental can have a comitative meaning (Xasanova & Pevnov 2003: 245). There are two additional suffixes with a comitative meaning in Negidal (found primarily in the Lower dialect): one is the suffix *–čil* also found in Èven, e.g. *emen beye asi–čil jav–ji ŋene–ča–l* [one person woman–COM boat?–INS go–PST–PL] 'one man crossed over by boat with a woman' (Xasanova & Pevnov 2003: 246). The other Negidal Comitative marker is the complex suffix *–l.γeli.ji*. The final element of this is clearly the Instrumental suffix; the first element is probably a plural suffix, and Xasanova & Pevnov compare the middle part to the Èven Comitative suffix *–gli/–γli* (Xasanova & Pevnov 2003: 245f). However, in my opinion this can also be compared to the Evenki suffix *–gali*

which can have a comitative meaning, although it mainly expresses circumstance, as discussed above.

In Udihe, an Amur Tungusic language, the Instrumental case has a wide range of meanings, covering the expression of instrument (60a), manner, means, circumstance and material, as well as commonly having a comitative function (N/T: 126). In addition to the Instrumental case, comitative meanings in Udihe can be expressed with any of four Comitative postpositions, one of which is restricted to inanimate nouns, referring to an entire object together with its parts (N/T: 412f, 496f, 499, 570). One of the postpositions with a comitative meaning, *mulə* 'with', is restricted in its functions in a similar way as the Evenki Comitative suffix *–nAn* and the Èven suffix *–čil*: it is used only with animate nouns and presupposes an 'inalienable' relationship between them, joining nouns referring to relatives or friends (N/T: 412). The Instrumental case in its comitative function does not appear to differ in use from the Comitative postpositions: both the case-marked noun phrase and the postpositional phrase can serve as nominal modifiers (60b, c) (in which case the verb shows plural agreement, although the subject is singular) or as comitative adjuncts (60d, e) (with singular verb agreement).

(60a)

loxo–zi	*mafa–wa*	*wa:–mu*		*g'ə:*
saber–INS	bear–ACC	kill–IMPRS.INF		bad

'It is a sin to kill a bear with a saber…'

[N/T: 567]

(60b)

bi	*anda*	***mulə***	*ŋeixi*	*ŋənə–u*
1SG	friend	with	to.the.river	go–1PL.EXCL

'I go to the river with my friend.'

[N/T: 496]

(60c)

kuti	*kəige–zi*	*anana*	*aya*	*bi–si–ti*
tiger	cat–INS	long.ago	good	be–PST–3PL

'Tigers and cats were friends a long time ago.'

[N/T: 500]

(60d)

bi	*anda*	***mulə***	*ŋənə:–mi*
1SG	friend	with	go.PST–1SG

'I left with my friend.'

[N/T: 413]

(60e)

> ya: mənə sita–**zi:**   xokto cu–li–ni   əmə–ini
> cow REFL child–INS.REFL   road   along–PROL–3SG come–3SG

'The cow is coming down the road with her calf.'

<div align="right">[N/T: 570]</div>

In Nanay the Instrumental case covers both prototypical instrumental functions of expressing an instrument, material, and circumstance as well as comitative functions, e.g. *boatoŋo miočan–**di** mioča–la–xa–ni* [hunter gun–INS gun–VR–PST–3SG] 'the hunter shot with the gun', *Lete brigadir–**di** pravlenie–či pulsi–xe–ni* [L. brigade.leader–INS administration–ALL go–PST–3SG] 'Lete went to the administration with the brigade leader' (AvrI: 180; AvrIII: 160ff). Furthermore, there exists a derivative suffix *–molia/–mulie* which expresses joint actions; this is restricted in use to kinship terms and similar nouns denoting a close relationship between humans. This suffix appears to be cognate to the Udihe postposition *mulə* 'with' described above, both in its form and function. However, it differs from the latter in that it does not mark the dependent noun of a complex subject NP, but derives nouns with a meaning of 'shared action of a pair of individuals' (cf. section 3.4.5), e.g. *neumulie* (< *neu* 'younger brother/sister') 'older brother with younger brother, older brother with younger sister, older sister with younger brother, older sister with younger sister' (Onenko 1980: 301), cf. *neu–**mulie** boato–nda–mari ene–xe–či* [younger.sibling–DER hunt–PURP–SIM.CVB.PL go–PST–3PL] 'The brothers (older and younger) went hunting' (AvrI: 113). Note that the verb agrees in number with the implicit (plural) subject, not the syntactic (singular) subject. However, it should be pointed out that Boldyrev (1976: 103) gives two Nanay examples with this suffix functioning as a 'normal' comitative, i.e. marking the dependent noun within a complex subject, e.g. *botamdi ekten–gu–**mulie** oni–la ene–xe[7]–či* [fisherman woman–??–DER river–LOC go–PST–3PL] '(The) fisherman and the woman floated downriver' (Boldyrev 1976: 103). What makes it difficult to assess these examples is the suffix *–go/–gu* preceding the Comitative suffix in both cases; this suffix is a verbal suffix (either iterative aspect or a variant of the purposive converb), and it is also the marker of the Designative case (Avrorin 1961: 281). Neither of these suffixal meanings makes sense in the construction above (which is why I glossed the suffix with a question mark); it is thus not clear whether this mysterious suffix adds some meaning to the construction that is not discussed by Avrorin (AvrI: 113).

---

[7] Boldyrev writes this as *ene–xo–či*, but that must be a typo, since the past tense marker should be *–xe* in this case, as can be seen in the previous example.

The Oroč Instrumental case has functions very similar to those of the Nanay Instrumental, marking both instruments, material, and circumstances, as well as expressing a comitative relation (A/B: 118ff), e.g. *biti ama–ji eńe–ji gese bi–je:–pi* [1PL father–INS mother–INS together be–FUT–1PL] 'we will live together with father and mother' (A/B: 120). The suffix *–muna* derives nouns with a meaning of 'people connected by kinship ties', e.g. *aki–muna eme–xe–ti* [older.brother–DER come–PST–3PL] 'the younger sibling came together with the older brother' (A/B: 72); this appears to be the same as the Nanay suffix *–molia/–mulie*.

Orok appears to be the only Amur Tungusic language that makes a distinction between an Instrumental and a Comitative case. The Instrumental case (marked by the general Tungusic instrumental suffix *–zi*; cf. Appendix 4) expresses instruments, means of transport and material, e.g. *bi mo:–βo pupun–zi pupu–la–xa–mbi* [1SG tree–ACC saw–INS saw–VR–PST–1SG] 'I sawed the tree with a saw' (Pet: 50), while the Comitative case marks the dependent noun of complex subject NPs 'playing an active role in the event', e.g. *bi gi:da uylta–ndo ŋene–xe–mbi* [1SG one Orok–COM go–PST–1SG] 'I went with one Orok' (Pet: 50). Even though the Orok Comitative suffix *–ndo/–ndu* differs formally from the Evenki and Ėven suffix *–nun/–ńun*, Petrova (Pet: 51) appears to think they are cognate. The verb can show either singular or plural subject agreement. In cases of plural verb agreement the main actant is often marked by a 'Formal Proprietive' suffix[8] ("… название главного действующего лица часто имеет формальную принадлежность *–ja/–jo…*"; Pet: 51), e.g. *Omboloto–yo sama–ndu zimbe ineŋi gobdo–mori…* [O.–PROP shaman–COM four? day hunt–SIM.CVB.PL] 'Omboloto and the shaman, hunting for four days….' (Pet: 51). The minor participant in Comitative NPs can be marked with the suffix *–muna/–mune*, which formally appears to be cognate to the suffix found in Nanay and Oroč as well as the Udihe postposition *mulə* 'with'. However, in Orok this suffix does not appear to be restricted to kinship terms, e.g. *tari nari apkatči–či ča boyo–ŋu–muna* [that person lie–PRS.3PL? that bear–??–DER] 'that person lay down to sleep with that bear' (Pet: 51). As can be seen, this suffix attaches to the same mysterious suffix *–go/–gu* (assimilated to an oblique *–n* in the example) as found in the Nanay examples given by Boldyrev (1976: 103).

Manchu lacks both a separate instrumental as well as a separate comitative case. Instrumental functions are fulfilled by the Genitive case, e.g. *ere suhe–i sači–mbi* [this axe–GEN chop–IPFV] '(I) shall chop with this axe' (Gor: 181), while comitative relations are expressed with the postpositions *emgi* and *sasa* 'with' (Gor: 349), e.g. *bi šin–i emgi gene–mbi* [1SG 2SG–GEN with go–IPFV] 'I shall go with you' (Gor: 350). For a summary of the instrumental and comitative cases in Siberian languages cf. Table 3.10.

---

[8] This suffix also occurs in Comitative constructions in Evenki, cf. section 3.2.3.7.

3.2.4.5 Instrumental and comitative cases in Siberian languages

Yukaghir makes a distinction between an Instrumental case that expresses instruments, means and manner (61a) (Mas: 104f) and a Comitative case which can function as a marker of coordination and which expresses joint actions (Mas: 101f). There are two different Comitative suffixes, *–ńe* and *–ńit/–ńut*, neither of which is restricted in use to humans or animates. The suffix *–ńe* can conjoin only two participants of an event (61b), either subject or object NPs (Mas: 313), while *–ńit/ –ńut* can conjoin more than two participants and is used preferably to mark objects (61c) (Mas: 103, 315f).

(61a)

> *ta:t* *ńumuji:–le* *čine–y–m*
> CA axe–INS chop–PFV–TR.3SG

'Then he chopped it with an axe.'

[Mas:104]

(61b)

> *met* *irk–in* *paype uö–ńe* *yuode–ye*
> 1SG one–ATTR woman child–COM play–INTR.1PL

'I played with a girl.'

[Mas: 314]

(61c)

> *uör–pe–ńit* *paype–ńit* *čumu ley–l'el–u–m* *pat–telle*
> child–PL–COM woman–COM all eat–INFR–E–TR.3SG cook–PF.CVB.SS

'The children and the women, he appears to have eaten all of them, having cooked them.[9]'

[Mas: 316]

Proto-Ob-Ugrian is believed to have made a distinction between an Instrumental and a Comitative case, but the Comitative case has been lost in most current-day dialects of both Khanty and Mansi (Honti 1998: 345). The Sosva dialect of Mansi has an Instrumental case which expresses both instruments, means and manner, e.g. *ńa:wram a:kań–əl yonγi* [child doll–INS play.3SG] 'the child plays with the doll' (Rse: 26) as well as comitative relations (Rmb: 51ff, 54). In its comitative function the Instrumental case suffix can mark both subjects and objects; it attaches to a possessively marked form of the noun, which distinguishes it from the instrumental function, e.g. *an' o:yka e:kwa–te a:γi–te:n–təl sas ra:t–saγ(e)* [that? man wife–POSS.3SG girl–POSS.DU:SG–INS ?? beat?–PST.DU.OBJ.3SG] 'that man beat

---

[9] Note the structural parallel with the Sakha Comitative used as an inclusive coordination marker (47a–c).

his wife together with his daughter' (Rmb: 54). Note that Riese (Rse: 26) claims that the Mansi Instrumental does not generally express comitative relations, and that the postposition *yot* 'with' is used instead, e.g. *piγ aːγi **yot** at yonγi* [boy girl with NEG play.3SG] 'the boy is not playing with the girl' (Rse: 26). Murphy (1968: 43, 86f) mentions both the Instrumental case and the postposition *yot* 'with' for expressing accompaniment in Sosva Mansi; in addition, he discusses a so-called 'Sociative' marker which can fulfill the same function, e.g. *akw oːyka aːγi–ńś yuw śaltə–səγ* [old? man girl–SOC into enter?–PST.3DU] 'an old man came in with his daughter' (Murphy 1968: 87); however, this is used less frequently than the other two means of expressing comitative relations.

Khanty has a number of widely divergent dialects (which on linguistic grounds could be considered separate languages; Comrie 1981: 106). The northern dialects, which form the basis of the literary language, have only three cases, while the eastern dialects have 10-11 (Nik: 4; Abondolo 1998a: 361). Since northern Khanty has only three different cases, it distinguishes neither the instrumental nor the comitative. Instrumentality is expressed by the Locative case (Nik: 13); it is not mentioned in the grammar sketch how comitative relations (i.e. the joint action of two nouns of equal standing) are expressed. In the eastern Vakh dialect both instrumentality and accompaniment are expressed by one case, called the 'Instrumental-Comitative' by Tereškin. The comitative function of this case is restricted to animate nouns, most frequently those denoting humans (Trš: 51f). A further case is called 'Instrumental-Objective' by Tereškin (Trš: 52) and 'Instructive' by Abondolo (1998a: 378); however, this case does not cover any of the functions usually expressed by instrumental or comitative cases, notwithstanding the term chosen for it by Tereškin. It expresses a benefactive relationship when used with ditransitive verbs, the case-marked object being the entity that benefits the patient.

The case system reconstructed for Proto-Samoyedic lacks both an instrumental and a comitative case (Janhunen 1998a: 469). Selkup used to distinguish between an Instrumental and a Comitative case, but that distinction has been lost (MSJ: 254ff). Nowadays, there is one case that jointly expresses manner, means, and accompaniment or joint action (MSJ: 266f, 269f), e.g. *wəči–p paŋi–se maːt–iŋiti* [meat–ACC knife–INS cut–3SG.OBJ] 'he cut the meat with a knife', *man ima–ni–se il–ak* [1SG woman–POSS.1SG–INS live–SBJ.1SG] 'I life with my wife' (OSJ: 175).

Nganasan lacks a dedicated instrumental case, with the Locative expressing both spatial and temporal location, as well as instrumentality and means; compare for example *basuʔsi–ʔ muŋku–**tini** basu–tu–ʔ* [hunter–PL tree?–LOC/INS hunt–PRS–

3PL] 'hunters hunt in the forest' and *ta:–**tini**–tuŋ tuy–süo* [reindeer–LOC/INS–POSS.3PL arrive?–PST[10]] '(they) arrived with (on) their reindeer' (Ter: 87). Comitative relations in Nganasan are primarily expressed by a postposition *nanu* 'with', which follows the noun in the unmarked case, e.g. *kuodümu nï **nanu** ta:–tini tuy–süo–gəy* [man woman with reindeer–LOC/INS arrive?–PST–3DU] 'a/the man and a/the woman arrived by reindeer' (Ter: 111). Furthermore, comitative relations with a reciprocal nuance can be expressed by the suffix *–na* (called a Comitative case by Helimski 1998b: 496, 499), e.g. *desï–mə deruʔtuomə basuʔsi–**na** buo–tu* [father–POSS.1SG unknown? hunter–COM talk?–PRS.3SG] 'my father is talking with an unknown hunter' (Ter: 111). Lastly, accompaniment and association can also be expressed by the suffix *–səptə*, e.g. *kuodümu nüo ban–**səptə** xuaŋgü–ʔö* [man child dog–SOC run.away?–PRS.3SG] 'the boy ran away with the dog' (Ter: 111).

Similar to Selkup, the neighbouring (but unrelated) language Ket expresses both comitative relations (joint actions) and instrumentality by a single case form (Wer 1997: 115f).

Nivkh has an Instrumental case that expresses an instrument or manner of an action as well as material, e.g. *ń–ïtïk qaχ–**kir** qʰotr kʰu–d'* [1SG–father spear–INS bear kill–FIN] 'my father killed the bear with a spear' (Grz: 21). There is also a Comitative suffix, which however is not classified as a case marker, since it can combine with further case suffixes. This suffix attaches to both conjoined nouns when it is stressed that both are equal participants in an action, e.g. *pʰ–at'ik–**xe** pʰ–nanak–**xe** hum–yaʁan* [REFL–younger.brother–COM REFL–older.sister–COM live–FIN] 'her younger brother lived with his elder sister' (Grz: 33). When the two nouns are viewed as a single agent, only the second noun is marked, e.g. *ń–ïkïn pʰ–oγla–**ge** uγrït ńe-rχ pʰrï–d'–γu* [1SG–older.brother REFL–child–COM together 1SG–ALL come–FIN–PL] 'my elder brother came to me with his child' (PanI: 166). The Comitative suffix can attach to subjects (in this case, it is restricted in use to animate, and predominantly human agents), direct and indirect objects – in these latter cases there is no restriction on the animacy of the case-marked nouns (PanI: 166f).

Siberian Inupik Eskimo has an Instrumental case (marked by the suffix *–mïŋ*) which expresses direct objects, instruments and ablative-like locative functions (Men: 51f). Furthermore, it has a Comitative suffix *–tuma* which expresses circumstance, e.g. 'they sent away the boy with his book', 'the woman gave me meat with a sack' (Men: 63). Since this suffix is always followed by the Ergative suffix in the possessive declension, Menovščikov (Men: 63) does not consider it a

---

[10] Judging from the plural possessive marking on the case-marked noun, there is a glottal stop, which is the 3PL person agreement marker of the subjective (intransitive) verbal inflection, missing at the end of the verb *tuy–süo–(?)*.

case suffix. True comitative relations, i.e. the joint action of two nouns of equal standing, are expressed by the particle –*lu*, which can attach to both nouns participating in the action, or only to the second one (Men: 187). The difference in marking appears to be whether the agents are perceived to be equal participants, or whether one is perceived as being more important than the other. However, this is difficult to judge based on the single example given for each situation by Menovščikov. Central Siberian Yupik and West Greenlandic both have an additional affix –*kkut(e)* that expresses an association [glossed by de Reuse (1994: 33) as 'have.N.as.associate.or.partner', and translated as 'and fellows', 'and company/ family' by Fortescue 1984: 129, 215)]; whether this is also found in Siberian Inupik Eskimo is not clear.

Itelmen has an Instrumental case which is restricted in use to inanimate referents and cannot combine with personal pronouns or with animate nouns; it expresses instruments, means, and manner (G/V: 80). It furthermore has two Comitative cases that express accompaniment. The first Comitative joins independent noun phrases; it is not restricted in use to animate nouns and sometimes its functions are similar to those of the Instrumental case. The second Comitative mainly joins noun phrases that show a part/whole relationship, though this is not a strict rule (G/V: 83f, 85).

In Chukchi, the Ergative and Instrumental case are marked by the same suffix, but are distinguished by their syntactic functions (Dnn: 112). The Instrumental is mostly used to express instruments, mainly on inanimates (Dnn: 113f). Furthermore, like Itelmen, Chukchi has two cases marking comitative relations and accompaniment: the Comitative case joins equally ranked nominals, while the Associative case marks accompaniment by something that is part of the head or a typical possession of the head (Dnn: 116f). Table 3.10 summarizes the distribution of an instrumental and a comitative case in Siberian languages.

Table 3.10: Overview over the distinction between instrumental and comitative in northern Eurasian languages

Language (group)	Instrumental	Comitative
Old Turkic	–*(X)n/–(I)n*	–*l(X)gU/–lUgUn*
present-day Turkic	*birlen/bile/ile/menen*	
Sakha	–*(I)nAn*	–*LI:n/–InA:n/–InI:n*[*]
Written Mongolian/ Middle Mongolian	–*iyAr/–(ni)'Ar*	–*lUgA/–lU'A* (–*tAy*)
present-day Mongolic	–*A:r*	–*tAy* (–*lA:*)

Table 3.10: Overview over the distinction between instrumental and comitative in northern Eurasian languages, cont.

Language (group)	Instrumental	Comitative
Evenki	*–(i)t/–di*	*–nun*
		–nAn
		(–gAli)
Ėven	*–č*	*–ńun*
		–gli
		–čil
Negidal	*–ji* (so-called Instrumental case)	
		–čil
		–lγaliji
Udihe	*–zi* (so-called Instrumental case)	
		mulə (postposition)
Nanay	*–di* (so-called Instrumental case)	
		(–mulie)
Oroč	*–ji* (so-called Instrumental case)	
		(–mune)
Orok	*–zi*	*–ndo*
		(–mune)
Language (group)	**Instrumental**	**Comitative**
Yukaghir	*–(l)e*	*–n'e*
		–ńit/–ńut
Mansi	*–(ə)l/–təl* (so-called Instrumental case)	
		yot (postposition)
		–ńś (Sociative)
Khanty – northern	neither, only 3 case forms	
Khanty – eastern	*–ne/–na*	
Nganasan	[*–(n)tənu/–(n)tini*]	*nanu* (postposition)
	(=Locative +	*–na*
	Instrumental)	*–səptə*
present-day Selkup	*–SA*	
Ket	*–as'*	
Nivkh	*–Kir/–Kiř/–Kis*	*–Ke/–Kin*
Siberian Inupik Eskimo	*–mïŋ*	*–lu*
Itel'men	*–ḷ*	*k–...–ḷ/x–...–ḷ*
		k–...–çom/x–...–çom
Chukchi	*–e^{-VH}*	*γe–...e^{-VH}*
		γa–...–ma^{+VH}

[*]The suffixes from the possessive declension are included here to show their similarity to the Evenki suffixes.

$^{-VH}$/$^{+VH}$: these superscripts indicate absence or presence of vowel harmony prosody.

3.2.4.6 The Sakha distinction between an Instrumental and Comitative case in the light of Eurasian case-marking patterns

From the data presented above, it is quite clear that Sakha differs strongly from the other Turkic languages in having retained both the Instrumental and the Comitative case to distinguish between an action performed using an inanimate entity as an instrument or means and a joint action performed by two (primarily animate, human) referents of equal standing. Although there have been suggestions that the Sakha Comitative suffix *–LI:n* is an innovation (Btl: 259; Radloff 1908: 32) or a copy from Evenki (Ubr: 123f), there cannot be any doubt in my opinion that it is in fact a retention of the Old Turkic suffix *–lUgUn*, as suggested by Erdal (Erd: 180), since intervocalic velars commonly get lost in Sakha, resulting in a lengthening of the vowel. However, even though the case form and function are inherited, the fact that Sakha retained the distinction between Instrumental and Comitative, when all other Turkic languages lost it, can most plausibly be explained by contact influence. Tracing the source of this influence, however, is once again not as easy a task as it might seem.

As was discussed above, and as can be seen in Table 3.10, the presence of both an instrumental and a comitative case is widespread in northeastern Eurasia, especially among eastern Siberian languages. As a matter of fact, the distinction is not restricted to Siberia, but is widespread worldwide, being found in 66% of a cross-linguistic sample; it is only rare in the languages of Europe (Stolz 1996; Stolz et al. 2005). Given the fact that a distinction between two cases to mark instruments, means and circumstances on the one hand and accompaniment or joint action on the other can be viewed as the default value, especially in northeastern Siberia, one might argue that Sakha followed a 'natural inclination' in retaining the distinction found in Old Turkic. However, not a single other Turkic language (with the exception of Dolgan, which, however, is closely related to Sakha) makes a formal distinction between these meanings, indicating that the 'natural inclination' of Turkic languages was to give up the distinction. Thus, an explanation based on contact influence is more plausible than one based on language-internal development.

Ubrjatova (1966: 49f; Ubr: 123) suggests that the Sakha Comitative case was copied from Evenki, deriving the suffix of the simple declension *–LI:n* from the Evenki suffix *–nun* (with dissimilation of the initial *–n–* in Sakha), and deriving the low-vowel variant of the possessive declension from the Evenki suffix *–nAn*. Since Ubrjatova claims that the variant with the low vowel in Sakha is used with kinship terms, she sees a close parallel to the use of the Evenki suffix *–nAn*. However, as mentioned above, I do not think that there are any good arguments for postulating a

copied origin of the Sakha Comitative suffix –*LI:n*, which can easily and plausibly be derived from the Old Turkic Comitative suffix –*lUgUn* (cf. Erd: 180).

We can thus assume that the contact influence was of a structural nature, leading to the retention of the case distinction, but not involving the copying of forms. Initially, Mongolic influence may have played a role in the retention of this case in (pre-)Sakha, since Middle Mongolian distinguished an Instrumental from a Comitative case. As to the Sakha possessive declension, the suffix variant with a high vowel can easily be derived from the suffix of the simple declension, since the case suffixes of the possessive declension are characterized by an initial –*n*– (cf. Appendix 3). This is especially clear in the Partitive, Comitative and Comparative cases, where we find the simple suffix –*TA*, –*LI:n*, and –*TA:γAr*, respectively, while the case marker in the possessive declension can be analyzed as –*nA*, –*nI:n*, and –*nA:γAr*, respectively. However, the Sakha Comitative case differs from the other cases in that it has two suffix variants in the possessive declension: –*nI:n* and –*nA:n*. While the variant with the high vowel can be explained by internal derivation, the situation is not so clear for the variant with the low vowel. This could have arisen by contamination of the possessive Comitative suffix –*nI:n* with the Instrumental suffix –*nAn*; instrumental and comitative meanings are semantically quite close, as demonstrated by the reasonably large number of languages in which these cases are marked by syncretic forms. However, the mere fact that only this case has two suffixal variants in the possessive declension is odd and calls for an explanation other than internal development. In this context, the similarity of the Evenki Comitative suffixes –*nun* and –*nAn* to the Sakha possessive variants –*nI:n* and –*nA:n* is rather striking. Even though Ubrjatova's claim that the low-vowel variant in Sakha marks kinship terms cannot be substantiated with modern data (cf. section 3.2.4.1.2), the large similarity in form still makes Evenki influence in the development of this suffix in Sakha quite plausible. Thus, to summarize, we might postulate two layers of influence in the retention of the Sakha Comitative case: first, Mongolic influence facilitated the retention of the case distinction, and this was later reinforced by Evenki influence, which furthermore led to the development of the low-vowel variant of the suffix in the possessive declension.

However, there is a factor that complicates the nice scenario sketched out above. This is the fact (described in detail in section 3.2.4.4) that Evenki stands alone amongst the Tungusic languages in having a Comitative suffix –*nAn*. Not even the closely-related languages Ėven and Negidal have this. One might therefore argue that it was rather Sakha influence that led to the development of this suffix in Evenki, if the suffix was not copied outright from Sakha. On the other hand, it appears to be a characteristic of the Tungusic languages, especially of the Northern Tungusic branch, to have a form marking specifically kinship terms as participants

of complex subjects. Thus, as described in section 3.2.4.4, the Evenki Comitative suffix *–nAn* fulfills this role, as does the Èven Comitative suffix *–čil*, the Udihe Comitative postposition *mulə* 'with', and the Nanay and Oroč derivational suffixes *–mulie* and *–muna*, respectively. This indicates that the influence of a notion of inalienable relationship on morphosyntax is common to the Tungusic languages (as also evinced by the formal distinction made by these languages between alienable and inalienable possession), making it more plausible that the Evenki suffix *–nAn* is an independent innovation in this language, rather than a copy from Sakha. This is in good accordance with Heath's suggestion (1978: 75) that the more archaic morpheme may have developed specialized functions, while in the process of copying the functions are simplified. On the other hand, one could also argue that Evenki had the means of expressing a joint action with an inalienably connected person (i.e. kin or close friends), such as the Comitative suffix *–čil* found in Èven, and that this was simply replaced later by the suffix *–nAn* copied from Sakha.

However, since the suffix *–nAn* in Evenki fills a functional slot also found in other Tungusic languages, while it is very unusual for Sakha to have two variants of a suffix in the possessive declension, it might be somewhat more plausible to argue for Evenki influence in Sakha rather than vice versa. On the other hand, if the retention of the Old Turkic Comitative in Sakha should truly be due to Evenki influence, then it is rather surprising that Dolgan does not have a suffixally-marked comitative case, as already noted by Ubrjatova (1966: 50; Ubr: 123). Since Dolgan is assumed to have been more heavily influenced by Evenki than Sakha has, one would expect a feature that was retained in the latter due to contact influence with Evenki to be present in Dolgan as well. Ubrjatova (1966: 50) explains the lack of this case in Dolgan by its having been copied into some Sakha dialects at a late stage, from where it spread to other dialects, but not to Dolgan. It is quite clear, of course, that given the fact that a copied source of the Sakha Comitative cannot be accepted, this argument cannot explain the lack of this case in Dolgan. On the other hand, it is possible that the functions of the comitative case have been taken over in Dolgan by the postposition *gïtta* 'with' for purely language-internal reasons – a tendency which is discernible in Sakha, too, where the postposition *kïtta* 'with' appears to be used as frequently as the actual Comitative case suffix.

Thus, to summarize, it appears plausible that the retention of the Old Turkic Comitative case in Sakha was facilitated by contact influence. Initially, this may have been Mongolic influence, but after the migration of the Sakha to the middle reaches of the Lena river Evenki influence arguably may have played a role, too, leading to the development of the low-vowel suffix variant of the possessive declension. In Dolgan the case may have been replaced by alternative means of expressing joint actions for language-internal reasons.

3.2.5 The origins of the Sakha Comparative case

As can be seen from Table 3.2 (section 3.2), there is yet another case in which Sakha differs from the other Turkic languages. This is the Comparative case, which is found only in Sakha. In the following, I shall give a brief overview over the form and function of this case and examine whether it may have originated as the result of contact influence.

3.2.5.1 The Sakha Comparative case

The Comparative case in Sakha is marked by the suffix *–TA:γAr* in the simple declension, and *–(I)nA:γAr* in the possessive declension (cf. Appendix 3). It marks the standard of comparison in comparative constructions (62a, b), while the adjective, for which no comparative or superlative forms exist, remains invariant. An exception is made, however, for comparative constructions regarding age: here, the standard of comparison is marked by the Ablative case, while the relative age ('younger', 'older') is expressed by the words *bïra:t* 'younger.brother' or *balïs* 'younger.sister' and *aγa* 'father' (62c). Nowadays, however, *bïra:t* and *balïs* are being replaced more and more by *kïra* 'small'.

(62a)

anï	*ikki*	*bi:r*	*oburgu*	*emie*	*hïtar*	*on**no:γor***	*kïra*
anï	ikki	bi:r	oburgu	emie	sït–Ar	on–TA:γAr	kïra
now	two	one	fairly.big	again	lie–PRSPT	that.OBL–COMP	small

'Now one or two fairly big ones are still lying there, smaller than that one.'

[BesP133]

(62b)

ontuŋ	*χannïk*	*da*	*χopruon*	*hap**ta:γar***	*böγö*	*buolar*
ontu–(I)ŋ	χannïk	da	kapron	sap–TA:γAr	böγö	buol–Ar
that–POSS.2SG	of.what.sort	PTL	kapron	thread–COMP	strong	be–PRSPT

'That is stronger than any kapron thread.'

[MatX243]

(62c)

ulaχan	*eji:ybit*		*mi:gitten*	*uon*	*aγïs*	*du:*	*hette*	*du:*
ulaχan	eji:y–BIt		mi:gitten	uon	aγïs	du:	hette	du:
big	older.sister–POSS.1PL		1SG.ABL	ten	eight	DISJ	seven	DISJ

hïl	*aγa*	*ete*
hïl	aγa	e–TA
year	father	AUX–PST.3SG

'Our oldest sister was seventeen or eighteen years older than me.'   [BesP11]

The Comparative case suffix *–TA:γAr* is clearly complex. Some researchers have suggested that the Comparative suffix derives from the Old Turkic Locative case with the suffix *–dA* (which had an ablative meaning as well, cf. section 3.2.3.3) and the Old Turkic Allative case with the suffix *–GArI* (Xaritonov 1947: 113; GSJa: 138; Teniševev 2002: 666). This receives further weight by the fact that Old Turkic always used the Locative case, not the Ablative case, in comparative constructions e.g. *anta–**da** taqi*[1] *yėg–rek* [that–LOC even good–COMP] 'even better than that' (Erd: 372). However, while it makes sense that an element with an ablatival meaning would grammaticalize to a comparative case (cf. section 3.2.5.2), the meanings of the Locative/Ablative and Allative case are complete opposites, making this combination rather unlikely. On the other hand, since one of the strategies used to mark the standard of comparison is to use a marker expressing direction (Stassen 2005: 490; Heine 1997: 116), one might argue that the original Old Turkic Locative/Ablative suffix was reinforced by the Allative suffix. However, while the 'source schema' (i.e. the use of a marker expressing movement away) for comparative constructions is very frequent in Asia, the 'goal schema' (using a marker expressing movement towards) is very rare to non-existent in this region (Heine 1997: 128), casting serious doubt on the supposed derivation of the Sakha Comparative suffix from an Allative suffix.

Ubrjatova (1976: 121f) suggests that the Comparative case suffix is derived from a 'Common Turkic Comparative case' suffix *–tay/–taγ* and the Turkic adjectival comparative suffix *–rAK*, followed by metathesis of the morpheme *–rAK* and subsequent lenition of the *–k–*. What she calls the 'comparative case' suffix is presumably the Old Turkic Similative or Equative postposition *teg* 'like' that grammaticalized to a pronominal case suffix *–tAg*, as found in constructions such as *an–tag* [that.OBL–SIML] 'like that', or *sizin–teg* [2SG.OBL–SIML] 'like you' (Erd: 336). Reflexives of this are still found in some modern Turkic languages, such as the Similative suffix *–TAγ* in Baskhir, e.g. *imen–dey* [oak–SIML] 'like an oak' (GSBJ: 175), and the postpositions *tek(i)* and *tektes* 'like, as' in Azerbaijanian and Kazakh, respectively (Schönig 1998: 257; Isengalieva 1957: 98). The suffix *–rAK* marks the comparative degree of adjectives in the Turkic languages, as described in section 3.2.5.2 below. Ubrjatova's proposal clearly makes the most sense from a semantic point of view, although it does seem somewhat far-fetched formally (I would rather expect the combination of *–tAg* and *–rAK* to grammaticalize to *–TA:rAχ* than *–TA:γAr*). Thus, the origins of the Sakha Comparative case suffix may require further study.

---

[1] Erdal (Erd: 150, 372) writes this with a *–k–*, since *q* is only an allophone of *k* (Erdal pers. comm.); the *Drevnetjurkskij Slovar'* (1969: 536) writes it with *–q–*. In general, in Turkic languages the palatal velar *–k–* does not combine with back vowels (Johanson 1998a: 31).

3.2.5.2 The expression of comparison in Turkic, Tungusic, Mongolic, and other Siberian languages

As mentioned above, most of the Turkic languages have a suffix –*rAK* to mark the comparative degree of adjectives, e.g. Old Turkic *yėg–rek* [good–COMP] 'better' (Erd: 150, 372), Azerbaijanian *böyük–rek* [big–COMP] 'bigger' (Schönig 1998: 251), or Uzbek *širin–råq* [sweet–COMP] 'sweeter' (Boeschoeten 1998: 361). The standard of comparison is expressed by the Ablative case, e.g. Uzbek *nåq ålma–**dan** širin–råq* [pear apple–ABL sweet–COMP] 'the pear is sweeter than the apple' (Boeschoeten 1998: 361). Turkish, however, has lost the suffix –*rAK*; here, an adjective in the comparative degree is modified by the adverb *daha* 'more'. As in the other Turkic languages, however, the standard of comparison is expressed by the Ablative case-marked noun (G/K: 198). Since the Ablative (or, in Old Turkic, the Locative) case-marked noun primarily carries the meaning of comparison, the Comparative suffix or adverb can be omitted from the adjective (Erd: 372; Berta 1998: 286; Boeschoeten 1998: 361; G/K: 199). The South Siberian Turkic languages lack the Comparative suffix –*rAK* that marks adjectives in the other Turkic languages; here, only the noun expressing the standard of comparison takes Ablative case-marking, while the comparee and the adjective remain unmarked, e.g. Tofa *neš ög–**den** bedik* [tree house–ABL high] 'the tree (is) higher than the house' (Ras: 97).

In the Tungusic languages as well, the standard of comparison is marked by the Ablative case, while the adjective in general remains unmarked; however, frequently it takes an intensifying suffix. In Evenki and Negidal, however, the adjective in the comparative degree is always marked with a Comparative suffix (–*tmAr/–dymAr* in Evenki), while the noun expressing the standard of comparison is marked by the Ablative case, e.g. Evenki *tar oron murin–**duk** hegdï–**tmer*** [that reindeer horse–ABL big–COMP] 'that reindeer is bigger than a horse' (Ned: 278; JaN5: 115). In Èven and Udihe, on the other hand, there is no comparative suffix to mark adjectives; in comparative constructions, the standard of comparison is marked by the Ablative case, while the adjective remains unmarked, e.g. Èven *Anna Marya–**duk** nose:gče:n* [A. M.–ABL young] 'Anna is younger than Maria' (Ben: 66). Frequently, the adjective takes an intensive marker (Ben: 66; N/T: 180, 189), e.g. Udihe *min–digi 'ai–**ndima** bi:–ni* [1SG.OBL–ABL elder.brother–INTS be–3SG] 'my brother is older than I am' (N/T: 189). This construction is similar to that found in Ulča and Orok, where, however, the standard of comparison is expressed by the Instrumental case (JaN5: 158; Pet: 62f). In Manchu the noun expressing the standard of comparison is marked by the Ablative case, while the adjective (classified as a 'noun of quality' by Gorelova) remains unmarked, e.g. *tere ere **či** sain* [that this ABL good] 'that is better than this' (Gor: 157). Comparative constructions in Mongolic

languages are formed in the same way as in most Tungusic languages, with the standard of comparison being marked by the Ablative case and the adjective remaining unmarked, e.g. Buryat *zürxeny–i:nï shulu:n–**ha:** xatu:* [heart–POSS.3 stone–ABL hard] 'his heart is harder than stone' (Skr: 109).

In Kolyma Yukaghir, the standard of comparison is marked by the Ablative case; the adjective remains unmarked, or can optionally be emphasized by a connective particle *a:y*, e.g. *tudel mit–**ket** a:y omoç modo–y* [3SG 1PL–ABL CP well sit–INTR.3SG] 'he lives even better than we do' (Mas: 364). In Ket, as well, the adjective remains unmarked, while the standard of comparison stands in the Ablative case (Wer: 124). In Mansi, there is a Comparative suffix *–nuw* that attaches to adjectives in the comparative degree, e.g. *karəs* 'tall', *karəs–nuw* 'taller'. However, the adjective can also remain unmarked in comparative constructions; the meaning of comparison is then expressed solely by the Ablative case-marking on the noun expressing the standard of comparison, e.g. *am kol–əm naŋ kol–ən–**nəl** janiɣ* [1SG house–POSS.1SG 2SG house–POSS.2SG–ABL big] 'my house is bigger than your house' (Rse: 29). In Obdorsk Khanty the standard of comparison is marked not by the Ablative case (which is lacking in this dialect), but by a postposition *e:wəlt* 'from' (Nik: 20f). In Vakh Khanty, on the other hand, there is a specialized suffix *–niŋət* which marks the standard of comparison (Gulya [1966] 1997: 54), with the adjective remaining unmarked, e.g. *tim kat tom kat–**niŋət** əllə–ki* [this house that house–COMP big–PRD] 'this house is bigger than that house'. The Ablative case can also serve to mark the standard of comparison, but this is less frequent (Gulya [1966] 1997: 68). In Nganasan, the standard of comparison is marked by the Ablative case, with the adjective remaining unmarked (Ter: 89f, 132); however, a few qualitative adjectives have suppletive comparative forms, e.g. *tanəgə:* 'wide', *tandudjə* 'wider' (Helimski 1998: 497). In Chukchi, comparative constructions are formed by deriving analytical verbs from the adjective. The standard of comparison, if overtly present, is marked by the Locative case (Dnn: 298). Although Itelmen consultants apparently willingly produce sentences expressing a comparison, such constructions do not occur in natural texts (G/V: 112; Stefan Georg pers. comm.). The elicited sentences do not contain a standard of comparison, and thus it is not known how Itelmen may have expressed this. The adjective in the comparative degree takes a suffix *–čeye*, e.g. *atx–q le–win a potom atx–**čeye*** [light–ADVR become–3SG but then light–COMP] 'it became light and then lighter' (G/V: 113). In Siberian Yupik Eskimo the adjectival meaning is expressed by a stative verb; one possible comparative construction is to have the comparee and the standard of comparison in the Absolutive and Relative case, respectively, with the stative verb taking the form of a possessive-marked participle; the second option is for the stative

verb to be intransitive, with the comparee standing in the Absolutive case and the standard of comparison in the Instrumental/Ablative case (Men: 72). In Nivkh there exists a special suffix to mark the standard of comparison, i.e. with a function similar to the Sakha Comparative case and the Vakh Khanty Comparative suffix. As in Sakha, this form is classified as a case suffix in Nivkh, e.g. *tlaŋi qanŋ–**ak** eɣ–d* [reindeer dog–COMP be.fast–FIN] 'a reindeer is faster than a dog' (Grz: 19). There is no separate adjectival class in Nivkh, where adjectival meanings are expressed by qualitative verbs (Grz: 16). Interestingly, in Nanay the standard of comparison is marked by the so-called 'Comparative form' and not by the Ablative case, which is missing in this Tungusic language, e.g. *morin ida–**duy** masi* [horse dog–COMP strong] 'a horse is stronger than a dog' (AvrI: 184). Avrorin suggests that the Comparative form in Nanay developed out of the obsolete Ablative case suffix, which marks the standard of comparison in all the other Tungusic languages, as mentioned above. Thus, this form is the result of the retention of one of the functions of the former Ablative case in a slightly modified form (*–duy* < *–duyi* < *–duxi* < *–duki*; AvrI: 185f). Although it is tempting to speculate that Nivkh contact influence may have played a role in the development of this Comparative form, other Tungusic languages have been spoken in closer contiguity with Nivkh than Nanay (Negidal, Ulča, and Orok; cf. Figure 1.3) without showing this development, so that an independent innovation is more likely.

3.2.5.3 The origins of the Sakha Comparative case in the light of Eurasian comparative constructions

As demonstrated by the above discussion, the most widespread construction to express comparisons in Siberia is for the standard of comparison to take Ablative case-marking, while the adjective remains invariant. There are some exceptions, where languages add a comparative suffix to the adjective as well as marking the standard of comparison by the Ablative case; these are most notably the Turkic languages and Mansi (where the comparative suffixes are, however, optional), as well as Evenki and Negidal. In other Tungusic languages, the adjective can, but need not be, marked with an intensifying suffix. The 'locational comparative' is characteristic of all of Asia with the exception of East and mainland Southeast Asia (Stassen 2005: map 121); however, in his coding Stassen does not distinguish between languages that use an ablative case, an allative case, or a locative case to mark the standard of comparison. Heine (1997: 128) provides a finer-grained classification of comparative constructions (based on data from Stassen 1985), and from this it becomes clear that it is precisely the 'source schema' (i.e. the use of a

marker expressing motion from) that is the most widespread in Asia (being found in 66% of the 24 languages of the sample), and the most frequent 'schema' world-wide.

There are only three languages apart from Sakha that have a specialized (case) suffix to mark the standard of comparison; these are: Vakh Khanty, Nivkh, and Nanay. For Nanay, a derivation of the Comparative form from the obsolete Ablative case is very plausible; this probably took place independently. Since Sakha is not and has not been in any known long-term contact with either Vakh Khanty or Nivkh, there are no grounds for postulating contact influence from these languages in the development of the Sakha Comparative case. This therefore represents an internal independent innovation in Sakha.

3.3 The Sakha Distant Future Imperative

Sakha differs from the other Turkic languages by making a formal distinction between Immediate Future and Distant Future Imperatives, while Turkic languages in general have only a single imperative paradigm that does not make a tense distinction. The similarity of the tense distinction expressed by the Sakha imperative paradigm to the Evenki imperative paradigm was already noted by Schönig ([1988] 1990: 54). The origins of the Sakha Distant Future Imperative are very complex, but it can be shown that areal contact influence, with Evenki as the primary source, may have played some role in its development.

Since imperatives generally have future, rather than present time reference (as a command/request can only be fulfilled after it has been uttered), often the terms 'immediate future imperative' and 'distant future imperative' are preferred (e.g. GSJa: 320; Avrorin 1961: 122). However, for the sake of brevity I will use the term 'present imperative' to designate the immediate future imperative and 'future imperative' to designate the distant future imperative in the following discussion, even though this may not be the most precise terminology.

3.3.1 The Sakha imperative forms

3.3.1.1 The Present Imperative

The Sakha Present Imperative is expressed in all three persons, both singular and plural. In addition, in the first person there is a minimal inclusive form, used when adressing one other person, as opposed to the augmented inclusive form, which is used when adressing a group of people (Dobrushina & Goussev 2005: 192). Sometimes the opposition between the forms is analyzed as an exclusive vs. inclusive imperative, or as a dual vs. plural imperative (Nasilov et al 2001: 190f; Ubr: 181f; Korkina 1970: 148). The Imperative mood is characterized by different person agreement suffixes from those found in the Indicative, cf. Table 3.11.

Table 3.11: Comparison of subject agreement suffixes used in the Indicative and Imperative moods in Sakha

	Indicative mood				Imperative mood		
	Predicative suffixes		Possessive suffixes				
	SG	PL	SG	PL	SG	MIN.INCL	PL
1	–BIn	–BIt	–(I)m	–BIt	–I:m	–IAχ	–IaɣIŋ
2	–GIn	–GIt	–(I)ŋ	–GIt	–Ø		–(I)ŋ
3	–Ø	–LAr	–(t)A	–LArA	–TIn		–TInnAr

It is possible to analyze the Present Imperative marker as being equal to the bare stem of the verb, as in all Turkic languages. In the second person singular subject agreement remains unmarked (or is marked by a zero morpheme[1]) (31a, repeated here for convenience as 63a), while the suffix –(I)ŋ marks agreement with the second person plural (63b). This analysis is further supported by the forms of the first person augmented inclusive Present Imperative (65c) and by the Future Imperative (67a,b). The suffix marking agreement with the third person singular –TIn (64a) is cognate to the Common Turkic suffix –sIn (cf. Table 3.12); to this the plural suffix –LAr is added in the third person plural (64b).

(63a)

bïra:kkïn	*ïskuolaγa*	*aγal*	*ere*
bïra:t–GIn	oskuola–GA	aγal–ø	ere
younger.brother–ACC.2SG	school–DAT	give[PRXIMP]–2SG	PTL

di:ller
die–Ar–LAr
say–PRSPT–PL

'Send your nephew (lit. your little brother) to school, they said.'

[PotP9]

(63b)

ehigi	*emie*	*kömölöh**üŋ***	*germaniyalar*
ehigi	emie	kömölös–(I)ŋ	germanya–LAr
2PL	also	help[PRXIMP]–2PL	Germany–PL

'You help too, Germans.'

[LukP241]

(64a)

je	*hïrït**tïn***	*törö:büt*	*doydutugar*	*dien*
je	sïrït–TIn	törö:–BIt	doydu–(t)IgAr	die–An
well	walk–IMP.3SG	be.born–PSTPT	land–DAT.3SG	say–PF.CVB

'Saying: "Well, let him walk around in his home country".'

[IvaP316]

---

[1] In this and the following section I analyze the 2SG person marking as zero to illustrate the structure of the imperative paradigms. Elsewhere, I analyze the 2SG imperative person marking as being inherent in the root or stem.

(64b)

je	*ol*	*beyeŋ*		*oɣolor*	*χayït**tïnnar***
je	ol	beye–(I)ŋ		oɣo–LAr	χayït–TIn–LAr
well	that	self–POSS.2SG		child–PL	chop–IMP.3SG–PL

dieχterin	*onnugar*
die–IAχ–LArIn	onnugar
say–FUTPT–ACC.3PL	instead.of

'…instead of them saying "Let your own children chop (wood)"…'

<div align="right">[LukP183]</div>

The first person agreement suffixes are Sakha innovations, with the non-singular forms (and possibly the singular, too) being based on the Future participle –IAχ[2]. The first person singular Present Imperative is marked by the suffix –I:m (65a), the origins of which are somewhat unclear. Böhtlingk (Btl: 303) suggests an origin out of the Future Participle marked for first person singular possessor (e.g. bïh–ïaɣ–ïm [cut–FUTPT–POSS.1SG]); this is supported by Korkina (1970: 151). However, it should be noted that this possessive-marked Future Participle (and the contracted form in –IAm, e.g. bïhïam) expresses the Indicative Future for the first person singular, i.e. bïhïaɣïm means 'I will cut'. The precise path of development from the Future Participle to the first person singular Present Imperative is therefore still not clear. The first person minimal inclusive Imperative has two forms that are in free variation; on the one hand the bare Future Participle (65b), and in addition the Dative-marked Future Participle. The first person augmented inclusive is derived from the short form of the minimal inclusive (i.e. the bare Future Participle) through addition of the Imperative second person plural marker –(I)ŋ, giving –IAɣIŋ (65c) (cf. Nasilov et al. 2001: 191; Dobrushina & Goussev 2005: 196).

(65a)

če	*ere*	*kör**ü:m***	*ere*
če	ere	kör–I:m	ere
well	PTL	look–IMP.1SG	PTL

'Well, let me have a look!'

<div align="right">[spontaneous utterance, Ver]</div>

---

[2] Ubrjatova (Ubr: 181f), however, disputes this analysis. Instead, she regards this as consisting of a suffix –IA followed by an old agreement marker for the first person plural in –χ. Schönig (2003 [1987]: 9) analyses it differently again, reconstructing a form *–AIIQ for this, and *–AIIQIŋ for the first person plural.

(65b)

manan   *barïaχ*
manan   bar–IAχ
this.INS   go–IMP.MIN

'Let us (two) go this way!'

[spontaneous utterance, Ver]

(65c)

če   *oγolor*   *taχχan*   *ohuoχaydïaγïŋ*   *dietim*
če   oγo–LAr   taγïs–An   ohuoχay–LA:–IAχ–(I)ŋ   die–TI–(I)m
well child–PL   go.out–PF.CVB   ohuoxay–VR–IMP.MIN–2PL   say–PST–POSS.1SG

'"Hey, kids, let's go out and do the ohuoxay (circular dance for summer solstice festival)", I said.'

[YmyE12]

The Present Imperative conveys an exhortation in the first and third persons and an immediate command or request in the second person. It can be negated by the standard negative suffix *–(I)mA* (example 32a, repeated here as 66):

(66)

oγoloru   *χa:yan*   *hïljïma*   *di:r*
oγo–LAr–(n)I   χa:y–An   sïrït–(I)mA–ø   die–Ar
child–PL–ACC   lock.up–PF.CVB   go[PRXIMP]–NEG–2SG   say–PRSPT

'"…don't leave the children alone", he said.'

[RaxA271]

3.3.1.2 The Future Imperative

The Sakha Future Imperative, which is marked by the suffix *–A:r*, is restricted to the second persons. As in the Present Imperative mood, agreement with second person singular is unmarked (67a), and agreement with the second person plural is marked with the suffix *–(I)ŋ*. The Future Imperative, too, is regularly negated with the negative suffix *–(I)mA*, with the final vowel of the negative suffix merging with the initial vowel of the Future Imperative suffix (67b). It conveys a command that is to be fulfilled at a later point in time (67c), or after some other action (67d), and is also used for weaker commands or polite requests (67e).

(67a)

je	*noχo:*	*torbosto:χ*	*ïnaχpïtïn*
je	noχo:	torbos–LA:χ	ïnaχ–BItIn
well	boy	calf–PROP	cow–ACC.1PL

illihe:r		*ere*	*dien*
ilt–(I)s–A:r–ø		ere	die–An
take.away–REC–DSTIMP–2SG		PTL	say–PF.CVB

'Saying "Hey, boy, just take our cow with calf (there) with me".'

[PotP40]

(67b)

ehigi	*sieme:rïŋ*	*dien*	*aɣam*
ehigi	sie–(I)mA–A:r–(I)ŋ	die–An	aɣa–(I)m
2PL	eat–NEG–DSTIMP–2PL	say–PF.CVB	father–POSS.1SG

keriehin	*eppite*
keries–(t)In	et–BIt–(t)A
memory–ACC.3SG	say–PSTPT–POSS.3SG

'"Don't you eat (it)," my father said as a testament.'

[XatR136]

(67c)

bu	*tülüppüönünen*	*kepsete:r*	*dien*
bu	tülüppüön–(I)nAn	kepse:–T–A:r–ø	die–An
this	telephone–INS	tell–CAUS–DSTIMP–2SG	say–PF.CVB

'"Call me on the telephone", he said.'

[YmyE53]

(67d)

hötüöle:n	*kelen*	*baran*	*χoskun*
sötüöle:–An	kel–An	bar–An	χos–GIn
bathe–PF.CVB	come–PF.CVB	SEQ–PF.CVB	room–ACC.2SG

χomuya:r
χomuy–A:r–ø
tidy.up–DSTIMP–2SG

'After coming back from swimming, tidy up your room.'

[translation, Sun (very similar in all districts)]

(67e)

sibe:s	*tuhunan*	*kepse:ri:y*
sibe:s	tus–(t)InAn	kepse:–A:r–I:y
communications	side–INS.3SG	tell–DSTIMP[2SG]–EMPH

'Tell about the (phone) connection.'   [LukP188[3]]

---

[3] Actually, this was said to P. Lukinov by a third party, a younger woman working for the village administration.

3.3.2 The imperative in other Turkic languages

All Turkic languages have a full person/number paradigm in the imperative mood, with a number of shared cognate forms (cf. Table 3.12). In Old Turkic, the second person singular is unmarked, as is common throughout the Turkic languages, e.g. Turkish *bu para–yı baba–n–a ver* [this money–ACC father–2SG.POSS–DAT give[IMP.2SG]] 'Give this money to your father!' (G/K: 360). In the Orkhon Turkic inscriptions the second person plural is marked by the suffix *–(X)ŋ* (Erd: 237). However, in Uygur texts this latter form was used only as a polite form for the second singular, with its plural-marked form *–(X)ŋlAr* being used for the second person plural, e.g. *tur–**uŋ*** 'stand up' used in an address to a king, and *odun–**uŋlar*** 'wake up' said to a large number of people (Erd: 520). Variants of this suffix are common throughout the Turkic languages to mark second person plural imperatives (Nasilov et al. 2001: 185). The Old Turkic third person form is *–zUn(lAr)*, which is found in all modern Turkic languages except for Chuvash and Khalaj, e.g. Turkish *çocuk–lar burada kal–**sın(lar)*** [child–PL here remain–IMP.3(PL)] 'Let the children stay here' (G/K: 360). The Sakha suffix *–TIn(nAr)* is clearly cognate to this. In the first person singular we find *–(A)yIn* in Orkhon Turkic, reflexives of which are found in a number of modern languages, and in the first person plural *–(A)lIm* (Erd: 235). This may be a secondary development in Old Turkic from Proto-Turkic *–(A)lI*, which is found in several modern languages. In all the Turkic languages, the negative imperative is regularly formed by adding the negative suffix *–mA/–BA* to the stem, e.g. Tatar *bar–**ma**–ɣïz* 'do not go (PL or polite)', Shor *pas–**pa**–ay* 'let me not write' (Nasilov et al. 2001: 193).

A minimal/augmented inclusive distinction is made in the first person imperative forms in the Siberian Turkic languages Tuvan, Khakas, Shor, Chulym, and Altay as well as in Turkmen, Khalaj and Bashkir dialects (Schönig [1987] 2003: 8; Nevskaya 2005: 342ff). In some cases, the augmented inclusive suffix is formed from the minimal inclusive by addition of the plural suffix *–LAr*, e.g. Shor *par–**a:ŋ*** 'let you and me go', *par–**a:ŋ–nar*** 'let all of us go'. In others, as in Sakha, Tofa, or Tuvan, the augmented suffix is formed from the minimal inclusive plus the imperative plural marker (Nasilov et al. 2001: 184, 191; Schönig [1987] 2003: 8; Nevskaya 2005: 347f).

Table 3.12: Person agreement suffixes of the (present) imperative mood in some Turkic languages

	1SG	2SG	3SG	1min incl	1PL/ 1aug incl	2PL	3PL
Old Turkic	*-(A)yIn*	*–ø, –(X)ŋ*	*-zUn*		*-(A)lIm*	*-(X)ŋ, (X)ŋ(lAr)*	*-zUn(lAr)*
Turkish	*-(y)AlIm*	*–ø*	*-sIn*		*-(y)AlIm*	*-(y)In(Iz)*	*-sIn(lAr)*
Kazakh	*-(A)yIn*	*–ø*	*-sIn*		*-(A)yIK*	*–(I)ŋdAr*	*-sIn*
Uzbek	*-(A)y(In)*	*–ø,(-gil)*	*-sIn*		*-(A)ylIK*	*–(I)ŋ(lAr)*	*-sInlAr*
Tuvan	*-Ayn*	*–ø*	*-zIn*	*-(A)Al(I)*	*-A:lIŋAr*	*-IŋAr*	*-zIn(nAr)*
Khakass	*-Im*	*–ø*	*-ZVn*	*-Aŋ*	*-AŋAr*	*-(V)ŋAr*	*-ZVnnAr*
Tofa	*–(V:)yIn*	*–ø*	*–sIn*	*–V:lI*	*-A:lIŋ(Ar)*	*–(I)ŋAr*	*-sIn(nAr)*
Chuvash	*–Am*	*–ø*	*–tĂr*		*–Ar*	*–Ăr*	*–ččĂr*
Khalaj	*–dVm*	*different suffixes*	*– (I)tA*		*–dVk*	*-(i)dUŋ(iz)*	*–(I)tAlAr*
Dolgan	*–I:m*	*–ø*	*–TIn*	*–IAk*	*–IagIŋ*	*–(I)ŋ*	*–TInnAr*
Sakha	*–I:m*	*–ø*	*–TIn*	*-IAχ(χA)*	*–IaɣIŋ*	*–(I)ŋ*	*–TInnAr*

From the above discussion it becomes clear that Sakha resembles the other Turkic languages with regard to its Present Imperative paradigm. However, Sakha differs from its sister languages (with the exception of Dolgan) in that none of the Turkic languages make a formal distinction between commands that are to be fulfilled immediately and those that may be fulfilled at a later point in time (cf. Table 3.16).

3.3.3 The imperative in Mongolic languages

The comparison of the Sakha Imperative to the imperative mood in Mongolic languages is complicated by the Mongolianist tradition of giving every type of suffix a different name, which creates the impression of there being a huge variety of different imperative forms. Upon closer scrutiny, however, it becomes clear that only in the second person do we find different imperative forms (generally conveying ever higher degrees of politeness), while there is only one form for the first and third persons, respectively (cf. Table 3.13). Furthermore, not all the languages have all the second person imperative forms which are typical of Khalkha Mongolian and its closely related dialects/sister languages.

Table 3.13: Imperative forms in Mongolic languages[4] (named following traditional Mongolianist terminology)

	1st person	2nd person	3rd person
Imperative/Hortative	Voluntative	Imperative	Concessive
		Precative	
		Benedictive	
Future Imperative?		Prescriptive	

As can be seen from Table 3.13, the so-called Voluntative is restricted to the first person, while the so-called Concessive is restricted to the third person. These forms convey exhortations (Sanžeev 1964: 110, 104). The so-called Imperative in the Mongolic languages is equal to the stem of the verb; it conveys an immediate command or request (68a), and is the least polite of the imperative forms (Kuzmenkov 2001: 105). The Precative conveys an emphatic request (68b), while the Benedictive conveys a very polite request (68c), i.e. we find an ever-increasing degree of politeness in these forms. The different forms of the Khalkha Mongolian Imperative Mood are negated with a separate negative particle *bitgiy* (rarely *bü:*; 68d) which is used only in this function (Vietze 1988: 39; K/TS: 175, 334).

(68a)
  ta_nar odo: or
  2PL   now   enter[IMP]
  'Go in now!'

                                       [Vietze 1988: 39]

(68b)
  *či   ene   ažl–ïg   xiy–**ge:č***
  2SG   this   work–ACC make–PREC
  'Please do this work!'

                                       [K/Ts: 178]

(68c)
  *ta   su:–**gtun***
  2SG   sit–BND
  'Be seated, please.'

                                       [K/Ts: 181]

---

[4] It should be noted that Mongolianists (e.g. Sanžeev 1964: 95ff) include the Optative and Dubitative in the imperative forms as well. I have excluded them from the table as I focus specifically on imperative (command/request) meanings.

(68d)

ta_nar	**bitgiy**	su:
2PL	NEG	sit[IMP]

'Don't sit down!'

[K/Ts: 177]

In the light of the discussion concerning the origins of the Sakha Future Imperative, the so-called Prescriptive form is of particular interest. This form is marked by the suffix –A:rAy, the similarity of which to the Sakha Future Imperative suffix –A:r was noted already by Böhtlingk (Btl: 304). Different authors differ in the emphasis they place on nuances of meaning of this form; for instance, Sanžeev (1964: 98) and Poppe (1955: 254) emphasize a future imperative meaning. Similarly, Poppe (PopWM: 90) writes that in Written Mongolian the Prescriptive (which occurs only under dialectal influence) "expresses a commission or a wish addressed to a second person. The action which the person concerned is ordered to perform may be done later on, if not immediately." On the other hand, Kuzmenkov (2001: 99, 105) emphasizes the higher degree of politeness conveyed by the Prescriptive in (Khalkha) Mongolian.

According to Janhunen (2003d: 22f), the Prescriptive was already present in Proto-Mongolic, with the reconstructed form *–xA–rA.(y)I. This may have differed from the other Proto-Mongolic second-person form, the Benedictive, "by the degree of politeness, the prescriptive being more casual and the benedictive more polite." (Janhunen 2003d: 23). However, Janhunen's reconstruction of the Prescriptive to Proto-Mongolic is somewhat surprising given the fact that this form is not attested in Middle Mongolian[5], the language spoken in the period of Chinggis Khan's Empire (Rybatzki 2003: 74; Poppe 1955: 254), and is nowadays found in a very restricted number of Mongolic languages: Khalkha, Buryat, Khamnigan Mongol, and Oirat (Svantesson 2003, Skribnik 2003, Janhunen 2003c, Birtalan 2003). It therefore seems to be a fairly recent innovation amongst the languages spoken in the vicinity of Lake Baykal.

In Khalkha Mongolian, the Prescriptive is described as expressing a weak command, a demand, or a request (69) (Vietze 1988: 99; K/Ts: 179). Kuzmenkov (2001: 105) ranks the Mongolian imperative forms by increasing degree of politeness as follows: Imperative, Precative, Prescriptive, and Benedictive.

---

[5] All the more so since Janhunen himself writes: "It happens that a language basically identical with Proto-Mongolic is historically attested in a multitude of written sources dating from the Yuan and early Ming dynasties. As a documented idiom this language is conventionally termed Middle Mongol (Middle Mongolian) […]" (Janhunen 1996: 145f).

(69)

*ene em–iyg   xo:l–ny   ömnö   u:–**ga:ray***
this medicine–ACC meal–GEN before   drink–PRESCR

'Please take this medicine before meals.'

<div align="right">[K/TS: 179]</div>

Similar descriptions of the Prescriptive as being a more polite imperative marker, not a future imperative, are also given for Khamnigan Mongol and Oirat (Janhunen 2003c: 93f, Birtalan 2003: 222). In Khamnigan Mongol, "[…] the prescriptive and benedictive, as compared with the basic unmarked imperative, express successively more polite requests addressed to the second person, both singular and plural, e.g. imp. *yabu* 'go!', prescr. *yabo:rie* '(please) go!', ben. *yabu**gtui*** '(would you please) go!'." (Janhunen 2003c: 93f).

Buryat is the only modern language for which the Prescriptive is described as having not so much a nuance of added politeness, but a future imperative meaning (PopB: 60), e.g. *yab–**a:rai**–t* [go–PRESCR–PL] 'you.PL can go (later)' (Skr: 113).

3.3.4 The imperative in Tungusic languages

3.3.4.1 The Northern Tungusic languages

3.3.4.1.1 Evenki

Evenki has two person-number paradigms for the Imperative mood, one of which has immediate future time reference and is more categorical (70a), the other of which expresses commands/requests which may be fulfilled at a later point in time, often after some other action (70b, c), and which is more polite. However, this distinction between the two imperative paradigms is not a strict rule, merely a tendency in use (Ned: 18f; B/G: 36).

(70a)

*ju–la–vi   himat   eme–**kel***
house–LOC–PREFL   fast   come–PRXIMP.2SG

'Come quickly to my place.'

(70b)

　ju–la–vi　　　　　*(gočin)*　　　*eme–**de:**–vi*
　house–LOC–PREFL　(next.year)　come–DSTIMP–PREFL.SG

　'Come to my place (next year)[6].'

<div align="right">[Ned: 19]</div>

(70c)

　bira–va　　　*dag–mi*　　*guluvun–ma　ila–**da:**–vi*
　river–DEF.ACC cross–CVB　fire–DEF.ACC burn–DSTIMP–PREFL.SG

　'(Upon) crossing the river make a fire.'

<div align="right">[Ned: 262]</div>

　　　　The Present Imperative in Evenki takes a different set of subject agreement markers from those found in the Indicative mood, with portmanteau suffixes expressing both person and mood; these are attached to the bare stem of the verb. The forms are 1SG *–ktA*, 2SG *–kAl*, 3SG *–gin*, 1PL.excl *–ktA–vun/–vvun*[7], 1PL.incl *–gAt*, 2PL *–kAllu* and 3PL *–ktïn* (Ned: 19, 262). These may go back to common Tungusic forms *–kal/*–ki* and *–ga:/*–gi* with personal suffixes, e.g. 1SG *–ktA* < *–ki–ta* or *–ka–ta* (which in turn may have led to the 1PL.incl form *–gAt*) (Sunik 1962: 192). The Future Imperative, however, is marked by a separate suffix to which the same subject agreement markers as those found in the Indicative mood are added (cf. Table 3.14). Furthermore, the Future Imperative paradigm is split, with the Future Imperative suffix in the first and third person being *–ŋnA:* plus type 1 subject agreement markers found in the Indicative mood (Ned: 262). The 3PL suffix *–tïn* is an exception, since this is the 3PL possessive suffix (also used as a subject agreement marker in some tense-aspect forms). In the second person, however, the Future Imperative is marked by the Purposive converb *–dA:* plus reflexive possessive subject agreement markers (Ned: 262; Sunik 1962: 339; cf. Table 3.14). It is very interesting in this respect that Bulatova & Grenoble (B/G: 37) give only second person, and not any first or third person forms, for the Future Imperative. Konstantinova (Kon: 184) however, gives a full paradigm for both the Present Imperative and the Future Imperative; she, too, points out that the Future Imperative is often used to express an action that is to be fulfilled after some other action:

---

[6] Note the discrepancy between the subject of the imperative verb (2SG) and the possessor of the house (1SG according to the translation), which is nevertheless encoded as a reflexive possessive.

[7] Bulatova & Grenoble (B/G: 36) give the suffix *–kvun* for the Present Imperative 1PL.excl.

(71)

derumki–t–čele:–tïn	*ju:–dï–va*	*hava:–ve*	*me:rïn*
rest–??–ANT.CVB–3PL	house–ADJR–DEF.ACC	work–DEF.ACC	self.3PL

o:–ŋna–tïn
make–DSTIMP–3PL

'After they have rested let them do the housework themselves.'

[Kon: 183]

Table 3.14: Comparison of subject agreement suffixes used in the Evenki Indicative
and Future Imperative (Ned: 259, 262)

	'Type 1', Indicative		Future Imperative	
	SG	**PL**	**SG**	**PL**
1incl	*–m*	*–p*	*–ŋnA–m*	*–ŋnA–p*
1excl		*–v*		*–ŋnA–v*
2	*–nni*	*–s*	*–dA:–vi*	*–dA:–ver*
3	*–n*	*–ø*	*–ŋnA–n*	*–ŋnA–tïn*

As mentioned above the second person Future Imperative suffix is *–dA:*, which is identical to the Purposive Converb. This converb has variable subject agreement forms (cf. section 3.5), with a non-coreferential subject of the purpose clause being expressed by personal possessive suffixes on the converb. Coreferential subjects, on the other hand, are expressed by the reflexive possessive suffixes *–vi* (singular) and *–ver* (plural) (72) (Ned: 52).

(72)

bi	*nuŋan–dula–n*	*tuksa–ča–v*
1SG	3SG–LOC–POSS.3SG	run–PST–1SG

dukuvun–ma–s	*bu:–**de:**–**vi***
letter–DEF.ACC–POSS.2SG	give–PURP–PREFL.SG

'I ran to him to give him your letter.'

[Ned: 52]

As can be seen in Table 3.14, the Future Imperative forms for the second person consist of the Purposive converb plus the possessive reflexive suffixes; thus, the purposive form in example (72) *bu:–**de:**–vi* actually has two readings: '(x did something) in order that x give' and '(you.SG) give (later)'. This may give rise to ambiguity; for example, in a sentence 'Go (later) to the neighbour to give him his axe', both the verb 'go' and the verb 'give' would probably carry the same suffixes *–dA:* and *–vi*, and the reading would be ambiguous between the above with a Future

Imperative and a purposive clause, and 'Go (later) to the neighbour and give (later) him his axe', with two Future Imperatives (Igor' Nedjalkov, pers.comm.). Of course, the actual difference in meaning between the two sentences is minimal, as is further shown by the fact that in French the first reading is grammaticalized ('va lui donner le livre' with an infinitive), while in English the latter reading is grammaticalized ('go and give him the book'; Bernard Comrie, pers. comm.). This ambiguity between same-subject purpose clauses and Future Imperatives is also found in Èven (see section 3.3.4.1.2).

Evenki has a further Monitory Imperative that expresses warnings, which is restricted to the second person, e.g. *er–tĭki tar–tĭki iče–t–**ne*** [this–ALL that–ALL see–PROG–MON] '(Be careful and) look in different directions.' (Ned: 266). The Imperative mood in Evenki is negated in the manner of the Indicative, with the negative auxiliary *e–* taking the imperative marking and the main verb taking a connegative suffix (73). Negative imperatives are restricted to the Present Imperative and Monitory Imperative (Ned: 19f).

(73)

tar	*beye*	*e–**gin***	*eme–**re***
that	man	NEG–PRXIMP.3SG	come–CONNEG

'Don't let that man come.'

[Ned: 20]

In Khamnigan Evenki, an Evenki language spoken in close contact with Khamnigan Mongol in Manchuria, the Present Imperative has separate forms for all persons and numbers excepting the first person plural exclusive. The second person Present Imperative suffixes are based on the common Evenki form *–kAl*, while the first singular and third person suffixes are based on the element *–gi* followed by possessive person suffixes; the first person plural inclusive marker has the suffix *–gA:r* (Janhunen 1991: 85). The Future Imperative, which is based on the Purposive converb *–da:* as in Evenki proper, appears to be restricted to the first and second persons. Interestingly, all person-number combinations of the Future Imperative appear to take only one subject agreement suffix, the singular reflexive possessive suffix *–bi*. This may be due to influence from Khamnigan Mongol which, like the other Mongolic languages, has invariant forms in the Imperative mood (Janhunen 1991: 86).

3.3.4.1.2 Èven

In Èven there are three imperative paradigms differing in their temporal and pragmatic usage. There are two different analyses of the forms belonging to the different paradigms, as can be seen in Table 3.15, with Novikova (1980: 74ff) proposing a full paradigm each for the Present Imperative and Future Imperative, while Rišes & Cincius (1952: 736f) and Malchukov (2001) analyze the Present Imperative as being restricted to the second person singular and plural and the first person plural inclusive. As can be seen from the table, Novikova's analysis presents a mixture of the paradigms suggested by Malchukov (Novikova 1980: 76ff). This may be an indication of the rather recent grammaticalization of some of these forms, which are based on the Purposive Converb –*da* and the Future Participle –*jiŋaβ* (~ –*ńŋaβ*); possibly, fully distinct meanings of the separate forms have not yet crystallized. Furthermore, the confusion may be due to the proximity in meaning between future imperatives and polite imperatives, as discussed above for Sakha (section 3.3.1.2) and Evenki (section 3.3.4.1.1).

Table 3.15: Èven imperatives: mood and person suffixes (comparison of Malchukov 2001 and Novikova 1980)

	Immediate future		Distant future		Polite	
	Mal	Nov	Mal	Nov	Mal	Nov
1SG		–*da–ku*	–*da–k.u*	–*ńŋaβ*	–*jiŋa–β*	
2SG	–*li*	–*li*	–*da–y*	–*da–y*	–*ŋa–nri*	–*ŋa–nri*
3SG		–*da–n*	–*da–n*	–*ńŋaβ–ən*	–*jiŋa.βa–n*	
1PLin	–*gar*	–*gar*	(–*da–t/* –*da–βur*)	–*ńŋaβ–ur*	–*jiŋa–βur*	
1PLex		–*da–kun*	–*da–k.un*	–*ńŋaβ–un*	–*jiŋa–βun*	
2PL	–*lra/* –*lilra*	–*lilre*	–*da–βur*	–*da–vur*	–*ŋa–san*	–*ŋa–san*
3PL		–*da–tan*	–*da–tan*	–*ńŋaβ–u–tən*	–*jiŋa.βu–tan*	

The Present Imperative expresses commands, requests or invitations to joint actions that are to be performed immediately (74). The subject agreement forms for the second persons and first plural inclusive are different from those found in other parts of the verbal inflection.

(74)
  *il–**li**=si*   *hinmač   učik–tuki–y*
  stand.up–PRXIMP.2SG=CLIT   quickly reindeer–ABL–PREFL.SG
  'Come, quickly dismount your reindeer.'

                                                    [Malchukov 2001: 163]

The first and third person of the imperative forms that are based on the Purposive Converb *–da* take personal possessive subject agreement markers. They often express less of an imperative meaning than rather the wish of the speaker, and can be used to ask for permission or to give consent. In their imperative meaning they do not necessarily have distant future temporal reference (75a, b) (Malchukov 2001: 164f), which may explain why Novikova considers them Present Imperative forms (Novikova 1980: 77).

(75a)

*bi–**de**–n*   *hiŋke–**de**–n*
be–DSTIMP–3SG   crackle–DSTIMP–3SG

'Let it alone, let (the fire) crackle!'

[Malchukov 2001: 164]

(75b)

*koye–**li**!*   // *koye–**de**–ku=lu*
look–PRXIMP.2SG   look–DSTIMP–1SG=CLIT

'Look! // OK, let me have a look.'

[Malchukov 2001: 166]

Like the corresponding Evenki forms, the second person Future Imperative (which is based on the Purposive Converb *–da* and analyzed as a Future Imperative by both Malchukov and Novikova) takes reflexive possessive person markers, which might be an indication that it originated from a purposive form coreferential with a second-person Imperative (Malchukov 2001: 167). These forms express primarily a future imperative meaning (76a). The imperative forms that are based on the Purposive Converb *–da* are both morphologically and syntactically identical to the latter and are occasionally ambiguous in meaning. Thus, example (76b) can be read either as 'go and repeat' or as 'go in order to repeat' (Malchukov 2001: 166).

(76a)

*ilan   dolbani–ß   bi–siji   emu–**de**–y*
three   night–ACC   be–ANT.CVB   bring–DSTIMP–PREFL.SG

'Bring it in three days.'

[Malchukov 2001: 165]

(76b)

*tiek hor–**lilre**   ju–tki–ßur   mer   hunŋi–tki–ßur*
now go–PRXIMP.2PL   house–ALL–PREFL.PL   self's   chief–ALL–PREFL.PL

*Menenje–tki   ere–ß   gon–če–ß   alma–**da**–ßur*
M.–ALL   this–ACC   say–PSTPT–ACC.1SG repeat–DSTIMP–PREFL.PL

'Now go to your chief Menenje (in order to/and) repeat what I said.'

[Malchukov 2001: 162]

The third set of imperative forms is based on the Future Participle –*jiŋaβ/* –*ńŋaβ* and in this form is restricted to the first and third persons (Novikova 1980: 79; Malchukov 2001: 169). Novikova analyses it as the Future Imperative for the first and third persons, and gives corresponding examples (Novikova 1980: 79); see for instance (77a). Malchukov analyses these forms as part of the Polite Imperative paradigm; however, he, too, gives an example with a future meaning (77b). The second persons of the Polite Imperative paradigm are formed with the suffix –*ŋa*. As can be seen from Table 3.15, in Novikova's analysis the Polite Imperative is restricted to the second persons (Novikova 1980: 76). Novikova gives very polite translations for these forms, e.g. *gö–**ŋe**–nri* [say–POLIMP–2SG] '**please** say', *duk–**ŋa**–san* [write–POLIMP–2PL] '**please** write' (Novikova 1980: 76, emphasis mine), while the translations given by Malchukov (2001: 168) appear to convey more of a future time reference (77c). As was seen in the above discussion of the Sakha and Evenki Future Imperative, the meanings of future imperatives and polite imperatives are semantically quite close; thus, both the Sakha and the Evenki Future Imperative have a more polite meaning than the Present Imperative (cf. section 3.3.1.2 and 3.3.4.1.1). This may explain the different analyses given by Novikova and Malchukov, even though judging from the examples they appear to have analyzed partly the same data.

(77a)

min–u   *mö–le*   *ule–**de**–y!*   *// ey*   *ule–**ńŋe**–β*
1SG–ACC   water–LOC throw.DSTIMP–PREFL.SG   OK   throw–POLIMP–1SG

'Throw me into the water (later)! // OK, let me throw you (later)!'

[Novikova 1980: 79]

(77b)

timinak   *beyji*   *hor–**jiŋe**–βu*
tomorrow   oneself   go–POLIMP–1SG

'Tomorrow I shall go myself.'

(77c)

ŋi=de   *hin–teki*   *dagam–raka–n*   *bogukla–**ŋa**–nri*   *ninkami–ji*
who=CLIT   2SG–ALL   approach–COND–3SG   beat–POLIMP–2SG   staff–INS.REFL

'If anybody approaches you, hit (him) with your staff.'

[Malchukov 2001: 168]

In addition to the above-mentioned imperative paradigms, Èven has a full Preventive paradigm that expresses warnings not to perform actions that might adversely affect the addressee (or speaker, in the case of the first person) (Malchukov 2001: 178).

The Ėven imperative forms are negated periphrastically with the help of the negative auxiliary *e–*. In the second person Present Imperative, the auxiliary takes a special prohibitive suffix (not identical to the imperative suffix), and the main verb furthermore takes a special Prohibitive Converb marker (Malčukov 2001: 176; Novikova 1980: 76), e.g. *duk–li* [write–PRXIMP.2SG] 'write!', *e–ji duk–kil* [NEG–PRXPROH write—PROH.2SG] 'don't write!' (Novikova 1980: 76); see also example (78). The second person Future Imperative is negated with yet another prohibitive form of the negative auxiliary, in conjunction with the standard connegative form of the main verb, e.g. *e–miken asaŋ–gara–r* [NEG–DSTPROH.SG be.angry–ITER–CONNEG] 'don't be angry with me (later)' (Malchukov 2001: 177; Novikova 1980: 76f). The first and third person negative imperative forms, on the other hand, are formed syntactically from the negative auxiliary *e–* with the imperative suffixes and the main verb with the standard connegative participle *–r*, e.g. *e–ger ma–r* [NEG–PRXIMP.1PL.INCL kill–CONNEG] 'let us not kill', *e–de–n ma–r* [NEG–DSTIMP–3SG kill–CONNEG] 'don't let him kill (afterwards)' (Malchukov 2001: 177).

(78)

eń=e	*e–ji*	*min–u*	*bagak–kil*
mother=VOC	NEG–PRXPROH	1SG.OBL–ACC	beat–PROH.2SG

'Mum, don't beat me!'

[Malchukov 2001: 176]

3.3.4.1.3 Negidal

Negidal also distinguishes between a Present and a Future Imperative, the first occurring with all persons and numbers, while the latter appears to be restricted to the second person. As is the case for the second person Future Imperative in Evenki and Ėven, the Negidal Future Imperative is formed from the Purposive Converb *–da:* with reflexive possessive suffixes to mark the distinction between second person singular and plural (Cincius 1982: 35, 36). Unfortunately, Cincius only gives tables with paradigms and does not discuss the use of the Imperative forms.

3.3.4.2 The Amur Tungusic languages and Manchu

In Udihe, a dedicated imperative form exists only for the second person, both singular and plural (39d, repeated here as 79a); the person-mood suffixes –*ya* (2SG) and –*ya–u* (PL) differ from other parts of the verbal paradigm. For particularly categorical orders it is also possible to use just the bare stem of the verb in the second singular (79b) (N/T: 221). For exhortations in the first and third person the Subjunctive mood is used; this, however, does not have a dedicated imperative meaning, but fulfills a number of other functions as well, such as expressing necessity and uncertainty. With the second person, the Subjunctive expresses a less categorical imperative than the Imperative mood (79c) (N/T: 265f). Nikolaeva & Tolskaya do not mention any future imperative form for Udihe.

(79a)

min–du   ulə:–wə   xauliə bu–ya
1SG–DAT   meat–ACC please give–IMP.2SG

'Please, give me some meat!'

(79b)

baulima–wa   diga
corn–ACC   eat[IMP.2SG]

'Eat the corn.'

(79c)

mama   xələ–ʐə–i   guŋ–ki–ni
grandmother   hurry.up–SUB–2SG   say–PST–3SG

'She said: "Grandmother, please hurry up." '

<div align="right">[N/T: 264, 265]</div>

Nanay has a Present Imperative with a full paradigm, and a Future Imperative in the second persons. While the former is used to express commands that are to be fulfilled immediately after the speech act (80a), the latter expresses commands that are to be fulfilled at a later point in time or after some other action (80b) (Avrorin 1961: 122). The suffix marking the Future Imperative, –*xAri*, is restricted to this use (Sunik 1962: 339). In the first person singular, the Purposive Converb is used with exhortative meaning, e.g. *mi taosi iče–nde–**gui**–ve* [1SG there see–GL–PURP–1SG] 'let me go and see' (Avrorin 1961: 129).

(80a)

*esi-tul   min–či   di–**du***
right.now   1SG–ALL   move–PRXIMP.2SG

'Come to me right now.'

(80b)

čimana	*xoton–či*	*ene–**xeri***
tomorrow	town–ALL	go–DSTIMP.2SG

'Tomorrow go to town.'

Ulča and Orok, both of which are very closely related to Nanay, also make a distinction between a Present Imperative and a Future Imperative (Sunik 1962: 339), while Oroč, which is closely related to Udihe, does not (A/B: 317ff). The Future Imperative suffixes in Ulča [*–sAr(I)*] and Orok (*–sAri*) are similar to the Nanay form (Sunik 1985: 45; Pet: 109). It thus appears that the Northern Tungusic languages use the Purposive Converb to express the future imperative, while amongst the Amur Tungusic languages the closely related Udihe and Oroč lack a future imperative and the closely related Nanay, Ulča and Orok use a cognate suffix restricted to this function.

Manchu has several different imperative forms; however, these do not make a temporal distinction, but rather social distinctions, being used to address equals, people of lower standing, or people of higher standing (Gor: 296ff). The bare verb stem is the most frequently used imperative form; this distinguishes Manchu from the other Tungusic languages (Gor: 297), which generally have suffixally marked imperative forms, as described above (sections 3.3.4.1 and 3.3.4.2) (although in Udihe the bare verb stem may be used in imperative function as well, as noted above).

3.3.5 The imperative in other Siberian languages

Kolyma Yukaghir has a Present Imperative paradigm for all person-number combinations with the exception of the first person singular; in addition, it has a Future Imperative restricted to the second person singular and plural. The Present Imperative is marked by the suffix *–ge* in the first and third persons, with the suffix *–n* marking agreement with the third person (and the plural marker *–ŋi* additionally marking number agreement for 3PL); in the first person plural the Present Imperative consists solely of the verb root plus the Imperative suffix *–ge*, with no further person-number marking. In the second person the Present Imperative is basically unmarked, consisting of the verb root followed by the second person suffix *–k* and the plural marker *–ŋi* for 2PL (Mas: 140, Table 14). Thus, the Present Imperative paradigm for the verb *yaqa* 'reach' is: *yaqa–ge* (1PL), *yaqa–ge–n* (3SG), *yaqa–ŋi–ge–n* (3PL), *yaqa–k* (2SG), and *yaqa–ŋi–k* (2PL). The Future Imperative, which, as mentioned above, is restricted to the second person, is marked by the

suffix *–ge* also found in the first and third person Present Imperative; this is followed by the suffix *–k* to mark agreement with the second person, while the 2PL Future Imperative additionally takes the plural marker *–ŋi*, e.g. *yaqa–ge–k* (Future Imperative 2SG 'reach'), *yaqa–ŋi–ge–k* (Future Imperative 2PL). Thus, the Future Imperative forms formally resemble the Present Imperative forms of the third and first person, since these all carry a suffix *–ge–* that is lacking in the Present Imperative for the second persons (Mas: 140). The Different-Subject (DS) Converb is marked by a suffix *–ge* as well; however, these two suffixes appear to be homonyms rather than one polysemous suffix. The converbal suffix *–ge* appears to be the Locative case suffix, with the DS Converb going back to a frozen Locative of the action nominal form of the verb. Furthermore, the DS Converb does not take person marking (Mas: 158), while the imperative forms do take suffixes that agree with the person and number of the subject. These, however, differ from the person suffixes found in the Indicative (Mas: 140).

The Present Imperative expresses commands, proposals, requests, and admonitions (81a, b), while the Future Imperative is used to express commands or requests that are to be fulfilled at a later point in time (81c). There is just one negative imperative paradigm (Mas: 140); this is formed with the regular negative prefix *el–*, although the second person forms generally take a special Prohibitive suffix *–le–* (81d) rather than the Imperative suffix *–ge–* found in the first and third persons (Mas: 176f).

(81a)

tit	*lebie*	*titt–in*	*qodo*	*omo*	*ta:t*	*a:–**ŋi–k***
2PL	land	2PL–DAT	how	good[ITR.3SG]	CA	make–PL–2

'Arrange your.PL land as it is good for you.PL'

(81b)

*pie-de-**ge**-n*
burn–DETR–IMP–3SG

'Let it burn!'

[Mas: 176]

(81c)

čemey–delle	*kel–**ge**–k*	*je*
finish–PF.CVB.SS	come–IMP–2	DP

'Finish.SG up and then come.SG!'

(81d)

*met–ul   el+kudede–**le**–k*
1SG–ACC   NEG+kill–PROH–2

'Do.SG not kill me!'

[Mas: 177]

Like Kolyma Yukaghir, Tundra Yukaghir has a Present Imperative in all three persons and two numbers (excluding 1SG), and a Future Imperative restricted to the second person singular and plural. As in Kolyma Yukaghir, the second person Present Imperative is basically unmarked, with the person agreement suffix *–k* (and the plural marker *–ŋu* for 2PL) attaching to the verb root. The imperative suffix found in the Present Imperative for the 1PL and third persons is *–ha–*; in the Future Imperative, however, it is *–hane–* (Maslova 2003a: 18). Interestingly, the Future Imperative is formed with the regular future marker *–t(e)* preceding the imperative suffix, e.g. *lew–**te–hane**–k* [eat–FUT–IMP–2] 'eat later!'. The Future Imperative is often used to express actions that are to be carried out after some other action (Maslova 2003a: 21f).

In Nivkh, person agreement is found only in the Imperative mood; in the Indicative, verbs do not agree with their subjects in person, and only optionally agree in number (Grz: 60). Nivkh has only one imperative paradigm that conveys immediate commands and requests (Grz: 64ff). This shows agreement with three persons and two numbers; a previous form for the first person dual is practically not used anymore (Grz: 64, 66). Imperatives are negated by a special prohibitive particle, or by incorporation of a negative verb into the imperative verb form (Grz: 68). In addition, there is a Preventive mood restricted to the second person which expresses warnings (Grz: 69ff).

In the Chaplino dialect of Central Siberian Yupik, imperative meanings can be expressed by the Imperative and the Optative mood (Vaxtin 2001b: 129). The Imperative paradigm has different imperative markers for the three persons, and additional agreement markers for subject (for intransitive verbs) and subject and object (for transitive verbs) (Vaxtin 2001b: 131). The Imperative expresses direct commands, requests, permissions, and offers (Vaxtin 2001b: 134f). The Optative mood can occur only with intransitive verbs, where it agrees with the subject in person and number (Vaxtin 2001b: 135). Although the Imperative as such does not make a tense distinction, having a basic prescriptive meaning of immediate fulfillment of the command or request, the meaning of the Optative focusses more on the wish of the speaker and allows for a certain delay of the action after the moment of speech (Vaxtin 2001b: 136):

"In AE [Asiatic Eskimo] this distinction [between urging the listener to act immediately and allowing for/prescribing an interval between the moment of speech and the action – B.P.] is partly realized in the opposition between the meanings of imperative and optative verb forms. However, within each single system of verb forms this distinction is unmarked." (Vaxtin 2001b: 132)

Although it is therefore possible to make a distinction between a Present Imperative and a Future Imperative in Asiatic Eskimo, this is restricted to intransitive verbs, since there are no optative forms for transitive verbs. The expression of commands or requests with transitive verbs is therefore restricted to the Imperative mood, so that such commands/requests automatically carry a prescriptive nuance of immediate action.

In Chukchi, the imperative is expressed by the Intentional mood and seems to be restricted to the second person. In the first and third person the Intentional mood is used to express hypothetical or intended actions (Dnn: 188f). There is no mention of any temporal distinction in the imperative meaning.

The Itelmen Imperative is characterized by a full paradigm with a separate set of subject agreement affixes. It can be formed both from the Perfective and Imperfective Aspect; the latter gives an ingressive or durative meaning, e.g. *brawo xn–anseʔ–qzu–z–en* [good IMP.3–study–IPFV–PRS–3SG] 'he must always study well!' (G/V: 157). The Imperative paradigm exists with both the (zero-marked) preterite and the present tense marker, but not with the future tense. The difference in meaning between the Preterite and Present Imperative seems to be that the former conveys a temporally unspecified command, while the latter refers specifically to actions that are to be performed at the time of speech, e.g *q'–nu–s–xç* [IMP.2–eat–PRS–IMP.2SG] 'eat now!' (G/V: 155ff).

In the Obdorsk dialect of Northern Khanty, the Imperative is restricted to the second person, with separate sets of suffixes marking agreement with the subject in number and, for transitive verbs, with the object in number, as well. There is no tense distinction (Nik: 26). In the northern Sos'va dialect of Mansi, the Imperative is also restricted to the second person, with number agreement for the subject and, in transitive verbs, object. No temporal distinction is made in the Imperative mood (Rmb: 126).

In Nganasan there is a full person-number imperative paradigm from a semantic perspective (i.e. suffixes marking exhortations, commands, or requests exist for all three persons and three numbers); however, formally the first person is distinct from the second and third person paradigm and is therefore treated as a separate mood (Ter: 212ff, 216f). The second and third person imperative forms are marked by an Imperative suffix and separate person-marking suffixes not found in the other moods. These agree with the person and number of the subject and, for

transitive verbs, with the number of the object. There is, in addition, a separate Future Imperative suffix used to express commands/requests that may be fulfilled later, or after a specified prior event, e.g. *tuy–**kuo**–ri* [come–DSTIMP–2DU] '(you two) come sometime' (Ter: 214). This suffix, which seems to be restricted to expressing future imperatives, combines with the second and third person in all three numbers (Ter: 214; Helimski 1998b: 503). This Future Imperative is lacking in Nenets and Enets (Ter: 214), as well as in Selkup (OSJ: 247f; Helimski 1998a: 566). The first person Hortative is marked by a different suffix from the second and third person Imperative marker; furthermore, in this mood the same person marking is used as in the Indicative. There does not seem to be a tense distinction in the first person Hortative (Ter: 216f).

In Ket the synthetic Imperative forms are restricted to the second person singular and dual (Dul: 452), while the Hortative forms of the first and third person singular and plural are formed analytically with the help of a particle *qan'* 'let' (Wer: 288). There are a large number of different forms for individual verbs built from different aspects, with transitive verbs agreeing with the object in person and number; however, there are no tense distinctions (Wer: 302ff).

For an overview of the presence or absence of a future imperative see Table 3.16.

Table 3.16: Presence of a future imperative in languages of Eurasia

Language	Present Imperative	Future Imperative
Turkic	full paradigm	---
Sakha	full paradigm	2SG/PL
Mongolic	full paradigm, several forms for 2SG/PL	---
Buryat	full paradigm, several forms for 2SG/PL	2SG/PL
Evenki	full paradigm	full paradigm
Ėven	full paradigm??	full paradigm
Udihe[*]	2SG/PL	---
Nanay[**]	full paradigm	2SG/PL
Yukaghir	full paradigm (no 1SG)	2SG/PL
Nivkh	full paradigm	---
Chukchi	2SG/PL	---
Itelmen	full paradigm	---
Eskimo	full paradigm	(Optative for intransitive verbs)
Khanty	2SG/PL	---
Mansi	2SG/PL	---
Nganasan	2+3SG/PL; Hortative for 1SG/PL	2+3SG/PL
Ket	2SG/PL (synthetic forms)	---

[*]and Oroč
[**]and Ulča and Orok

As becomes clear from Table 3.16, there is a very strong areal bias in the distribution of the future imperative amongst Siberian languages: it is those languages that have been and/or still are in contact with Evenki and Ėven that make a formal distinction between a present imperative and a future imperative. It is possible that this areal influence stems from the Northern Tungusic languages, since a future imperative is found in all the Northern Tungusic languages as well as in a number of languages from the Amur Tungusic branch which are closely related to each other. On the other hand, Nganasan is the only Samoyedic language with a future imperative, and Sakha and Dolgan are the only Turkic languages that have this category. Buryat also seems to be the only Mongolic language with an imperative marker with a clear future imperative meaning as opposed to a meaning of greater politeness. This areal distribution attains even more weight in a world-wide comparison of languages with a future imperative made available to me by Ewa Schalley, based on a typological sample of 408 languages:
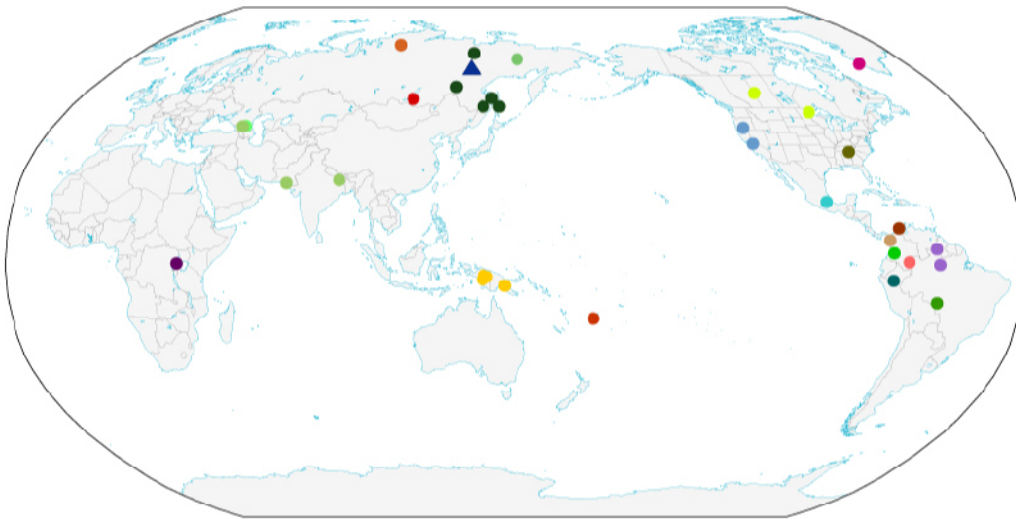
Figure 3.1: World map showing the distribution of languages that express a meaning of future imperative (based on Ewa Schalley's data and a sample of 408 languages). In Siberia, genera are coloured by different colours (dark green for Tungusic, red for Mongolic, dark blue for Turkic (triangle = Sakha), orange for Samoyedic, and light green for Yukaghir); in the rest of the world, different colours represent different families. This shows that the presence of a future imperative is a clearly areal, and not genealogical, feature in Siberian and South American/Amazonian languages.

Mapped here are languages that have some verbal form (not necessarily a morphologically dedicated Future Imperative) that expresses predominantly a future imperative meaning in all verbs; i.e. languages such as Siberian Eskimo, where the Optative is restricted to intransitive verbs, are excluded. This distribution shows that Siberia and South America/Amazonia are two major areas of languages that have a future imperative, a feature which is quite rare worldwide, being found in a frequency of only 10-15% (Ewa Schalley, pers.comm.; Gusev 2005: 62). This areal distribution is very interesting and important in the light of possible contact influence in the development of the Sakha Future Imperative.

3.3.6 The origins of the Sakha Future Imperative suffix –$A{:}r$

3.3.6.1 Possible copy from Mongolic

Poppe (1959: 682) and Sanžeev (1964: 99) suggest that the Sakha Future Imperative suffix might have been copied from Mongolic, a suggestion that was first made by Böhtlingk:

> "The Turkish-Tatar languages do not have a corresponding form [Future Imperative, B.P.] in the second and third person[8]; but the Mongolian colloquial language and Kalmyk have a second person Future Imperative that even strikingly matches the Yakut [form] …." (Btl: 304; translation mine).[9]

However, there are several issues that cast some doubt on this suggestion. Firstly, it is important to note the difference in meaning between the Sakha Future Imperative and the Mongolic Prescriptive: the Sakha form predominantly expresses a command to be fulfilled at a later point in time, while judging from the more recent descriptions, the Mongolic Prescriptive conveys primarily a higher degree of politeness than the Imperative and Precative forms (e.g. Kuzmenkov 2001). Buryat is a notable exception to this, since here the Prescriptive clearly has a future imperative meaning; however, Buryats have been and are in close contact with

---

[8] Böhtlingk (Btl: 302) gives a Future Imperative form for the first person singular and third person as well as the second person; the third person form is also given by Ubrjatova (1972: 585). It is based on the Future Participle –$IA\chi$ and the Present Imperative 3rd person agreement suffix –TIn. However, Böhtlingk's 1SG Future Imperative is in actual fact the 1SG Present Imperative form; and the third person form he and Ubrjatova give is part of a rarely-used separate paradigm with 'imperative-optative' meaning (GSJa: 324).

[9] Original: "Die türkisch-tatarischen Sprachen haben keine entsprechende Form in der 2ten und 3ten Person; aber die mongolische Volkssprache und das Kalmückische besitzen eine zweite Person Imperativi Futuri, die sogar in der Endung auffallend zum Jakutischen stimmt: [...]"

speakers of Evenki. The shift of meaning in Buryat can therefore well be due to contact influence, a fact further supported by the distinctly areal distribution of the future imperative in Siberian languages, as discussed in section 3.3.5.

As to the formal similarity between the Sakha Future Imperative suffix *–A:r* and the Mongolic Prescriptive *–A:rAy*, this is not so great as it would seem. There exists an emphatic form of the Future Imperative in Sakha (cf. example 67e, repeated here as 82) which ends in an (optionally long) high vowel; it is this form that is compared to the Mongolic Prescriptive (Btl: 304; Poppe 1959).

(82)

sibe:s	*tuhunan*	*kepse:ri:y*
sibe:s	tus–(t)InAn	kepse:–A:r–I:y
communications	side–INS.3SG	tell–DSTIMP[2SG]–EMPH

'Tell about the (phone) connection.'

[LukP188]

However, this emphatic form is (currently) not very widespread, but is found predominantly in the Suntar dialect[10] (own observation); Ubrjatova states that this is a dialectalism that has entered the standard language, without giving precise information about which dialect it may have come from (Ubr: 180). Furthermore, this emphatic form exists for the Present Imperative as well:

(83)

a:kkїn	*eti:y*
a:t–GIN	et–I:y
name–ACC.2SG	say[PRXIMP.2SG]–EMPH

'Say your name.'

[LukP3]

It might possibly go back to the content-question marking Interrogative Clitic *=(I)y* (cf. Btl: 302), and it is interesting to note that one speaker in the Suntar district used this clitic in discourse as an emphatic marker as well, e.g.

---

[10] Although the two examples presented here come from a narrative recorded not in the Suntar, but in the Olenëk district, they were both said by a young woman working for the village administration. I do not know anything about her origins, but it is possible that she was not a native of the Olenëk district.

(84)

itinnikteriy	*iti*	*ülüger*	*tïra:nspar*	*ülüger*	*biligin*	*bu*
itinnik–LAr=(I)y	iti	ülüger	tïra:nspar	ülüger	biligin	bu
such.a–PL=Q	this	INTS	transport	INTS	now	this

keteχ	*χa:lbït*
keteχ	χa:l–BIt
private	RES–PSTPT

'All of these, so much of the transport now became private possession.'

[IvaP232]

Since there is an emphatic variant of the Present Imperative as well as of the Future Imperative, and since a very similar morpheme is used as an occasional emphatic marker in discourse, it is possible to analyze the emphatic Future Imperative as consisting of the Future Imperative suffix –*A:r* and an emphatic marker –*I:y*, as I have done in the glosses. In this context it is important to note that Dolgan appears to lack the emphatic variant of the Present Imperative[11] (Ubr: 180). This provides further support for the analysis of the Future Imperative suffix as being separable from the emphatic suffix. These facts considerably weaken the case for copying of the Sakha Future Imperative from Mongolic, since the assumption here is that the emphatic Future Imperative –*A:rIy* was the form that was copied.

Of course, it is possible that the modern disyllabic Mongolic form actually originated through fusion of two morphemes, and that at the time of the proposed copying the two morphemes were still distinguishable[12]. This possibility is strengthened by the reconstruction of the Prescriptive given by Poppe (1955: 254): he claims that it consisted of "… the Altaic suffix –**r* of verbal nouns added to the suffix –**γa*– of the secondary verbal stem. The final –*aị* is a particle, an interjection." Ramstedt ([1903] 1968: 63) also analyzes the modern Mongolic Prescriptive as originally consisting of a morpheme –*A:r* plus an emphatic ending in –*ai/–a:*. If this reconstruction is correct, it might explain why the Sakha form is monosyllabic: if at the time of Sakha-Mongolic contact the two morphemes were still segmentable to the hearers, the ancestors of the Sakha could have copied the primarily Prescriptive morpheme without the emphatic ending. On the other hand, Kałużyński suggests that possibly only the emphatic ending was copied: "It cannot be ruled out that –*ār–yi* as a whole or only the second component was borrowed from Mongolian or formed according to the Mongolian pattern." (Kałużyński 1962: 111f, translation mine[13]). However, the form reconstructed by Janhunen (2003d: 23)

---

[11] From Ubrjatova's discussion it is not clear whether Dolgan has the emphatic variant of the Future Imperative; in any case, no examples of this are given.

[12] I thank Juliette Blevins for pointing this out to me.

[13] Original: "Es ist nicht ausgeschlossen, dass –*ār-yi* als ganzes oder nur seine zweite Komponente aus dem Mongolischen entlehnt oder nach mongolischem Muster gebildet ist."

does not lend itself to the same interpretation; although he, too, reconstructs a heteromorphemic form *–xA–ra.(y)I*, here the *–r* belongs to the second morpheme, and it is highly unlikely that the ancestors of the Sakha would have copied what amounts to one and a half morphemes. Furthermore, since the Mongolic Prescriptive is restricted to only four modern languages and is not attested in Middle Mongolian, and in Written Mongolian occurs only under influence of spoken dialects (cf. section 3.3.3), it seems to be a fairly recent innovation. Such a recent innovation might not have been present yet in the Mongolic dialect that the ancestors of the Sakha came into contact with. In addition, all the modern languages that have the Prescriptive have a final vowel or diphthong (Khamnigan Mongol: *–Arie*, Buryat: *–(g)A:rAy*, Khalkha: *–(g)ArAy*, Oirat: *–(γ)Arä*); it is thus unclear whether the original form really consisted of a morpheme plus separate interjection/particle, as suggested by Poppe and Ramstedt, and whether the ancestors of the Sakha would have been familiar with an obviously segmentable morpheme or not.

Lastly, it should be pointed out that among the Sakha verbal forms the Mongolic Prescriptive is in fact identical to the Sakha third person singular Voluntative-Potential (85a). This 3SG form is a portmanteau morpheme expressing both person and mood; it differs from the Voluntative-Potential paradigm as a whole, which is marked by a suffix *–A:yA* and takes predicative person markers (85b). It is unclear whether this 3SG Voluntative-Potential has anything to do with the Mongolic Prescriptive (probably not); nevertheless, the existence of this form in the Sakha verbal inflectional paradigm further weakens the case for copying of the emphatic Future Imperative suffix *–A:rI:y* from Mongolic.

(85a)

Maša	*Suntar–tan*	*kel–**e:rey***
Maša	Suntar–(t)tAn	kel–A:rAy
M.	S.–ABL	come–VPOT.3SG

'Perhaps (hopefully) Masha will come from Suntar.'

[back translation Sun, Ver]

(85b)

bihigi	*telley**de:ye**bit*
bihigi	telley–LA:–A:yA–BIt
1PL	mushroom–VR–VPOT–1PL

'Perhaps (hopefully) we could look for mushrooms.'

[back translation Sun, Tat, Oln]

In conclusion, although copying from Mongolic cannot be ruled out, this is definitely not the only possible source available for the Sakha Future Imperative.

3.3.6.2 Grammaticalization of a former analytical imperative

Another source of the Future Imperative suffix in Sakha proposed by Böhtlingk (Btl: 303) is through the grammaticalization of an analytical imperative still found in some Turkic languages (Korkina 1970: 161; Nasilov et al. 2001: 192). This analytical imperative is found already in Old Turkic, where the conjunction of a converbal form of the main verb plus the auxiliary *kör* 'to.see' has an exhortative meaning, e.g. *yel–ü kör* [gallop–IPF.CVB see[IMP]] 'See to it that you ride fast!' (Erd: 524). According to Böhtlingk, such an analytical form could have been the source of the Sakha Future Imperative suffix through a process of lenition of the intervocalic velar and assimilation of the vowel of *kör*, e.g. *bïh–a kör* [cut–IPF.CVB see] → *bïhaɣar*, and later loss of the intervocalic voiced velar and lengthening of the vowel → *bïha:r*. In the modern-day language, *–k–* voices to *–g–* in intervocalic position, e.g. *tik* 'sew' → *tig–e–bin* [sew–IPF.CVB–PRED.1SG], *bügün* 'today' < *bu kün* 'this day', and *g* lenites to *ɣ* after a low vowel, e.g. *kuoska* 'cat' → *kuoska–ɣa* [cat–DAT] 'to the cat' (contrast with *ïal* 'family' → *ïalga* [family–DAT] 'to the family'), *aɣa–ɣïn* [father–ACC.2SG] 'your father.ACC' (contrast with *ïal–gïn* [family–ACC.2SG] 'your family.ACC'). Thus, the lenition of intervocalic *–k–* to *–ɣ–* following upon a low vowel in such a grammaticalization process is not implausible. Intervocalic *–ɣ–* is quite unstable, for instance in fast-speech phenomena, but also in the standard allomorph of the singular Future Tense markers, i.e. *–IAm* instead of *–IAɣIm* (1SG), *–IAŋ* instead of *–IAɣIŋ* (2SG), and *–IA* instead of *–IAɣA* (3SG). Of course, as noted already by Böhtlingk, this grammaticalization process cannot explain the form of the Future Imperative for verbs ending in a vowel, since in such verbs the Imperfective converb takes on the form of a long high vowel, e.g. *sana:* 'think' → *sanï:* [think.IPF.CVB], and a form such as *sanï: kör* would not lead to the formation of a long low vowel in the Future Imperative suffix. However, it is quite possible to assume analogical levelling of the forms in vowel-final verbs following the pattern of consonant-final verbs (cf. Btl: 303).

According to Ragimov, analytical imperatives such as *yelü kör* used to cover a fairly wide range of imperative meanings; nowadays, however, they have developed individual meanings in different languages. Thus, in Turkmen this analytical form has a meaning of request, in Tatar it is used to convey orders, the fulfillment of which is extremely important, and in Uzbek the command has an added nuance of evil or well-wishing (Ragimov 1966, cited in Korkina 1970: 161). It is therefore quite plausible to assume a similar development in Sakha, with the individual nuance of a future imperative taking over. However, even though the analytical imperative may exist in a number of Turkic languages, it is important to note that Sakha stands out in that the analytical form grammaticalized to a suffix, with a distant future meaning not found in the other Turkic languages.

3.3.6.3 Connection with the Sakha Purposive Converb

In his discussion of the Sakha Future Imperative, Böhtlingk (Btl: 303) also proposes that it represents an old participle, the accusative form of which (–*A:rI*) nowadays has supinal meaning, e.g.

(86)

tugu	*da*	*aha:bat*	*buola*	*ïksa:n*
tugu	da	aha:–BAt	buol–A	ïksa:–An
what.ACC	PTL	eat–PRSPT.NEG	AUX–IPF.CVB	hurry–PF.CVB

bartïm	*kördörö:rü*
bar–BIt–(I)m	kör–TAr–A:rI
go–PSTPT–POSS.1SG	look–CAUS–PURP

'Not having been able to eat anything (I) hurried in order to show myself (to him).'

[Afny137]

Formally, the Purposive Converb does look like the Accusative of the Future Imperative suffix, and semantically there is a plausible link between the two forms as well, since the Purposive Converb expresses an intention, the fulfillment of which lies in the future, just as the Future Imperative conveys a command, the fulfillment of which can wait until a later point in time. However, the possible origin of the Purposive Converb as an Accusative-marked participle, the Nominative form of which later developed a meaning of future imperative still does not offer any solution to the origin of the supposed participle itself.

3.3.7 The Sakha Future Imperative as contact-induced grammaticalization

As can be seen from the discussion so far, there are two competing possibilities for the origins of the Sakha Future Imperative: either a copy from Mongolic, or Evenki contact influence. Facts in favour of the Mongolic copy hypothesis are the similarity in form of the suffixes, and also the restriction of the Mongolic Prescriptive and Sakha Future Imperative to the second person, while according to Nedjalkov and Konstantinova (Ned: 262; Kon: 184), the Evenki Future Imperative occurs with all person-number combinations [though Bulatova & Grenoble give only second person forms (B/G: 37)]. Furthermore, proponents of the copying scenario (Sanžeev 1964: 98; Poppe 1955: 254) claim that the Prescriptive has a future imperative meaning – in this case, both the form and meaning of the Sakha Future Imperative would be similar to the Mongolic Prescriptive, strengthening the case for copying. Facts that argue against copying from Mongolic

are the lack of the Prescriptive in Middle Mongolian (which is the language spoken in the Mongol Empire and thus the language the Sakha ancestors are most likely to have been in contact with) and the weak formal match between the probably heteromorphemic Mongolic Prescriptive suffix and the Sakha Future Imperative (see section 6.1). Furthermore, more recent descriptions of Mongolic languages do not mention a distant future meaning of the Prescriptive, but stress its added degree of politeness (see the individual chapters in Janhunen 2003b). This would leave only the similarity in form as proof of copying. However, since the difference in meaning between a future imperative and a more polite imperative/request is not very big, this counter-argument may not hold much weight. Thus, the Evenki Future Imperative, for example, expresses not only a command the fulfilment of which can be delayed, but also less categorical commands and requests: "In many contexts [the two imperative forms] can be used interchangeably, but change of the [Present] imperative for the [Future Imperative], as a rule, produces a softening of the imperative's force." (Ned: 19).

On the other hand, Table 3.16 and Figure 3.1 clearly show that Evenki and Ėven lie at the core of a linguistic area of Siberian languages that make a distinction between a present and a future imperative. This areal distribution favours the Evenki contact influence hypothesis. As to the question why Sakha under Evenki influence developed the Future Imperative only in the second person, and not in all persons, when Evenki appears to have a Future Imperative for all person/number combinations, a possible explanation lies in the fact that the Evenki Future Imperative paradigm is split, with the second person being marked by a different suffix and by different subject agreement markers than the first and third persons. The Evenki Future Imperative paradigm might therefore not have been perceived as a full paradigm by Sakha speakers in contact with Evenki. Furthermore, the function of second person imperatives is different from that of first and third person forms, since the latter do not convey a direct command or request; first and third person imperatives are often not considered imperatives at all (e.g. Isxakov & Pal'mbax 1961: 394; van der Auwera et al. 2005: 294). In addition, there is dialectal variation in the number of persons for which the Evenki Future Imperative exists: thus, amongst the dialects spoken in or adjacent to Yakutia, in the Učur dialect (spoken south of the river Aldan) the Future Imperative is restricted to the second and third persons (Myreeva 1964: 48), and in the Tommot dialect it is restricted to the second person (Myreeva 1962: 76), while the Ajano-Maja and Tokko dialects have full Future Imperative paradigms (Romanova 1964: 106; 1962: 34). Furthermore, as mentioned in section 3.3.4.1.1, Bulatova & Grenoble (B/G: 37) do not give any first and third person forms of the Evenki Future Imperative. Lastly, there exists the

possibility that the Turkic analytical imperative, which may have formed the basis of the Sakha Future Imperative suffix, may have been restricted to the second person. In Old Turkic, no forms other than second person have been attested (Marcel Erdal, pers. comm.); in Bashkir, however, this analytical imperative exists also for the third person (GSBJ: 284). Thus, Evenki influence in the development of the Sakha Future Imperative is quite plausible.

There is, however, yet again a further complicating factor, and this is the fact that the Northern Tungusic languages differ from their Amur Tungusic relatives. As was discussed in section 3.3.4.1, the Future Imperative in Evenki, Ėven, and Negidal is formed with the Purposive Converb plus reflexive possessive person suffixes for the second person singular and plural. This is quite different from the Amur Tungusic languages, which either do not have a future imperative (e.g. Udihe), or have a future imperative formed with a dedicated suffix (e.g. Nanay) (cf. section 3.3.4.2). Thus, the Future Imperative in the Northern Tungusic languages might itself be due to contact influence. One possibility for this would be Yukaghir, which distinguishes between a Present Imperative and a Future Imperative. However, since the Yukaghir Future Imperative is formally consistent with the first and third person Present Imperative, i.e. all of these forms appear to constitute a single paradigm (cf. Mas: 140, Table 14), it would seem as if Yukaghir developed the distinction between a Present and Future Imperative relatively recently, possibly under contact influence, rather than this being an old category in the language. Another possible source of influence for the development of the Northern Tungusic Future Imperative could be Sakha, in which the Future Imperative suffix and the Purposive Converb resemble each other formally. This could have triggered the extension of the Purposive Converb to a Future Imperative marker in the Northern Tungusic languages. Under this scenario, Sakha would have copied the suffix –*A:r* from the Mongolic Prescriptive, with a meaning of a more polite and future imperative, from which it would have developed the Purposive Converb. Speakers of Northern Tungusic dialects, who would have heard the Sakha use a similar-sounding form to express purpose as well as commands that are to be fulfilled at a later point in time, would have extended their Purposive Converb to a Future Imperative. A similar case of extension involving the Purposive Converb appears to have taken place in the Mongolic language Dagur, which was historically spoken along the middle reaches of the Amur river in contact with Evenki. This language has developed an 'Indirect Imperative' that expresses "delayed action or politeness" consisting of the Purposive Converb plus possessive subject agreement markers (Tsu: 143f), clearly under Tungusic influence.

However, as shown by Malchukov (2001: 167), the extension of purposive converb to future imperative is a very plausible development (87a, b), making it unnecessary to postulate a contact-induced change in the development of the Northern Tungusic Future Imperative. This is further supported by the fact that cross-linguistically languages use both strategies to convey the same information, as mentioned in section 3.3.4.1.1. This language-internal development of the Purposive Converb to Future Imperative in the Northern Tungusic languages is further strengthened by the similar development observable in Nanay, where the 1SG Purposive Converb is often used as a 1SG Present Imperative (Avrorin 1961: 129).

(87a)

ama	*edu*	*tegeči–d–li*	*nokle–**de–y***
father	here	sit–PROG–PRXIMP.2SG	shoot–PURP–PREFL.SG

'Father, keep sitting here in order to shoot (afterwards).'

(87b)

*nokle–**de–y***
shoot–DSTIMP–PREFL.SG

'Shoot (afterwards)!'

[Malchukov 2001: 167]

Furthermore, since Nanay, Ulča and Orok also have a Future Imperative marked by a dedicated suffix, the category of future imperative appears to be characteristic of more than one branch of the Tungusic languages. It is therefore quite plausible that the Northern Tungusic languages replaced a previous dedicated future imperative suffix by the Purposive Converb through language-internal grammaticalization.

Taking together all the evidence, I believe that it is possible to make a case for Evenki contact influence in the development of the Sakha Future Imperative. Thus, I suggest that the distinction between a Present and a Future Imperative in Evenki triggered the grammaticalization of a Future Imperative out of the analytical imperative also found in other Turkic languages. This proposal provides an explanation for why the Turkic analytical imperative grammaticalized to a synthetic form only in Sakha: this was encouraged by the specific contact situation the Sakha speakers found themselves in, while a similar trigger was lacking for the other Turkic languages.

It is furthermore possible that the Sakha Purposive Converb, too, may have developed under Evenki contact influence, though at a later stage: Once the grammaticalization of the Turkic analytical imperative to the Future Imperative in Sakha had taken place, speakers of Sakha in contact with Evenks may have noticed the use of one form to express both future imperatives and purpose in Evenki. Under

this influence, they may have made use of the Future Imperative marker *–A:r* to develop their Purposive Converb *–A:rI*. This process can be schematically represented as in Figure 3.2.

Evenki influence in establishing
the new category of Future Imperative

analytical
imperative ⟶ *bïha r* ⟶ *–A r* as Future ⟶ *–A rI* (Accusative)
(*bïha kör*)   Imperative suffix   as Purposive
                                                      Converb suffix

Evenki influence (Future Imperative
suffix 2SG/2PL = Purposive Converb)
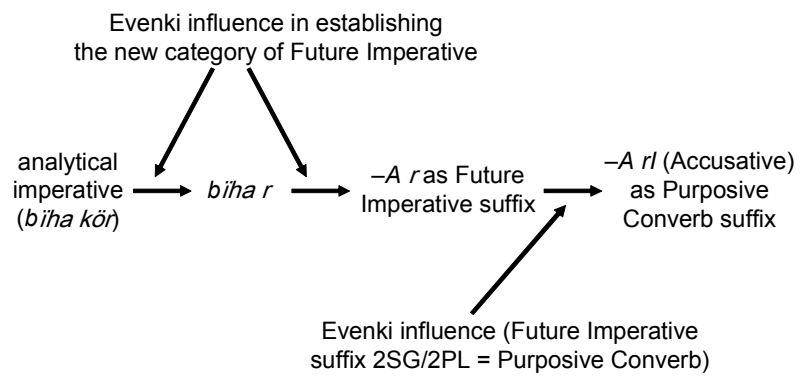
Figure 3.2: Schematic representation of the development of the Sakha Future Imperative and the role of Evenki contact influence in its development

3.4 Contact influence in the domain of possessive marking in Sakha

3.4.1 The 'non-possessive' use of possessive suffixes in Sakha

It is well known that the Turkic languages use the 3SG possessive suffix as a way of indicating referentiality, both situational and anaphoric (Johanson 1998a: 51; Fraurud 2001 *passim*). Grønbech ([1936/1979] 1997: 92ff) even refers to the Turkic 3SG possessive suffix as 'the article', and claims that its uses as a referentiality and definiteness marker were the initial ones, while the function of marking agreement with the possessor of a NP is a later derivation (Grønbech [1936/1979] 1997: 101). In Turkish, the 3SG possessive suffix has a very high occurrence on nearly 30% of over 3,000 nouns in a corpus of written non-fiction prose (Fraurud 2001: 251), making the assumption of an article-like function quite plausible. However, Fraurud (2001) convincingly shows that the 3SG possessive suffix has not yet grammaticalized into a definite article, neither in Turkish, nor in the Uralic languages.

In Sakha, too, 3SG possessive marking is used to signal situational identifiability on nouns referring to natural phenomena, e.g. *ńurguhun–**a** ol da buol–lar erde taχs–ar* [snowdrop–POSS.3SG that PTL AUX–COND early go.out–PRSPT] 'nevertheless, the snowdrops come out early' [Efmy803], *hayïn–**a** kïlgas* [summer–POSS.3SG short] 'the summer's short' [Efmy850]. However, in Sakha it is not only the 3SG possessive suffix that is used in 'non-possessive' functions, but other possessive suffixes as well, most especially the 2SG possessive suffix (88a-c), but also the 1SG possessive suffix (88d). These function mainly to establish anaphoric reference and to indicate that the entity in question will continue as the topic of discourse (88b, c), but they can also be used to mark contrastive focus (e.g. 88a) or just for emotional purposes, as shown here by the 1SG possessive marker (88d). In addition, they add a nuance of involving the addressee in what is being said. This is in stark contrast to other Turkic languages (with the exception of Dolgan, see below): in the corpus analysed by Fraurud, out of 3,492 nouns none carried second person possessive suffixes (Fraurud 2001: 251)[1].

---

[1] One of the reasons for this discrepancy may lie in the fact that my corpus consists of spoken narratives, while the corpus Fraurud analyzed consists of written non-fictional text, which would generally be expected to be much less emotional.

(88a)

otton	*sïlgïŋ*		*buollaɣïna*	*beyete*	*χahan*
otton	sïlgï–(I)ŋ		buollaɣïna	beye–(t)A	χas–An
CP	horse–POSS.2SG		however	self–POSS.3SG	dig–PF.CVB

ahï:r		*buolan…*
as–A:–Ar		buol–An
food–VR–PRSPT		AUX–PF.CVB

'But horses, however, since they eat by themselves by digging (for grass under the snow)…'

[XatR253]

(88b)

hiti	*hoɣotoχ*	*gïlabaŋ*	*barïtïn*
s(I)–iti	soɣotoχ	gïlaba–(I)ŋ	barï–(t)In
EMPH–this	single	headman–POSS.2SG	all–ACC.3SG

bas bille
bas_bil–TA
take.into.possession–PST.3SG

'Just this headman alone took everything into his possession.'

[IvaP56]

(88c)

ol	*emčitteriŋ*		*büteydi:*
ol	em–SIt–LAr–(I)ŋ		büteyde:–A
that	medicine–AGNR–PL–POSS.2SG		see.clearly–IPF.CVB

bileller	*ebit*
bil–Ar–LAr	e–BIt
know–PRSPT–PL	AUX–PSTPT

'Those healers are clairvoyants (lit. know how to see clearly), it seems.'

[Afny114]

(88d)

ol	*ölö*	*sïhan*	*baran*	*bili*
ol	öl–A	sïs–An	bar–An	bili
that	die–IPF.CVB	IRR–PF.CVB	SEQ–PF.CVB	that.one

χotonum	*mahïgar…*
χoton–(I)m	mas–(t)IgAr
cattle.shed–POSS.1SG	wood–DAT.3SG

'After having nearly died (I worked) in the cattle-shed's wood…'

[PotP148]

In none of the examples (88a) to (88c) is there a real possessor. The horses in (88a) were mentioned contrastively to cattle, and since I was the only hearer present, it is clear that *sïlgïŋ* cannot mean 'your horses', as I have never owned a horse in my life.

Similarly, the headman referred to in (88b) was not my headman, nor was the person referred to a headman anymore at the time the narrative was recorded; he does, however, feature quite extensively in the preceding sentences as having given away state farm property to his close relatives for nothing. Thus it is clear that here the 2SG possessive suffix functions only as an emphatic marker of anaphoric reference and does not refer to a real possessor. The same holds for the healers in (88c), who had no relationship to any of the people present, but had been referred to in the preceding sentence and were mentioned in the further discourse as well. As to (88d), the cattle-shed the speaker worked in was not his personal cattle-shed, but the cattle-shed belonging to the state farm; it was not mentioned before, nor after, and the use of the 1SG possessive suffix may have been for purely emotional reasons, since he had just passed the climax of his tale of how he nearly starved during the war.

A similar use of the second person possessive suffixes (mainly 2SG, but also one example of 2PL) to mark the identifiability of the referent was noted by Stachowski in Dolgan (Stachowski 1998), e.g. *ha:s küöl–**üŋ** arï–ll–an balïk tut–ta–lar* [spring lake–POSS.2SG open–PASS–PF.CVB fish hold–PST–PL] 'In the spring, when the lake (not 'your lake') became free of ice, they caught fish' (Stachowski 1998: 127). Since this use of the second person possessive is not found in other Turkic languages[2], Stachowski (1998: 128) suggests that it provides evidence of Nganasan contact influence in Dolgan.

The 'non-possessive' use of possessive markers other than 3SG is common in Nganasan, as in other Uralic languages (Nikolaeva 2003). Thus, Tereščenko writes: "Like in Nenets, in Nganasan the personal-possessive suffixes of the second person function in a number of cases as a kind of distinctive definite article, underlining that the noun they mark denotes that same person or that same object which was mentioned before and which will be mentioned further." (Ter: 95, translation mine[3]), e.g.:

---

[2] In Bashkir the 2SG suffix may be used to express delight and admiration (GSBJ: 124); however, from the examples it appears that the possessively-marked noun has some form of relationship to the addressee, e.g. 'how your boy does dance!', or 'your old man is fighting well, I said to the teacher'. Furthermore, the three examples given of this usage all refer to humans. Thus, this appears to be different from the use of the 2SG possessive suffix in Sakha and Dolgan.

[3] Original: "Как и в ненецком языке, в нганасанском лично-притяжательные суффиксы 2-го лица в ряде случаев выступают в виде своеобразного определенного артикля, подчеркивая, что оформленное ими существительное обозначает то самое лицо или тот самый предмет, о котором шла речь и о котором будет говориться в дальнейшем."

(89)

ba:rbəδuŋ	*xontï*	*koptua //*	*koptua–rə*	*t'etuamï*	*ńeńaŋku*
master.GEN	is	girl //	girl–2SG	very	pretty

'The master has a daughter. // This daughter is very pretty[4].'

[Ter: 95; transkription following Nikolaeva 2003: 138]

As shown by Nikolaeva (2003), the third person possessive suffixes in Uralic languages function to mark identifiability, both situational and anaphoric, similar to what is found in Turkic languages. The first and second person possessive suffixes, however, have pragmatic functions. The first person possessive suffix establishes "a kind of a close pragmatic relation between the speaker and the addressee" (Nikolaeva 2003: 136), while the second person possessive suffix indicates "that the speaker somehow pragmatically associates the listener and the referent of the corresponding noun" (Nikolaeva 2003: 136), creating an added emotional effect for the addressee. These suffixes can replace the 3SG possessive suffix functioning as an identifiability marker (e.g. 90a from Nenets, similar to example 88c above), and can replace the normal possessive suffix in its possessive function (e.g. 90b from Khanty).

(90a) Nenets:

puxacyako yilyewi°//	*puxacyako–nt°*	*syidya*	*søwa*	*nyu–da*
old.woman live.NARR.3SG //	old.woman–GEN.2SG	two	nice	son–POSS.3SG

'An old woman is living. // This old woman has two nice sons.'

[Nikolaeva 2003: 137f]

(90b) Khanty:

wŭl–li	*yis*	*uže*	*aś–en*	*pălat–ti*	*wanaməs*
big–TRNS	became.3SG	already	father–POSS.2SG	length–TRNS	reached.3SG

'He has become big already; he has reached the size of his father.'

[Nikolaeva 2003: 138]

Given the widespread use of second person possessive suffixes in non-possessive function in the Uralic languages, and the lack of such a function of second person possessive suffixes in Turkic, it might be plausible to suspect Uralic contact influence in Sakha, as was suggested for Dolgan by Stachowski. This is all the more plausible when the areal distribution of this phenomenon in Sakha is taken into account: the 2SG possessive suffix with a non-possessive function was used

---

[4] In the translation and interpretation of this example I follow Tereščenko, who claims that the 2SG possessive suffix functions as a marker of anaphoric reference. Nikolaeva (2003), however, interprets this as an example of the 2SG possessive marker replacing the 'normal' possessive marker and gives a corresponding translation: '<u>his</u> daughter is very pretty' (emphasis mine).

only once by one of the three speakers recorded in the Verxojansk district; in the Taatta district, three of the four speakers used it in this way, but only very rarely (one speaker showed two instances of this, the other two only one each), while in the Suntar and Olenëk districts all the speakers used 2SG possessive marking as a marker of identifiability and to establish an emotional link with the addressee. This feature was used most extensively in the Suntar district, especially by one speaker. We thus find a very skewed distribution of this phenomenon, with a heavy bias in the west and northwest and a complete lack in the northeast. Although at least in historical times no direct contact with speakers of Samoyedic languages has been documented, the west and northwest of Yakutia are regions that might most plausibly be expected to show some Uralic/Samoyedic substrate influence, since these are regions closest to the current and historical distribution of Uralic languages.

However, my main consultant in the Verxojansk district had no problems whatsoever in accepting and translating the sentences that showed this use of the 2SG possessive. She furthermore insisted on this being used just as much in her native Verxojansk dialect, and suggested that its lack in the narratives I had recorded in her native village of Tabalaax was due to a certain diffidence of the speakers, claiming that the feature was linked to how emotionally involved the speaker was. Although there does appear to be a certain link between emotional involvedness in the narrative and pragmatic use of the 2SG possessive suffix (thus, the speaker in the Suntar district who showed such a widespread use of this feature was very emotional, very indignant about the malpractises of former and current politicians), this cannot explain it entirely: on the one hand, there were a number of speakers in the Suntar and Olenëk districts who were not particularly emotionally involved in their narrative, and yet they used the 2SG possessive marker in pragmatic rather than possessive function; on the other hand, one of the speakers in Tabalaax was quite involved in his story without making use of this feature. However, it appears that the speech of the Suntar district may be more emotional in general. Thus, in this district I noticed very widespread use of the emphatic demonstratives *hu–bu* [EMPH–this], *h–iti* [EMPH–this.(intermediate distance)], *h–ol* [EMPH–that], and the adverb *hi–biligin* [EMPH–now], as well as derivations and inflections thereof, to the extent that they seem to have replaced the standard forms (cf. example 88b). Furthermore, the emphatic Imperative variant (ending in a long high vowel, cf. section 3.3.6.1) is widespread here and considered a characteristic feature of the Suntar/Vilyuy dialect by speakers of the Verxojansk and Taatta dialects. Thus, the more extensive use of the emphatic/emotional use of the 2SG possessive marker fits the general tone of Suntar speech.

However that may be, my consultant's judgement does indicate further that the phenomenon in question is not restricted to the dialects where I happened to record it, but that it is common enough in the northeast for her not to classify it as a foreign dialectal variant. Thus, if one wanted to postulate Nganasan or other Samoyedic contact influence, one would have to explain the spread of this feature beyond the plausible Samoyedic-Sakha substrate zone. A different explanation for this phenomenon might be contact influence from Tungusic languages, in which possessive marking also has pragmatic functions. However, here the possessed noun is additionally marked by the alienable possession suffix, as will be discussed in more detail below.

3.4.2 Uses of the Tungusic possessive suffixes in comparison to Sakha possessive marking

3.4.2.1 The alienable possession suffix in the Tungusic languages

The Tungusic languages stand out amongst the languages of Siberia in that they have a formally marked category of binary possessive classification (Comrie 1981: 79; Nichols & Bickel 2005: map 59[5]). Thus, judging from the information in *Jazyki narodov SSSR* (JaN5: 68-232) practically all the Northern Tungusic and Amur Tungusic languages have the Alienable Possessive suffix *–ŋV*. Although no category of 'indirect' alienable possession is mentioned in the brief sketch of Negidal given by Kolesnikova & Konstantinova (JaN5: 112-115), Cincius (1982: 20) provides evidence that in Negidal, too, the Alienable suffix *–ŋ(i)* is used to mark alienable body parts, e.g. *del–i–**ŋi**–nin* [head–E–ALN–POSS.3SG] 'his head (of an animal that he killed)'. There appear to be two exceptions amongst the Tungusic languages: Oroqen, for which Whaley et al. (1999: 301) explicitly state that it lacks a distinction between inalienable and alienable possession, and Manchu, which lacks a morphological category of possessive marking (Gor: 399). However, Manchu appears to have derivatives of the common Tungusic Alienable Possession suffix *–ŋV* in the 'Substantive-Possessive' suffix *–ŋge* and a suffix deriving quality adjectives *–ŋga/–ŋge/–ŋgo* (Gor: 152f). These, however, synchronically derive adjectives from nouns.

---

[5] It should be noted that in map 59 Nichols & Bickel show Nanay as having a system of binary possessive classification, but Evenki as not having it. This is due to a misinterpretation on their part of the grammars they consulted (Johanna Nichols, pers. comm.), since both Evenki and Nanay, like the other Tungusic languages, make the same distinction using a cognate suffix, as discussed in this section.

In Evenki, kinship terms are almost always marked for possessor, except when they are used vocatively. Body part terms, too, almost always take possessive-marking suffixes; inalienable body part terms are used with possessive suffixes (91a), while alienably possessed body parts (e.g. the parts of an animal which I killed and which are therefore my property, even though not part of my own body) are marked by the Alienable Possession suffix –*ŋ(i)* (91b). It is not possible in Evenki to mark kinship terms with this suffix. Nouns referring to natural phenomena, such as mountains, rivers, lakes, birds, and plants can also take possessive marking when they are understood to be important participants in the narrative event (91c) (Ned: 124f, 145). In this case they are additionally marked by the Alienable Possession suffix.

(91a)

minŋi	**halgan–mi**	*enu–jere–n*
my	leg–POSS.1SG	ache–PRS–3SG

'My leg aches.'

(91b)

minŋi	**halgan–ŋi–v**	*tadu*	*bi–si–n*
my	leg–ALN–POSS.1SG	there	be–PRS–3SG

'My leg (i.e. of an animal which I killed and cut off) lies there.'

[Ned: 144]

(91c)

tar	**munnukan–ŋi–v**	*tuksa–malča–ra–n*
that	hare–ALN–POSS.1SG	run–quickly–NFUT–3SG

'That hare (which I saw or tried to chase) ran away quickly.'

[Ned: 125]

In Ėven, the Alienable Possession marker –*ŋ*– is inserted before the possessive suffix if the possessive relation is "regarded as temporary or conventional" (Mal: 9). Interestingly, this alienabilizing suffix can be attached to kinship terms as well, in contrast to Evenki (Mal: 9; NovI: 145). Thus, in Ėven *min eńi–ŋ–u* [1SG.OBL mother–ALN–POSS.1SG] 'my mother' has the meaning of 'non-biological mother, woman who is considered to be my mother' (NovI: 141), while *hute–ŋe–n* [child–ALN–POSS.3SG] 'his/her child[6]' refers to a child that is not

---

[6] Malchukov contrasts *huteŋen* with *hut–u* [child–POSS.1SG] 'my (own) child', and glosses it with a 1SG possessive suffix, saying: "... someone who is considered to be **my** child." (Mal: 9, emphasis mine). However, the form *huteŋen* is clearly marked for 3SG possessor, since the 1SG possessive suffixes are –*β/–u*, –*mu*, or –*bu*, while the suffix –*n* is the 3SG possessive suffix (Mal: 9), as I have glossed it. The 1SG form would be *hute–ŋ–u* (Andrej Malčukov, pers. comm.).

someone's biological offspring, but is considered his or her child (Mal: 9). In general, however, kinship terms are marked with the plain possessive suffix, without the Alienable Possession suffix (NovI: 146).

In Udihe, an Amur Tungusic language, kinship terms are mostly marked with possessive suffixes "[…] unless used in generic contexts." (N/T: 127). As in all the languages discussed so far, the same possessive constructions are used for inalienable possession (i.e. with body part and kinship terms) and for prototypical possession of objects, e.g. *in'əi igi–ni* [dog tail–POSS.3SG] 'dog's tail', *mama təgə–ni* [grandmother gown–POSS.3SG] 'grandmother's gown' (N/T: 127). As in Evenki, the Alienable Possession suffix *–ŋi* is used with body part terms when they refer to the parts of an animal that was in some way obtained by the current possessor (92a). Furthermore, the nouns *na:* 'land', *mo:* 'tree', *ze:* 'money' and *ya:* 'cow' obligatorily take Alienable Possessive suffixes when denoting a temporary possession, e.g. when the land is somebody's property. Nouns that are considered important to the speech situation can also take possessive marking with the addition of the Alienable Possession suffix, similar to Evenki (92b) (N/T: 135ff).

(92a)

bi	**imo:–ŋi–wə–i**	*təu*	*diga:–i*	*si*	*susa:–i*
1SG	fat–ALN–ACC–2SG	all	eat.PST–2SG	2SG	run.PST–2SG

'You have eaten my fat and run away.'

[N/T: 136]

(92b)

bi	*oloxi–wə*	*me:usa–la:–mi*	*ə–si–mi–də*	*wa:*
1SG	squirrel–ACC	gun–VR.PST–1SG	NEG–PST–1SG–FOC	kill

oloxi–ŋi:	*sus'a*
squirrel–ALN.1/2SG	escape.PRF

'I shot at the squirrel but didn't get it. My squirrel escaped.'

[N/T:138]

As a rule, kinship terms do not combine with the Alienable Possessive suffix. An exception is made by the words *b'ata* 'boy/son' and *aziga* 'girl/daughter'; here, it is the Alienable suffix that expresses the meaning of progeny over that of young human (N/T: 139f):

(93)

bi	*zə–i*	**b'ata–ŋi–ni**	*yəxə–ini*
1SG	neighbour–1SG	boy–ALN–3SG	sing–3SG

'My neighbour's son is singing.'

[N/T: 140]

As discussed in section 3.2.1.3, Nikolaeva & Tolskaya analyze the suffix –ŋi that marks predicative possession as being identical to the Alienable Possession suffix (N/T: 141). However, given the fact that predicative possession is marked on the possessor, while alienable possession is marked on the possessum, and given the very different syntactic and semantic functions of the two constructions, I would rather consider them separate, homonymous suffixes.

3.4.2.2 Pragmatic uses of the Tungusic and Sakha possessive suffixes[7]

As mentioned in section 3.4.2.1, in Evenki and Udihe nouns that stand in some form of relationship to the contextually salient person in the speech event can be marked by the possessive suffixes with addition of the Alienable Possession suffix. Nikolaeva & Tolskaya describe this as follows:

> "In this case a situational non-possessive relation is presupposed between the 'possessor' and the head noun. The relationship can be established between virtually any two entities associated in the discourse. The exact form of this relation is not expressed overtly, and therefore the construction itself is semantically undetermined." (N/T: 138)

Examples from Evenki (91c repeated here as 94a; 94b below) and from Udihe (92b, repeated here as 94c; 94d below) show how the possessive plus Alienable suffixes function to stress the relationship between the 'possessor' and the 'possessed' "which is realized through [the possessor's] activity" (N/T: 137). It should be noted here that both Nedjalkov (Ned: 125) and Nikolaeva & Tolskaya (N/T: 137f) analyze this marking of salient discourse participants as a sole function of the Alienable Possession suffix. However, the relationship between the discourse participants is in actual fact expressed by the possessive marking, which agrees in person and number with the subject of the sentence, while the Alienable Possession suffix appears to stress the saliency of the 'possessed' entity.

(94a) Evenki:

tar	*munnukan–**ŋi–v***	*tuksa–malča–ra–n*
that	hare–ALN–POSS.1SG	run–quickly–NFUT–3SG

'That hare (which I saw or tried to chase) ran away quickly.'

---

[7] I thank Markus Lang for very fruitful discussion on this topic and for pointing out to me that it is the possessive suffix rather than the alienable suffix that establishes the relation between the discourse participants.

(94b)

asatkan	tatkit–tula	avdanna–**ŋi–l–vi**	emev–re–n
girl	school–LOC	leaf–ALN–PL–PREFL	bring–NFUT–3SG

'The girl brought the leaves (which she had gathered) to school.'

[Ned: 125]

(94c) Udihe:

bi	oloxi–wə	me:usa–la:–mi	ə–si–mi–də	wa:
1SG	squirrel–ACC	gun–VR.PST–1SG	NEG–PST–1SG–FOC	kill

oloxi–**ŋi:**	sus'a
squirrel–ALN.1/2SG	escape.PRF

'I shot at the squirrel but didn't get it. My squirrel escaped.'

(94d)

bi	tuduzə–**ŋi:**	sagde–ŋku
1SG	potato–ALN.1/2SG	big–PL

'My potatoes (those cultivated by me) are big.'

[N/T:138]

This possessive marking of situationally and contextually salient referents is found in Sakha as well, for example in translations of modified versions of Dahl's texts to elicit aspectual distinctions (Dahl 1985: 205, text B1 – 95). Here, the function of marking the relation between a highly salient participant in the discourse and an entity that plays an important role, but is not actually possessed, is performed by the normal possessive suffixes which agree in person and number with the main discourse participant.

(95)

beɣehe	min	oyu:rga	hïrïttïm	
beɣehe	min	oyu:r–GA	sïrït–TI–(I)m	
yesterday	1SG	woods–DAT	walk–PST–POSS.1SG	

aray	tuoχ	ere	hïp~hïmnaɣahï	ükte:tim
aray	tuoχ	ere	sïp~sïmnaɣas–(n)I	ükte:–TI–(I)m
suddenly	what	PTL	INTS~soft–ACC	step.on–PST–POSS.1SG

körbütüm		kutuyaχ	ebit
kör–BIt–(I)m		kutuyaχ	e–BIt
see–PSTPT–POSS.1SG		mouse	AUX–PSTPT

kuttanammïn		oχton	tüstüm
kuttan–An–BIn		oɣus–An	tüs–TI–(I)m
be.afraid–PF.CVB–PRED.1SG		fall–PF.CVB	descend–PST–POSS.1SG

kutuyaɣïm	kuotan	χa:lla
kutuyaχ–(I)m	kuot–An	χa:l–TA
mouse–POSS.1SG	flee–PF.CVB	RES–PST.3SG

'Yesterday I was walking through the woods. Suddenly I stepped onto something very soft. I looked and saw it was a mouse. I got a fright and fell down. The mouse (lit. my mouse) ran away.'

<div align="right">[translation, Sun]</div>

This little story was translated in basically the same way by nearly all my consultants in the Suntar, Olenëk and Taatta districts. In all cases the mouse was unmarked on its first appearance in the narrative, and all consultants added possessive suffixes agreeing with the main character of the narrative in person and number when describing that the mouse ran away. That is, in the version of the text (Dahl 1985: 205, text B3) where the first person singular narrator goes for a walk in the woods together with his/her brother, the noun referring to the fleeing mouse was marked for first plural possessor (once a variant with third singular possessor was offered, since it was the brother who stepped on the mouse, so 'his mouse' ran away), while in the version of the text where the brother goes for a walk by himself and the narrative is in the third person singular (Dahl 1985: 205, text B4), the noun referring to the mouse was marked with corresponding 3SG possessive suffixes. There was one exception to this rule in the Olenëk district, where one consultant used possessive marking, while another didn't; and in the Verxojansk district two consultants did not refer to the fleeing mouse with possessive marking, either[8]. Example (96) comes from a narrative about a group of bandits (the 'possessors') who are tricked by an old man (the 'possessum') into following him into a trap.

(96)

inńe	*die:čči*	*kihilere*		*hïrga*	*kölümmütterin*
inńe	die–A:ččI	kihi–LArA		sïarγa	kölüy–(I)n–BIt–LArIn
thus	say–HAB	man–POSS.3PL		sled	harness–REFL–PSTPT–ACC.3PL

kenne		*ol*	*oγonńordoro*
kelin–(t)A		ol	oγonńor–LArA
back.part–POSS.3SG		that	old.man–POSS.3PL

mi:nnerinen		*kelbit*
mi:n–ner–(I)n–An		kel–BIt
ride–CAUS–REFL–PF.CVB		come–PSTPT

'Saying this the old man, after they harnessed their sled, the old man came riding.'

<div align="right">[Afńy161]</div>

---

[8] This is in accordance with the lack of non-possessively used possessive suffixes in the narratives recorded in this district.

Furthermore, in Udihe use of the possessive suffixes in conjunction with the Alienable Possession suffix with proper names (97a) and nouns referring to humans (97b) presupposes that the person concerned is well known to both the speaker and the listener. Thus, the 'Sergey' referred to in (97a) "may be a relative or friend of the interlocutor", while the 'girl' in (97b) was the linguist come to work with her consultant; the sentence was said by the consultant's husband (N/T: 139). In Sakha, this function appears to be fulfilled by second person singular possessive marking (98a-c):

(97a)

si	*Sergey–**ŋi:***	*ə–s'ə*	*mamasa–la*
2SG	S.–ALN.1/2SG	NEG–PRF	wife–VR

'Hasn't your Sergey married?'

(97b)

si	*aziga–**ŋi:***	*əmə:–ni*
2SG	girl–ALN.1/2SG	come.PST–3SG

'Your girl has come.'

[N/T: 139]

(98a)

büteren	*bili*	*tüöttü:*	*hïl*	*a:har*	*bili*
büt–(I)Ar–An	bili	tüört–LI:	sïl	a:s–Ar	bili
end–CAUS–PF.CVB	that.one	four–DIST	year	pass–PRSPT	that.one

*Popova**ŋ***	*oskuolatïttan*	*iti*	*bardïlar*
Popova –(I)ŋ	oskuola–(t)IttAn	iti	bar–TI–LAr
P.–POSS.2SG	school–ABL.3SG	this	go–PST–PL

'They finished that Popova's school, where you pass four years (of curriculum) in one, they left from there.'

[IvaP190]

(98b)

kü:ste:χ	*Böčüöχöb**üŋ***	*otto*	*ol*	*hala:kka*	*baran*
kü:s–LA:χ	Böčüöχöb–(I)ŋ	otton	ol	sala:t–GA	bar–An
strength–PROP	B.–POSS.2SG	CP	that	soldier–DAT	go–PF.CVB

baran	*ol*	*huoχ*	*buolbuta*	*büten*
bar–An	ol	suoχ	buol–BIt–(t)A	büt–An
SEQ–PF.CVB	that	non.existence	become–PSTPT–POSS.3SG	end–PF.CVB

'Küste:x Böčüöxov (i.e. strong Böčüöxov) died (lit. became non-existent) after going to the soldiers.'

[P95_113]

(98c)

χata	ister	ebit	mantïkay**ïŋ**	haχalï:
χata	ihit–Ar	e–BIt	mantïka–(I)ŋ	saχa–LI:
PTL	hear–PRSPT	AUX–PSTPT	this–POSS.2SG	Sakha–ADVR

'Look, she understands Sakha, it seems.'

[Afny187]

The sentence in (98a) was said by the speaker to my host and consultant, who clearly knew about Mrs Popova and her school; while the sentence in (98b) was said by the speaker to her son. It was part of a whole enumeration of her relatives and neighbours and where they had lived and what had happened to them, and clearly the people she referred to were known to him, at least from stories. Finally, (98c) was said by one speaker to her sister in reference to me, who was sitting at the table with them. In all three cases the possessive marking clearly does not signal any kind of possession, nor any close relationship, and it does not have anaphoric reference, either. The use of the second singular possessive suffix in all of these cases is strongly reminiscent of the second singular possessive suffix plus alienable marking in the Udihe examples.

Thus, Sakha has clearly extended the functions of possessive marking beyond those found in Turkic languages in general. In addition to the 3SG possessive suffix serving to indicate situational and anaphoric referentiality, the use of other possessive suffixes has pragmatic functions, highlighting an association between important discourse participants or establishing an emotional link with the addressee. This is done most frequently with 2SG possessive marking, although 1SG and 3PL can fulfill such pragmatic functions as well. This is reminiscent of the pagmatic use of the possessive suffixes in Nganasan and other Samoyedic languages as well as the uses of the alienably marked possessive suffixes in Evenki, Èven, and Udihe. Sakha shows further similarities to these languages in the domain of possessive marking, such as a very strong tendency for possessive marking on inalienably possessed entities, as will be discussed below.

3.4.3 Possessive marking of inalienable entities in Sakha

3.4.3.1 Possessive marking of kinship terms in Sakha

An analysis of the life stories I recorded in the field shows that kinship terms in Sakha occur predominantly with possessive marking, as summarized by Table 3.17. However, there are some exceptions to this strong tendency that will be discussed below. In the table, nouns referring to kin are listed followed by the

number of tokens in my corpus of oral narratives and the number of times the kinship term was used without possessive marking. With respect to tokens for the nouns *oɣo* 'young.human/offspring', *uol* 'boy/son', *kï:s* 'girl/daughter', *eme:χsin* 'old.woman/wife', and *oɣonńor* 'old.man/husband' which have a double meaning, only those referring to kin were counted. In the last column 'Comment on unpossessed kinship terms' I give an abbreviated explanation of why these kinship terms may have not been marked for possessor agreement; these are discussed in more detail below.

Table 3.17: Number of unpossessed kinship terms in collected life stories

Kinship term	N	N$_{unposs}$	Comment on unpossessed kinship terms
jon 'family'	35	0	
aymaχ 'family'	9	1	used with a derivative suffix that marks relatives that are not related to the speaker
törüt 'ancestor'	6	1	general meaning, not referring to specific person's ancestors
ebe 'grandmother'	9	0	
ehe 'grandfather'	6	0	
töröppüt(ter) 'parents'	5	3	twice this refers to 'parents in general', not to a specific person's parents
iỹe 'mother'	59	3	in lexical compound *iỹe aɣa* 'parents'
mama	20	4	copied term; also, 1 example of code-switching
aɣa 'father'	83	3	in lexical compound *iỹe aɣa* 'parents'
papa	6	1	copied term
ta:y 'maternal uncle'	4	0	
eji:y 'older sister'	32	2	forms of address *eji:y Kïrïstïn* 'aunt Kristina'
balïs 'younger sister'	10	0	
ubay 'older brother'	16	0	
bïra:t 'younger brother'	12	0	
kergen 'spouse'	26	0	
eme:χsin 'wife'	12	3	possibly a way of showing dominance in a familial hierarchy
oɣonńor 'husband'	7	3	as above
oɣo 'child'	157	58	as above
kï:s 'daughter'	34	3	as above
uol 'son'	47	10	as above
kiyi:t 'daughter-in-law'	5	1	as above
kütüöt 'son-in-law'	3	2	as above
sien 'grandchild'	53	13	as above

As can be seen from the table, the kinship terms *aɣa* 'father' *iỹe* 'mother', *eji:y* 'older sister', *ubay* 'older brother', *balïs* 'younger sister', *bïra:t* 'younger brother', *ta:y* 'maternal uncle', *kergen* 'spouse', *ebe* 'grandmother', *ehe* 'grandfather', *aymaχ* 'family/relatives', and *jon* 'people, relatives, family' are practically always marked for possession[9]. It is notable that these are nearly all terms referring to older family members, siblings (i.e. kin at the same generational level as ego), and the rather formal term for 'spouse' *kergen*. Furthermore, all the instances of kinship terms occurring in the first 268 sentences of the mid-19th century memoirs of Uvarovskij (in Btl: 5-78) carry a possessive suffix. Unfortunately, Uvarovskij had very little family, being the sole surviving child of parents who appear not to have belonged to extensive families, so that the number of kin terms (only 39) mentioned in the first third of his narrative is relatively small, and these refer predominantly to his parents.

In my corpus of oral narratives, the lexical compound *iỹe aɣa* 'parents' and the term *töröppütter* (which literally means 'those who gave birth') form an exception to the observation that kinship terms referring to the older generation require possessive marking. Uvarovskij uses the expression *aɣa iỹe ikki* [father mother two] 'father and mother' (a coordinate noun phrase as compared to the compound occasionally used nowadays) when referring to his parents, marking both terms for possession of 1SG possessor, e.g.:

(99)

bu	*sirge*	*min*	***aɣam***	***iỹem***	*ikki*
bu	sir–GA	min	aɣa–(I)m	iỹe–(I)m	ikki
this	place–DAT	1SG	father–POSS.1SG	mother–POSS.1SG	two

Ejigeŋŋe	*bara*	*ilikterine*	*üčügey*	*nučča*
Ejige:n–GA	bar–A	ilik–tArInA	üčügey	nučča
E.–DAT	go–IPF.CVB	not.yet–COND.3PL	good	Russian

jietin	*tuttan*	*olorbuttara*
jie–(t)In	tut–(I)n–An	olor–BIt–LArA
house–ACC.3SG	build–REFL–PF.CVB	sit–PSTPT–POSS.3PL

'In this place my mother and father had built themselves a good Russian house and lived there before they went to Zhigansk.'

[Uvar64]

In two instances of my text count the word *töröppüt* is used with reference to parents in general, not to any specific person's parents, and this might be the reason for lack

---

[9] The three tokens each of *iỹe* 'mother' and *aɣa* 'father' that were not marked with a possessive suffix occurred in the lexical compound *iỹe aɣa* 'parents'.

of possessive marking. In two instances of the compound *iỹe aɣa* it is used by the speaker to refer to herself and her husband, e.g.:

(100)

onno	*otton*	*barïta*	*χaččï*	*ol*	***iỹe***	***aɣa***
on–nA	otton	barï–(t)A	χarčï	ol	iỹe	aɣa
that.OBL–LOC	CP	all–POSS.3SG	money	that	mother	father

öttütten
örüt–(t)IttAn
side–ABL.3SG

'Of course, there as well everything, all the money, came from her parents.'

[Efmy436]

In this case, it might be that the speaker was trying to play down the role she and her husband had played in supporting their children financially. But it is nevertheless not entirely clear why 'parents' as an entity should be viewed as being any less inalienable than a person's mother or father as separate individuals. One possibility might be that parents as a unit are not culturally salient[10], but that the individuals are of bigger importance. Furthermore, the current use of *töröppütter* might be a calque of Russian *roditeli* and as such might not underlie the requirement found for Sakha kin terms of marking inalienably possessed entities for possessor. Unfortunately, the compound as well as the synonym *töröppütter* occurs with too low a frequency in my data set to be able to make any inferences.

As to the Russian copies *mama* and *papa*, in my texts these were used very rarely; only four speakers used the term *mama*, three of whom used it only with possessive marking (6 tokens), while the fourth speaker, a woman who very frequently switched to Russian while telling me her story, used it 14 times. This speaker used it without possessive marking only four out of the 14 times; and one of these unmarked tokens was within a Russian clause. The term *papa* was used only six times by three speakers, and it was used only once without a possessive marker by the above-mentioned speaker when referring to her step-father: *on–no kel–en bihieχe **papa** buol–la* [that.OBL–LOC come–PF.CVB 1PL.DAT papa become–PST.3SG] 'he came and became a papa for us[11]'. This speaker's use of unmarked *mama* and *papa* may therefore simply reflect strong Russian influence. This is in good accordance with the fact that my primary consultants occasionally used *mama* and

---

[10] This was suggested to me by Bernard Comrie.

[11] The Russian translation of this sentence would be: Он стал у нас папой. This (as I understand it) is somewhat intermediate between 'he became our papa' and 'he became a papa for us'. There is definitely a bigger emotional distance than if the speaker had said *onno kelen papabït buolla* 'he came and became our papa' with possessive marking on *papa*.

papa without possessive marking in translations of Russian sentences; here, the original Russian sentence may have triggered the unmarked use of these terms.

Even though most kinship terms do follow the requirement of marking the possessor, it is interesting to note that the terms referring to kin of a younger generation, i.e. *oγo* 'child', *sien* 'grandchild', *kï:s* 'daughter', *uol* 'son', *kiyi:t* 'daughter-in-law', and *kütüöt* 'son-in-law', are quite frequently not marked for possession (although possessive marking occurs in the majority of instances), even when referring to somebody's progeny. Since *oγo*, *kï:s* and *uol* can mean just 'child' 'girl' and 'boy', respectively, possessive marking on these terms can actually help distinguish between the general 'young human' meaning and the 'progeny' meaning of these words. Thus, in some of the instances of unmarked *oγo* it is not entirely clear whether the focus of the speaker may not actually have been on the 'young human' rather than on the 'offspring' meaning of the word. Furthermore, the colloquial terms for 'husband' and 'wife' (*oγonńor* and *eme:χsin*, which literally mean 'old man' and 'old woman', respectively), are also found without possessive suffixes.

The lack of possessive suffixes on descendant-line kinship terms appears to be pragmatic. As explained to me by one consultant, a dominant family member can indicate his/her dominance by omitting the (expected) possessive suffixes on kinship terms. For instance, a woman who was very clearly the head of the family indicated this by frequently not marking the possessor on the terms denoting children and in-laws (101a). Another speaker, who had complained quite bitterly about her grandchildren being so greedy and demanding and exhausting (a very un-characteristic thing for a Sakha to do), also referred to these grandchildren five times out of six without possessive marking, e.g. (101b). Furthermore, the lack of possessive marking appears to indicate a form of emotional distance; thus, when it was mentioned that children died (101c) or suffered (101d) the possessive marking was omitted.

(101a)

onton	*bi:r*	*ere*	*tabaχsït*	*bi:r*	**kütüöt**	*tabaχ*	*taddar*
ol–(t)tAn	bi:r	ere	tabaχ–SIt	bi:r	kütüöt	tabaχ	tart–Ar
that–ABL	one	only	tobacco–AGNR	one	son.in.law	tobacco	pull–PRSPT

'Of those, only one is a smoker, one son-in-law smokes.'

[Efmy202]

(101b)

sietter[12] *emie buka barï üöreχtenen*
sien–LAr emie buka barï üöreχ–LA:–(I)n–An
grandchild–PL also completely all studies–VR–REFL–PF.CVB

ereller
er–Ar–LAr
PROG–PRSPT–PL

'The grandchildren are also all getting an education.'

[MalA6]

(101c)

oγolor *delbi ölön χa:llïlar*
oγo–LAr delbi öl–An χa:l–TI–LAr
child–PL INTS die–PF.CVB RES–PST–PL

'The children died terribly.'

[Chir75]

(101d)

hassïarda **oγolor** *čï:bïrγa:n turdaχtarïna ulaχan*
sarsïarda oγo–LAr čï:bïrγa:–An tur–TAχ–TArInA ulaχan
morning child–PL chirp–PF.CVB stand–MDL–COND.3PL big

üöri:[13]
üör–I:
be.glad–NR

'If/when the children stood up chirping in the morning the joy was great.'

[P90_8]

Example (101c) was said by an old man who had had ten children who had all died before the age of five, i.e. in reference to his own children; while (101d) was said by a woman in reference to the time of World War II, when people in Yakutia were starving to death; her own children had been small then, and she was saying how difficult it was and how very often she had had to put them to bed hungry.

Additional support for the very strong tendency to mark the possessor on kinship terms in Sakha comes from the word list published in the appendix of Martin Sauer's (1803) description of the expedition to northeastern Siberia under Captain

---

[12] The plural form in *–tt–* is found relatively frequently in nouns denoting humans ending in *–n*; it is standard in the semi-suppletive forms for 'girls' *kïrgïttar* (instead of expected *kï:star*) and 'boys' '*uolattar*' (instead of '*uollar*'), but I have also found it in forms such as *oyu:ttar* 'shamans' (from *oyu:n*) or *baraχsattar* 'the dears' (from the term of endearment *baraχsan*). These forms can be analyzed as doubly marked nouns, with the Turkic plural *–LAr* following on a plural suffix *–t*, which was probably copied from Mongolic (Kałużyński 1962: 116f; Xaritonov 1947: 102).

[13] The pronunciation of *üöri:* (instead of expected *üörü:*) is probably just a slip of the tongue.

Joseph Billings (cf. section 3.1.3). In this word list six out of eight kinship terms are given with 1SG possessive suffix. Thus we find in Sauer's spelling (i.e. with *j* representing the glide transcribed as *y* by me): *agam* 'father' (correctly *aɣam* 'my father'), *ija* 'mother' (correctly *iỹe*), *oal* 'son' (correctly *uol*), *kihsim* 'daughter' (correct modern-day form: *kï:hïm* 'my daughter'), *ubagim* 'brother' (correctly *ubayïm* 'my older brother'), *agasim* 'sister' (correct modern-day form: *aɣahïm* 'my older sister'), *erim* 'husband' (correctly 'my husband), and *jaghtarim* 'wife' (correct modern-day form: *ʤaχtarïm* 'my wife') (Sauer 1803: 317). This implies that although Sauer was probably asking for neutral non-possessed forms (as indicated by the translation), his informants by default marked the kinship terms with a possessive suffix. This is indirectly confirmed by Katja Potapova, who writes: „I think that *aɣa* always requires a further reference noun (or pronoun), e.g. *Bu min aɣam* ['this is my father'], *Bu kini aɣata* ['this is his father']. *Bu aɣa ['this is father'] sounds somehow incomplete." (Katja Potapova, e-mail 22.12.2006, translation mine[14]).

3.4.3.2 Possessive marking of body part terms in Sakha

In my corpus of oral narratives, terms referring to inalienably possessed body parts also occur predominantly with possessive marking. For the analysis, body part terms in idioms such as *ili: bat* [hand press] 'to sign (one's name)' [Efmy255] as well as those with derivational morphology, e.g. *kïp~kïhïl battaχ–ta:χ* [INTS~red hair–PROP] 'with very red hair' [MatX1_23] were excluded. There were a total of 97 tokens in my texts referring to body parts or bodily excretions. Of these, 14 tokens occurred in adnominal possessive constructions of the form N+N–POSS, i.e. the general form for all such adnominal constructions in Sakha; these were therefore excluded from the analysis. In the Instrumental case, there is no distinction between the third person singular possessive form and the unpossessed form of nouns ending in a consonant, e.g. *süreɣinen* could be analyzed as *süreχ–(I)nAn* 'heart–INS' or as *süreχ–(t)InAn* 'heart–INS.3SG'. There were seven such cases in my texts concerning the body part terms *ayaχ* 'mouth', *süreχ* 'heart', *kurtaχ* 'stomach' and *uŋuoχ* 'bone'. I checked these forms by asking how they would appear with a 1SG subject – in all cases I was given possessively marked forms, e.g. *min **kuttaχ–pïnan** ïaljï–bït–ïm* [1SG stomach–INS.1SG be.ill–PSTPT–POSS.1SG] 'I had stomach problems'. Of the 83 tokens that did not occur in adnominal possessive constructions, only 18 were

---

[14] Original: "Ich denke, "*aɣa*" verlangt immer nach einem weiteren Bezugsnomen (oder Pronomen): z.B. *Bu min aɣam. Bu kini aɣata.*Bu aɣa.* klingt irgendwie unvollständig."

clearly not marked for possessor. Six of these are modifiers in lexical compounds and as such do not refer to any specific person's inalienable body part, e.g. *ti:s bïra:h–a* [tooth doctor–POSS.3SG] 'dentist' [RaxA177], *χaraχ χolo:h–unan emte:–n* [eye measure–INS.3SG heal–PF.CVB] 'healing by eye's measure' [XatR214], *bïar ïarï:–tïgar illar–an* [liver illness–DAT.3SG be.ill.long–PF.CVB] 'suffering from liver problems for a long time' [Chir83], while the others can be explained by different factors. For example, the statement *ti:s–ter–i tur–a–γïn* [tooth–PL–ACC pull–IPF.CVB–PRED.2SG] 'you pull teeth' [RaxA179] refers not to the teeth of any one specific person, but to teeth in general. In other instances, the speaker broke off what he or she had started to say; and in one case, the speaker seemed to be getting confused about which person to use for his narrative: he started off in first singular, switched to first plural, then to second singular, then to third singular, and at the end of it all he said the following (102), where the lack of possessive marking on *ti:s* might be explained by the general confusion of his narrative:

(102)

onno	*χamna:bat*	*da*	**ti:s**	*buollar*
on–nA	χamna:–BAt	da	ti:s	buol–TAr
that.OBL–LOC	move–PRSPT.NEG	PTL	teeth	AUX–COND

ïarï:ta	*bert*
ïarï:–(t)A	bert
illness–POSS.3SG	INTS

'Even though the teeth weren't loose, they hurt unbearably.'

[Efmy587]

In another case, a person telling me his life-story responded to my stuttering a Sakha question whether he still hunted as follows:

(103a)

min	*bultu:r*	*bultu:r*	*etim*
min	bult–LA:–Ar	bult–LA:–Ar	e–TI–(I)m
1SG	bag/catch–VR–PRSPT	bag/catch–VR–PRSPT	be–PST–POSS.1SG

'I used to hunt.'

biligin	*biligin*	*huoχ*	**χaraχ**	*kuhaγan*
biligin	biligin	suoχ	χaraχ	kuhaγan
now	now	non.existence	eye	bad

'Not now, (my) eye is bad.'

[YmyE95, 96]

In this case, it was arguably my broken attempts at Sakha which elicited a 'pidgin-like' simplified response on his side; this is all the more probable, since in his further elaborations he marked χaraχ 'eye' for 1SG possessor, e.g.:

(103b)

bu	**χaraχpïn**	ïnaχ	muohugar	ölörbütüm
bu	χaraχ–BIn	ïnaχ	muos–(t)IgAr	öl–(I)Ar–BIt–(I)m
this	eye–ACC.1SG	cow	horn–DAT.3SG	die–CAUS–PSTPT–POSS.1SG

bïlïrï:n
bïlïrï:n
last.year

'I hurt my eye here on a cow's horn last year.'

[YmyE97]

However, there are four exceptions to the rule in my text collection which cannot be readily explained. For example, we find

(104)

kenniki	**uŋuoχ**	kïbïttïbït	dien
kenniki	uŋuoχ	kïbït–(I)n–BIt	die–An
afterwards	bone	squeeze–REFL–PSTPT	say–PF.CVB

'…afterwards the bone got squeezed.'

[Efmy306]

where uŋuoχ 'bone' is not marked for 3SG possessor, as expected, but is left completely unmarked. This could possibly be the idiosyncrasy of one couple of speakers, a husband and wife, who did not mark uŋuoχ 'bone', silge 'ligaments' and one instance of nierbe 'nerves' (in another instance they did mark nierbe for possessor).

Nevertheless, only a small fraction of the nouns referring to body parts in the narratives I recorded, which were all oral narratives susceptible to slips of the tongue and other errors, occurred without possessive suffixes, while practically all the body part terms occurring in the first 268 sentences of Uvarovskij's narrative (Btl: 5-78), which is a written text, are marked for possession[15].

---

[15] This is approximately one third of the complete text. A total of 56 nouns referring to body parts occurred; of these, only three were unpossessed – one of these formed part of an idiom (χaraχ bes 'envy'), and two were used in a very abstract sense and did not refer to specific body parts of specific persons.

3.4.3.3 Further inalienable possessions in Sakha

In addition to kinship terms and body part nouns, in Sakha the nouns *öy* 'mind, thought' and *sana:* 'thought' appear also to require possessive inflection. The number of tokens of these words used without derivational suffixes (such as the Proprietive suffix *–LA:χ*) is small, however, so this statement is somewhat uncertain. But the few examples of these words in their underived form in my corpus are all marked for possessor, e.g. ***öyü–m** daɣanï küččügüy* [mind–POSS.1SG PTL tiny] 'my memory is bad (lit. my mind is tiny)' [P90_19], *ïra baɣa **hana:–m*** [presentiment wish thought–POSS.1SG] 'my presentiment and wish is' [IvaP90].

Furthermore, terms denoting friends or comrades appear to take possessive marking as well when they are not used as a form of address, e.g. *Bieribey Kiristiepel die–n kïrjaɣas kihi e–te **taba:rïh–ïm** ol öl–büt–e* [first K. say–PF.CVB old person AUX–PST.3SG comrade–POSS.1SG that die–PSTPT–POSS.3SG] 'So-called First Xristofor was an old man, my comrade, he died.' [Chir40], *bu L. mama–ta **podruga–m*** [this L. mama–POSS.3SG girlfriend–POSS.1SG] 'now L.'s mother (is) my friend…' [Efmy343]. Of course, as with kinship terms, it is hard to imagine any context in which one might say 'a/the friend' rather than 'your friend', 'my friend', or 'Vasya's friend'. However, the fact that, like kinship terms, terms denoting friends and comrades require a special derivational suffix when they are not in a relationship with the speaker, indicates that they belong to the category of inalienable possession as well (cf. section 3.4.5).

A very good case can be made for a person's age being inalienably possessed: both the number of years and the word *sa:s* 'spring' are marked for possessor when they refer to a person's age (105a, b). This is invariable for the number of years, while there are three instances of *sa:s* occurring without a possessive suffix, e.g (105c). As was explained to me by my consultant in the Verxojansk district, this emphasizes the very substantial age of the speaker (88 years old); a younger person would not be able to use this unpossessed expression. Furthermore, periods of time in a person's life that are still continuing are marked for possessor (105d).

(105a)

min	*alta*	*uon*	***altam***	*B.*	*hette*	*uon*	***hettete***
min	alta	uon	alta–(I)m	P.	sette	uon	sette–(t)A
1SG	six	ten	six–POSS.1SG	P.	seven	ten	seven–POSS.3SG

'I'm sixty six, P.is seventy seven ….'

[Efmy539]

(105b)

ha:hïm	*kïaybat*	*da*	*buollar*
sa:s–(I)m	kïay–BAt	da	buol–TAr
spring–POSS.1SG	be.able–PRSPT.NEG	PTL	AUX–COND

hïssïhan	*ki:rbitim*
sïrïs–An	ki:r–BIt–(I)m
race–PF.CVB	enter–PSTPT–POSS.1SG

'Even though I wasn't old enough, I ran after (them).'

[IvaP14]

(105c)

aɣïhuon		*aɣïs*	**sa:s**	*tuolla*
aɣïs	uon	aɣïs	sa:s	tuol–TA
eight	ten	eight	spring	turn.(of.age)–PST.3SG

'I've turned eighty eight.'

[Chir72]

(105d)

bi:r	*oɣom*	*Uolbaɣa*	*ostuoruya*	*uču:tala*
bi:r	oɣo–(I)m	Uolba–GA	ostuoruya	uču:tal–(t)A
one	child–POSS.1SG	U.–DAT	history	teacher–POSS.3SG

buolan	*oloror,*	*ühüs*	**sïla**
buol–An	olor–Ar	üs–(I)s	sïl–(t)A
AUX–PF.CVB	sit–PRSPT	three–ORD	year–POSS.3SG

'One of my children has been living in Uolba for three years, working as a history teacher.'

[XatR318]

Thus, in Sakha inalienably possessed entities (not only kinship terms, friends, and body parts, but also age and perhaps the mind) show a very strong tendency of requiring possessive marking. This requirement, however, is not entirely obligatory, as it can be overridden by pragmatic considerations, especially in those instances when the kinship term denotes kin of a younger generation. Nevertheless, this extensive use of possessive marking on kinship and body part terms is certainly noticeable. It is, however, very difficult to judge whether this extensive possessive marking may be attributable to contact influence, since it is hard to find comparable data for other languages. Writers of grammars might not always mention these features, which are of a frequential nature and not morphologically marked. In the absence of textual data of a similar kind as my corpus of narratives for Sakha it is very hard to evaluate whether other languages might not show a similar tendency to mark the possessor on inalienably possessed nouns, or whether this might be omitted. The following can thus only be regarded as a very cursory survey of the feature in some Eurasian languages.

3.4.4 Similar possessive marking on inalienably possessed entities in neighbouring languages

It is perhaps no coincidence that Ubrjatova, a Turcologist, was so struck by the extensive use of possessive marking on inalienably possessed nouns in Dolgan that she explicitly mentions it in her description:

> "Some of them [nouns – B.P.] are practically not used without possessive affixes: body part terms, household items, kinship terms. For example, it is nearly impossible to use the nouns *ili:* 'hand/arm', *atak* 'foot/leg', *aga* 'father', *ińe* 'mother' without possessive affixes, since there are no hands in general (existing independently of a person), as there are no fathers in general, but there exist fathers of specific persons, the relationship to whom has to be indicated in speech without fail." (Ubr: 114; translation mine[16])

This might be an indication that the system found in Dolgan and Sakha differs from what is found in most Turkic languages.

Similarly, Anderson (2004: 7) writes that in both Khakas and Tofa certain body part terms and kinship terms always appear in possessive form and suggests that this might perhaps be due to a Samoyedic substrate. This might be found in other South Siberian Turkic languages as well, such as Chulym; however, the 'obligatory' possessive marking is not applied consistently to all kinship and body part terms that might be expected to take it (Gregory Anderson, pers. comm.).

Two further languages for which the extensive use of possessive marking is explicitly mentioned are Nganasan and Mansi. In Nganasan, inalienably possessed nouns such as kinship and body part terms and nouns denoting household items are, as a rule, used only with possessive suffixes (Ter: 94), which is reminiscent of the extensive possessive inflection on such entities in Sakha and Dolgan. In Mansi, too, kinship terms, body parts and clothing terms are used with possessive suffixes, which seems comparable to the case in Sakha and Nganasan. Furthermore, nouns denoting periods of time that started in the past and are still continuing are also possessively marked, e.g. 'She married three years ago', where 'years' would be marked for 3SG possessor (Rmb: 63). This is very similar to Sakha, where periods of time in a person's life that are still ongoing are possessively marked (cf. section 3.4.3.3 and example 91d). From the translations given for some of the examples for

---

[16] Original: "Некоторые из них без притяжательных аффиксов почти не употребляются: названия частей тела, предметов домашнего обихода, родственные термины. Например, почти невозможно употребить без притяжательных аффиксов существительные илии 'рука', атак 'нога', ага 'отец', иньэ 'мать', поскольку нет руки вообще (независимо от человека существующей), как нет отца вообще, а есть отец определенных лиц, отношение к которым должно быть обязательно в речи отмечено."

possessive-marked noun phrases in the Vakh dialect of Khanty it appears as if kinship terms might always take possessive marking; thus, *əpï–m Iŋki–m–nœti* [father–POSS.1SG mother–POSS.1SG–COM] is translated as (unpossessed) 'father and mother (отец с матерью)' (Trš: 42). However, without further data it is very hard to come to a conclusion on this point regarding Khanty.

Furthermore, as mentioned in section 3.4.2.1, in the Tungusic languages kinship and body part terms are generally marked with possessive suffixes. In Evenki, kinship terms only occur without possessive suffixes if they are used vocatively, while in Udihe, kinship terms not referring to any specific person but having only generic reference can be used without possessive suffixes.

For most of the other languages surveyed (Turkish, Tuvan, Chuvash, Yukaghir, Nivkh, Ket, Itelmen, Chukchi, and Eskimo) no mention is made whether inalienably possessed entities require possessive marking or not. For Bashkir, the fact that no mention is made of kinship or body part terms generally taking possessive marking in a rather detailed list of 'lexico-grammatically conditioned meanings of the possessive forms' in a Russian grammar of Bashkir (GSBJ: 124f) might be an indication that this is not a salient feature of the language. However, this can be at most a tenuous conclusion, since the authors of the grammar may not have considered such a feature worthy of mention. One language, however, where it is relatively clear that kinship terms at least do not require possessive marking, is Khalkha Mongolian. Although neither Kullmann & Tserenpil (1991) nor Vietze (1988) specifically make any statement about inalienably possessed items, from some of the examples given it becomes clear that kinship terms do not need to be marked for possession, an observation that was confirmed by Klaus Koppe (pers. comm.), e.g.:

(106)

aav	*egč–id*	*n'*	*xiče:l*	*za:–dag*
father	older.sister–DAT	POSS.3SG	lesson	show–HAB

'Father gives his/her older sister lessons.'

[K/Ts: 107]

Thus, for some languages spoken in Siberia there is evidence that they show a strong tendency of requiring possessive marking on inalienably possessed nouns. Apart from Sakha, these are Dolgan and some South Siberian Turkic languages, the Tungusic languages, and the Uralic languages Nganasan and Mansi. Judging solely from the descriptions available to me, amongst the northern Eurasian languages the Sakha system shows similarities to Nganasan, where kinship and body part terms as well as nouns denoting household items require possessive marking (Ter: 94). There are also similarities between Nganasan and Sakha in the 'non-possessive' use of the

second person possessive suffixes, as discussed in section 3.4.1. However, possessive marking in Sakha is also strongly reminiscent of the Tungusic system, with the strong tendency of possessive marking on inalienable nouns and the similarities in the pragmatic use of the possessive suffixes (with the addition of the Alienable Possession suffix in the Tungusic languages) discussed in section 3.4.2.2. As will be demonstrated in the following section, there is another point of similarity between Sakha and at least some Tungusic languages in the domain of possessive marking.

3.4.5 Referring to kin/friends that have no relationship to the speaker

In Sakha there exists a special derivational suffix –*LI:* which is used to refer to people who are relatives or friends, but who do not have a relationship to the speaker. This suffix is homonymous to one of the adverbializing suffixes, e.g. *ńučča–lï:* [Russian–ADVR] 'in Russian', as well as to the numeral distributive suffix, e.g. *ikki–li:* [two–DIST] 'two each'. It occurs with nouns denoting siblings, e.g. *bïra:t–tï:–lar* [younger.brother–DER–PL] 'brothers', *balïs–tï:–lar* [younger.sister–DER–PL] 'sisters', *eji:y–di:=balïs–tï:–lar* [older.sister–DER=younger.sister–DER–PL] 'sisters' (this emphasizes the fact that one is older, one is younger), with 'neighbour', 'family', and 'spouse', e.g. *ïal–li:–lar* [neighbour–DER–PL] 'neighbours', *aymaχ–tï:–lar* [family–DER–PL] 'relatives', *kergen–ni:–ler* [spouse–DER–PL] 'spouses', with terms denoting friends, e.g. *atas–tï:–lar* [chum–DER–PL] 'chums', *jüöge–li:–ler* [girlfriend–DER–PL] 'girlfriends', and with *sa:s* 'spring/age', e.g. *sa:s–tï:–lar* [spring–DER–PL] 'people of the same age'. Thus, example (107a) means that Fedya is playing with his own two younger brothers, while (107b) means that he is playing with two boys who are brothers, but not his.

(107a)

Fedya	*ikki*	*bïra:**tïn***	*kïtta*	*o:nńu:*	*hïjjar*
F.	ikki	bïra:t–(t)In	kïtïn–A	o:nńo:–A	sïrït–Ar
F.	two	younger.brother–ACC.3SG	join–IPF.CVB	play–IPF.CVB	IPFV–PRSPT

'Fedya is playing with his two younger brothers.'

(107b)

Fedya	*ikki*	*bïra:**ttï:larï***	*kïtta*
F.	ikki	bïra:t–LI:–LAr–(n)I	kïtïn–A
F.	two	younger.brother–DER–PL–ACC	join–IPF.CVB

o:nńu:	*hïjjar*
o:nńo:–A	sïrït–Ar
play–IPF.CVB	IPFV–PRSPT

'Fedya is playing with two brothers.'   [Sun, translation]

Similarly, in one of the narratives I recorded, the speaker is referring to other people who are related, and she excludes herself from this relationship:

(108)

iti	*Ispi:rep*	*kïrjaɣas*	*Ispi:rep*	*kergene*	*aymaχtï:lar*
iti	Ispi:rep	kïrjaɣas	Ispi:rep	kergen–(t)A	aymaχ–LI:–LAr
this	I.	old	I.	spouse–POSS.3SG	family–DER–PL

bïhï:la:χ	*ete*	*onno*	*kelen*
bïhï:–LA:χ	e–TA	on–nA	kel–An
appearance–PROP	AUX–PST.3SG	that.OBL–LOC	come–PF.CVB

a:har	*ete*
a:s–Ar	e–TA
pass–PRSPT	AUX–PST.3SG

'Old Spirov's wife was her (the grandmother's) relatives it seems, she (the grandmother) came there.'

[P95_98]

Here, it is interesting that the speaker is talking about her grandmother's relatives using the derivational suffix that excludes her from the relationship. The reason for this is most probably that she grew up in a different family and hardly knew her grandmother, as she had only seen her a few times at 'Old Spirov's' house as a child.

The use of this suffix is interesting for two reasons. First of all, it underlines the fact that nouns which are normally inalienably possessed (kinship terms as well as friends or comrades) cannot generally stand alone, without possessive marking. If possessive marking is inappropriate, because the persons referred to are not related to the speaker, the derivational suffix is inserted to mark this fact. Thus, it is not possible to say ***balti–lar** *ulu:sa–nan hïjj–al–lar* [younger.sister–PL street–INS walk–PRSPT–PL] '(a group of) sisters are walking down the street'. One either has to say **balti–lar–ïm** [younger.sister–PL–POSS.1SG] *ulu:sanan hïjjallar* 'my younger sisters are walking down the street' (or use a corresponding possessive suffix, depending on the context), or **balti–lï:–lar** [younger.sister–DER–PL] *ulu:sanan hïjjallar* '(a group of) sisters are walking down the street' (Margarita Ivanova, pers. comm.). Secondly, the use of this suffix is strongly reminiscent of a derivational suffix found in Èven, which "derives nouns from nominal roots denoting humans with a meaning of two persons who have something in common, a shared feature" (Nov: 228, translation mine[17]), e.g. *ge:–**nun*** [friend–DER] 'friends', *aqə–**nun*** [older.brother–DER] 'two brothers'. Unfortunately, Novikova does not say any more about the use of this

---

[17] Original: "… с помощью которого от именных основ названий людей образуются новые имена со значением двух лиц, имеющих какое-либо сходство или одниаковый признак..."

suffix, although she does suggest that it is related to the Comitative suffix –ńun, and neither Benzing (Ben: 20-34) nor Malchukov (Mal: 11) mention it. Interestingly, in Udihe there is a non-productive derivational suffix –mule which is clearly related to the comitative postposition mulə 'with' (cf. section 3.2.4.4). This suffix derives a "closed class of nouns meaning close symmetrical relationship between two or more people", e.g. *gagda–mulə* [other–DER] '(married) couple', *xunazi–mulə* [elder.sister–DER] 'sisters', *gətu–mule* [comrade–DER] 'comrades' (N/T: 163f). The nouns derived with this suffix do not occur with the regular possessive suffix, and from some of the examples they appear to function in a similar way to the Sakha nouns derived with –LI:, e.g.:

(109a)

   bagdi:–ti   *zu:*   *'aga–**mule***
   live–3PL   two   elder.brother–DER

   'There live two brothers.'

(109b)

   ni   *ugda–ni? //*   *zu:*   *xa:–**mule**–ŋi*
   who   boat–3SG //   two   sibling–DER–PRDPOSS

   'Whose is the boat? The two brothers'.'

<div align="right">[N/T: 164]</div>

Unfortunately, Nedjalkov (Ned: 297f) does not provide information whether a similar Comitative-related suffix exists in Evenki, but a suffix with similar functions is mentioned for Nanay and Oroč (cf. section 3.2.4.4). It is very interesting in this context that the Sakha derivational suffix –LI: might perhaps be related to the Comitative case suffix –LI:n, thus possibly sharing not only the function of the Èven and Udihe suffixes, but their source as well. This might be an indication that contact may have played a role in the development of the Sakha derivational suffix.

3.4.6 Contact influence in the domain of possessive marking in Sakha?

The pragmatic 'non-possessive' use of the possessive suffixes, especially the 2SG suffix, in Sakha is indicative of contact influence, since this feature is quite distinct from what is found in the Turkic languages, apart from Dolgan (Stachowski 1998). Whether the very strong tendency to mark the possessor on inalienably possessed entities might also be due to contact influence cannot really be ascertained given the data (or rather, the lack of it) available for comparison. However, the similarity between the very extensive possessive marking on inalienable nouns in Nganasan and Sakha, as well as the use of the 2SG possessive marker in both Sakha and Nganasan to mark anaphoric reference are striking and make Nganasan (or

Samoyedic) contact influence appear likely. However, an argument against this is the lack of known historical contact between speakers of a Samoyedic language and Sakha, and the fact that the phenomenon of the 'non-possessive' use of the 2SG possessive suffixes is not areally restricted in Sakha, judging from my Tabalaax consultant's evaluation. On the other hand, a link between Uralic populations and the Sakha is provided by the high frequency of the Y-chromosomal SNP Tat C in these groups (cf. section 1.1.1.2), indicating possible Uralic male admixture or substrate in the Sakha. However, this polymorphism is absent from Selkups and practically absent from Nganasans, although it is found in high frequency in Nenets (Karafet et al. 2002). This makes specifically Nganasan admixture or substrate rather unlikely.

The Sakha possessive system also exhibits a similarity to the Tungusic system, especially in the functions of the non-possessive uses of the possessive suffixes, which show striking parallels to the use of the possessive suffixes with further alienable marking in the Tungusic languages. Furthermore, the extensive possessive marking of inalienably possessed entities in Sakha may have been reinforced by contact with Evenki, a language in which inalienably possessed kinship terms and body part terms do not generally occur without possessive marking. In addition, the Sakha derivational suffix –*LI:* appears to show a similarity in function to the Èven and Udihe suffixes –*nun* and –*mulə*, and it further supports the importance of possessive marking on terms denoting inalienably possessed kin and friends. Given the known historical contact between speakers of Tungusic languages and Sakha, and given the prevalence of documented changes undergone by Sakha through contact with speakers of Evenki (cf. sections 3.2.1, 3.2.3, and others), Evenki contact influence may be more plausible than Samoyedic.

3.5 Person-marked converbs in Sakha

A notable and noted feature in which Sakha differs from the other Turkic languages is that converbs optionally take subject agreement markers. It has been suggested that this may be due to some degree of Evenki influence (Johanson 2001: 1732; Ubrjatova 1976: 45), since in Evenki converbs take subject agreement suffixes to participate in a system of syntactic referent tracking (I. Nedjalkov 1995). However, in the following I will show that the development of Sakha person-marked converbs is more likely to represent an independent innovation than the result of contact influence.

Before turning to the Sakha converbal forms, however, it may be of use to give a brief explanation of what converbs are. Following Haspelmath (1995), a converb can be defined as "a nonfinite verb form whose main function is to mark adverbial subordination" (Haspelmath 1995: 3). More specifically, converbs can function as verbal adverbs by modifying other verbs, as secondary predicates of coordinated clauses, and as the predicate of subordinate clauses (V. Nedjalkov 1995: 98f). They often derive from case forms of verbal nouns or from participles (Haspelmath 1995: 17), and play a large role in languages that do not make much use of conjunctions in syntactic coordination and subordination (V. Nedjalkov 1995: 100). In addition to what Vladimir Nedjalkov (1995: 102) calls 'canonical' converbs, converbal functions can also be performed by other verbal forms, especially by participles and infinitives. Such multifunctional verb forms in converbal function have been termed 'quasi-converb' (V. Nedjalkov 1995: 103; Janhunen 2003d: 26). Depending on the coreferentiality with the subject of the main clause, converbal constructions can be classified as same-subject (SS) constructions, in which the subject of the converb is coreferential with the subject of the main verb, or different-subject (DS) constructions, in which the subject of the converb is non-coreferential with the subject of the main verb. Some converbs obligatorily occur in either SS or DS constructions, while others can occur in both; these are termed variable-subject (VS) converbs (Čeremisina 1977: 24; V. Nedjalkov 1995: 110). See section 3.5.3 below for a description of the Evenki system, which makes use of all three kinds of converbs.

3.5.1 Converbs in Sakha

Like other Turkic languages, Sakha makes extensive use of converbs in coordination (110a), clause chaining (110b), adverbial modification (110c) and, in conjunction with auxiliaries, in the formation of aspectual meanings (cf. the third

converbal form *ölön* in example 110a, where the Perfective Converb *–An* in conjunction with the auxiliary *χa:l* (lit. 'stay') provides a resultative meaning). Some of these converbs, such as the Perfective Converb and the Imperfective Converb, are unanalyzable, while others, most importantly the Privative Converb *–BAkkA*, are historically derived from a case-marked participle[1]. As seen from examples (110a-c), the converbal subjects are most often coreferential with the subject of the main verb; however, occasionally converbs can occur in DS constructions as well (110d).

(110a)

oγobut …	*tïmnïyan*	*hötöllön*	*ölön*	*χa:lla*
oγo–BIt	tïmnïy–An	sötölün–An	öl–An	χa:l–TA
child–1PL	be.cold–PF.CVB	cough–PF.CVB	die–PF.CVB	RES–PST.3SG

'Our child … caught a cold and coughed and died.'

[RaxA39]

(110b)

ol	*Akkï:ray*	*a:rtïga*	*dien …*	*Boroγonton …*
ol	akkï:ray	a:rtïk–(t)A	die–An	Boroγon–(t)tAn
that	clergyman	mountain.pass–POSS.3SG	say–PF.CVB	B.–ABL

onon	*kelen*	*bu*	*Bayaγantay*	*ulu:hugar*
on–(I)nAn	kel–An	bu	Bayaγantay	ulu:s–(t)IgAr
that.OBL–INS	come–PF.CVB	this	B.	district–DAT.3SG

ki:ren	*Tompoγo*	*Kirieske*	*taχχan*	*Kiriesten*
ki:r–An	Tompo–GA	Kiries–GA	taγïs–An	Kiries–(t)tAn
enter–PF.CVB	T.–DAT	K.–DAT	go.out–PF.CVB	K.–ABL

tönnön	*kelen*	*Ta:tta*	*ulu:hunan …*	*kuoratïgar*
tönün–An	kel–An	Ta:tta	ulu:s–(t)InAn	kuorat–(t)IgAr
return–PF.CVB	come–PF.CVB	T.	district–INS.3SG	town–DAT.3SG

tönnör
tönün –Ar
return–PRSPT

'That so-called clergyman's road came from … Borogon, … coming through there it entered this Bajagantaj district and went on to Tompo and Krest', returning from Krest' it returned to town (Yakutsk) via the Taatta district, ….'

[XatR26]

---

[1]This can be analyzed as consisting of the negative Present Participle in the Dative case. Although this is of very recent origin (dating to within the last 100 years; Korkina 1985: 65f), it has by now grammaticalized to a 'canonical' converb. This can be seen from the fact that it can take predicative person markers that attach to the end of the entire suffix, e.g. *–BAkkABIn* for 1SG, instead of following the possessive declension expected for case-marked participles, which would be *–BAppAr* for 1SG (< *–BAt* 'PRSPT.NEG' + *–BAr* 'DAT.1SG').

(110c)

manna	*huolga*	*ölö*	*hïtar*
man–nA	suol–GA	öl–A	sït–Ar
this.OBL–LOC	road–DAT	die–IPF.CVB	lie–PRSPT

'Here they lay dying in the road.'

(110d)

ebe:	*ayannï:*	*hïjjan*	*emčitterge*
ebe:	ayan–LA:–A	sïrït–An	emp–SIt–LAr–GA
grandmother[VOC]	journey–VR–IPF.CVB	IPFV–PF.CVB	medicine–AGNR–PL–DAT

emčitter	*bileller*	*diebitiŋ*	*duo*
emp–SIt–LAr	bil–Ar–LAr	die–BIt–(I)ŋ	duo
medicine–AGNR–PL	know–PRSPT–PL	say–PSTPT–POSS.2SG	Q

ebe:
ebe:
grandmother[VOC]

'Grandmother, when you journeyed to the healers, the healers knew, you said, right?[2]'

In contrast to other Turkic languages, in Sakha the Perfective Converb –*An*, the Purposive Converb –*A:rI*, the Privative Converb –*BAkkA*, the Immediate Precedence Converb –*A:t*, and the Imperfective Converb –*A* optionally take subject agreement markers (Ubrjatova 1976: 43). Of these, the Perfective Converb, which often functions to coordinate verb phrases and sentences, is most often marked for subject agreement (approximately 28% of the instances of this converb in my corpus that do not occur in combination with an auxiliary; cf. examples 18a, 19b, and 95), while the Privative and Purposive Converbs are marked for person agreement in approximately 14% and 17% of the instances, respectively. The Imperfective Converb is used very rarely with subject agreement suffixes: of the approximately 400 instances of this converb in my corpus (once again excluding use with auxiliaries), only five instances are with person markers: twice with a 1SG predicative suffix and three times with the plural suffix. The Immediate Precedence Converb occurs only once in my corpus of texts, but it is easily obtained in elicitation, both with and without person marking. With the exception of the Immediate-Precedence Converb –*A:t*, which takes possessive suffixes, the subject

---

[2] This (and a few similar examples in my corpus) contradicts Čeremisina's claim that the non-coreferential use of the Perfective converb –*An* requires person-marking (Čeremisina 1977: 22). It also contradicts Efremov's claim that the use of this converb in DS constructions has an obligatory causal reading (Efremov 1979: 70).

agreement markers found on converbs belong to the set of predicative person markers (cf. Table 3.11 in section 3.3.1.1, and Table 3.19). Since the 3SG predicative person suffix is zero, 3SG subject agreement cannot be marked on converbs with the exception of the Immediate Precedence Converb. Person-marked converbs in Sakha are most often coreferential with the subject of the main clause; the additional subject agreement marking is therefore redundant (111a-c).

(111a)

ol	*ebiet*	*buhar**aŋŋïn***
ol	ebiet	bus–(I)Ar–An–GIn
that	dinner	ripen–CAUS–PF.CVB–PRED.2SG

ahattalï:gïn
as–A:–T–(I)tAlA:–A–GIn
food–VR–CAUS–MULT–IPF.CVB–PRED.2SG

'So you prepare their dinner and feed them all.'

[MatX2_27]

(111b)

en	*ol*	*kihieχe*	*kuolaskin*	*bier**bekkeɣin***
en	ol	kihieχe	kuolas–GIn	bier–BAkkA–GIn
2SG	that	person.DAT	vote–ACC.2SG	give–PRV.CVB–PRED.2SG

jonu	*taŋnarï*	*tarda*	*hïtaɣïn*	*dien*
jon–(n)I	taŋnarï	tart–A	sït–A–GIn	die–An
people–ACC	downwards	pull–IPF.CVB	PROG–IPF.CVB–PRED.2SG	say–PF.CVB

'"By not giving that person your vote you are pulling the people downwards", he said.'

[IvaP73]

(111c)

tïl	*eten*	*turan*	*biere:ribin*	*χaččïbïn*
tïl	et–An	tur–An	bier–A:rI–BIn	χarčï–BIn
word	say–PF.CVB	stand–PF.CVB	BEN–PURP–PRED.1SG	money–ACC.1SG

eŋin	*ïla*	*kelen*	*olorobun*
eŋin	ïl–A	kel–An	olor–A–BIn
etc.	take–IPF.CVB	come–PF.CVB	sit–IPF.CVB–PRED.1SG

'I came in order to say a speech and to receive my money etc.'

[IvaP91]

The Imperfective Converb can be reduplicated to give a meaning of duration, iterativity, or distribution; and it is only in this reduplicated use that it takes subject agreement markers (112a). This restriction in person-marking is probably due to the fact that in conjunction with predicative first and second person suffixes the Imperfective Converb expresses present tense for the first and second person (112b).

(112a)

onno	*tïllana*	*tïll**abïn***
on–nA	tïl–LA:–(I)n–A	tïl–LA:–(I)n–A–BIn
that.OBL–LOC	word–VR–REFL–IPF.CVB	word–VR–REFL–IPF.CVB–PRED.1SG

ha:hïm	*kïaybat*	*da*	*buollar*
sa:s–(I)m	kïay–BAt	da	buol–TAr
spring–POSS.1SG	be.able–PRSPT.NEG	PTL	AUX–COND

hïssïhan	*ki:rbitim*
sïrïs–An	ki:r–BIt–(I)m
race–PF.CVB	enter–PSTPT–POSS.1SG

'There I begged and begged, and even though I wasn't old enough, I ran after them.'

<div align="right">[IvaP14]</div>

(112b)

nehi:le	*χa:mabïn*
nehi:le	χa:mp–A–BIn
barely	walk–IPF.CVB–PRED.1SG

'I walk with great difficulty.'

<div align="right">[BesP109]</div>

The only Sakha converb that has a function of tracking referents in complex clauses is the Immediate-Precedence Converb –*A:t*, which takes no subject agreement markers when its subject is coreferential with the subject of the main verb (113a), and Possessive-Accusative suffixes plus optionally the postposition *kïtta* 'with' when its subject is non-coreferential with the subject of the main clause (113b).

(113a)

min	*jieber*	*kel**e:t***	*küöspün*
min	jie–BAr	kel–A:t	küös–BIn
1SG	house–DAT.1SG	come–IMM.CVB	pot–ACC.1SG

küöstü:bün
küös–LA:–A–BIn
pot–VR–IPF.CVB–PRED.1SG

 'As soon as I come home I start cooking.'

<div align="right">[translation, Ver]</div>

(113b)

jieber	*kele**:ppin***	*kïtta*	*iỹem*
jie–BAr	kel–A:t–BIn	kïtïn–A	iỹe–(I)m
house–DAT.1SG	come–IMM.CVB–ACC.1SG	join–IPF.CVB	mother–POSS.1SG

ebieti	*belemne:bitinen*	*barar*
ebiet–(n)I	belem–LA:–BIt–(t)InAn	bar–Ar
lunch–ACC	ready–VR–PSTPT–INS.3SG	INCP–PRSPT

'As soon as I come home, my mother starts cooking lunch.'

[translation, Tat]

3.5.2 The functions of person-marked converbs in Sakha

Given the redundancy of the subject agreement marking on most converbs, the question is what function the person-marking has. Not all instances of person-marked converbs in my corpus of spoken life stories are functionally motivated, and it seems that their use is due in part to an idiosyncratic preference of individual speakers. This individual preference is not bound to a specific dialectal area, nor is it gender-related, since three women and three men from all four districts stand out in the frequent use they make of person-marked converbs. There is a correlation between degree of animatedness of the narrator and frequency of use of person-marked converbs, in that especially those speakers who got quite carried away by their narrative made frequent use of this feature, while more diffident speakers made less use of it. However, all speakers use person-marked converbs at least occasionally.

Nevertheless, there are three clearly discernible trends in the function of person-marked converbs: firstly, they emphasize the main actant of the sentence; as such, one occasionally finds strings of person-marked converbs and one or even several instances of free pronouns in the same sentence, i.e. an 'overload' of subject reference (114). Secondly, they facilitate reference tracking in discourse, being used after (and occasionally before) a switch in discourse referent (115). This contrasts with the use of unmarked converbs in sequences of sentences where no switch in discourse referent has taken place (116). And lastly, there is a discernible, though low-frequency tendency for person-marked Perfective Converbs to take the place of finite verbs, suggesting grammaticalization of a new TAM form (117a, b).

(114)

je	*onon*	*bu*	*kihieχe*	**min**	*bert*	*ereyinen*
je	on–(I)nAn	bu	kihieχe	min	bert	erey–(I)nAn
well	that.OBL–INS	this	person.DAT	1SG	INTS	labour–INS

χargïyan	*toŋon*	*öl**böχχöbün***	*če*	*hin*
χorguy–An	toŋ–An	öl–BAkkA–BIn	če	sin
be.hungry–PF.CVB	freeze–PF.CVB	die–PRV.CVB–PRED.1SG	well	fairly

χas	*da*	*hïl*	*ol*	*kurduk*	*olorboχto:n*
χas	da	sïl	ol	kurduk	olor–MAχtA:–An
how.many	PTL	year	that	like	sit–INTS–PF.CVB

*ihe**mmin***	*ïalï*	*keriye**mmin***
is–An–BIn	ïal–(n)I	keriy–An–BIn
DUR–PF.CVB–PRED.1SG	family–ACC	make.the.round–PF.CVB–PRED.1SG

ïal	*χonuk*	*mahïn*	*mastï:r*	*buolan*
ïal	χonuk	mas–(t)In	mas–LA:–Ar	buol–An
family	24.hour.period	wood–ACC.3SG	wood–VR–PRSPT	AUX–PF.CVB

*bar**ammïn***	*je*	*baran*	*χa:llïm*	*bu*
bar–An–BIn	je	bar–An	χa:l–TI–(I)m	bu
go–PF.CVB–PRED.1SG	well	go–PF.CVB	RES–PST–POSS.1SG	this

gihitten	*teyen*	*bukatïn*
kihi–(t)tAn	tey–An	bukatïn
man–ABL	leave–PF.CVB	completely

'Well, with that person it was difficult, I was hungry and cold, but I didn't die, well, I lived like that for a few years, I made the round of the people (I went from house to house), I chopped a day's worth of wood for people, well I left, I left this person completely.'

[PotP26]

(115)

onu	*bierbeteχtere*	*//*	*sa:ha*	*ïra:tta*
onu	bier–BAtAχ–LArA		sa:s–(t)A	ïra:t–TA
CP	give–PSTPT.NEG–POSS.3PL		spring–POSS.3SG	move.off–PST.3SG

'But they didn't send me.' // '"She's too old."'

üören	*bar**a:rïbïn***	*köppütüm*	*aγay*
üören	bar–A:rI–BIn	köt–BIt–(I)m	aγay
learn	go–PURP.CVB–PRED.1SG	fly–PSTPT–POSS.1SG	INTS

'Thinking I would go to school, I was very happy (lit: flew).'

onu	*jonum*	*a:χ*	*bierbeteχtere*
onu	jon–(I)m	a:χ	bier–BAtAχ–LArA
CP	people–POSS.1SG	COLL	give–PSTPT.NEG–POSS.3PL

'But my family didn't let me (lit: give me).'   [BesP208-211]

(116)

üs	sïl	ustata	onno	bara
üs	sïl	usta–(t)A	on–nA	bar–A
three	year	length–POSS.3SG	that.OBL–LOC	go–IPF.CVB

hïrïttïm	min	ol	Hutuoŋŋa	allara	Yanïskayga
sïrït–TI–(I)m	min	ol	Sutuon–GA	allara	Yanïskay–GA
IPFV–PST–POSS.1SG	1SG	that	S.–DAT	down	Y.–DAT

'For three years I was there, in Zaton, down in Yanskiy.'

je	ol	onton	ïla	je	üleni	χamnahï
je	ol	on–(t)tAn	ïl–A	je	üle–(n)I	χamnas–(n)I
well	that	that.OBL–ABL	take–IPF.CVB	well	work–ACC	salary–ACC

kïayan	je	kihi	kebiger	je	ki:ren	jon
kïay–An	je	kihi	kiep–(t)IgAr	je	ki:r–An	jon
be.able–PF.CVB	well	person	form–DAT.3SG	well	enter–PF.CVB	people

keχχetiger	je	hïllar	buollum
kekke–(t)IgAr	je	sïrït–Ar	buol–TI–(I)m
row–DAT.3SG	well	go–PRSPT	AUX–PST–POSS.1SG

'Well, from there on, being able to work and earn money, well I became a person (lit: entered the form of people), I became a grown-up, on equal footing with the others (lit: went in the rows of people).'

honton	ïla	je	bu	hïrïttaγïm	di
s(I)–on–(t)tAn	ïl–A	je	bu	sïrït–TAχ–(I)m	di:
EMPH–that.OBL–ABL	take–IPF.CVB	well	this	go–MDL–POSS.1SG	ASS

ülele:n	χamna:n
üle–LA:–An	χamna:–An
work–VR–PF.CVB	earn–PF.CVB

'Well, from then on I lived like this, working.'

[PotP79-83]

(117a)

bihigi	otto	bïlïr	Huχa:na	nehilieger	olorommut
bihigi	otton	bïlïr	Suχa:na	nehiliek–Ar	olor–An–BIt
1PL	CP	in.the.past	S.	nasleg–DAT.3SG	sit–PF.CVB–1PL

'Well, in the past we lived in the Suxana nasleg.'

[Afny45]

(117b)

ol	ihin	bihi	kömölöhömmüt	bastakï	üöreγïn	haγala:n
ol	ihin	bihigi	kömölös–An–BIt	bastakï	üöreχ–(t)In	saγala:–An
that	for	1PL	help–PF.CVB–1PL	first	studies–ACC.3SG	begin–PF.CVB

'So therefore we helped her when she first began her studies.'   [Efmy454]

3.5.3 Converbs in Turkic languages

There are a number of converbs in Old Turkic, some of which are morphologically unanalyzable, while others have clearly developed from other verbal forms. The converbs' subjects are mostly coreferential with the subject of the main verb, or alternatively the subject has to be inferred from the context. In some cases when converbs are used in DS constructions, the converbal subject stands in the Nominative (Erd: 308f). The 'contextual' converbs –*(X)p* and –*V* (corresponding in function to the Sakha Perfective and Imperfective Converbs) are most often coreferential with the subject of the main verb, while the converb –*gInčA* (with a meaning of 'as long as', 'until') is mostly used in non-coreferential clauses (Erd: 318). In addition to such canonical converbs there exist quasi-converbs in Old Turkic, which are case-marked forms of verbal nominals, such as the Locative case-marked Perfect Participle –*dOkdA*. These secondary converbs can take possessive person agreement markers in the standard position for possessive suffixes within the NP, i.e. before the case suffix, e.g. *tütsüg yïd–ïn tuy–dok–**umuz–da*** [incense smell–POSS.3SG notice–PTCP–1PL–LOC] 'when we feel the smell of incense' (Erd: 318). In his detailed description of the form and function of the various Old Turkic converbal forms, Erdal (Erd: 308-320, 458ff) does not mention person-marking on the canonical converbs, and he contrasts finite verb forms with non-finite verb forms by saying that the former "normally expresses the person and the number of its subject(s)…" (Erd: 232). I therefore assume that canonical converbs in Old Turkic do not take person marking. The quasi-converbs are an exception to this, since they can take possessive person-marking, as mentioned above.

In Turkish most converbs cannot take person marking (G/K: 95f). Only three suffixes that can be used in converbal constructions (–*DIK,* –*(y)AcAk, and* –*mA*) can take possessive subject agreement markers; however, these suffixes are 'multifunctional subordinating suffixes' that function as masdars, participles, and converbs (G/K: 91ff, 467), i.e. they function as quasi-converbs and can thus not be compared to the Sakha person-marked canonical converbs discussed in section 3.5.1 and 3.5.2. The situation is similar in Uzbek, where unanalyzable converbs cannot take possessive inflection, while case-marked participles take possessive subject agreement markers (Bdr: 589ff, 601ff).

In the South Siberian Turkic languages, verbal forms with a subordinating function can take person-marking when they originate in case-marked participles, while the canonical converbs remain unmarked. The person-marked forms are used in DS constructions, while the unmarked forms are used in coreferential constructions (Gregory Anderson, pers. comm.). However, in Tuvan, one suffix that is traditionally considered a converb (–*GAš*) takes Genitive case-marking (A/H: 56);

furthermore, a complex converb –*BIšaan* combines with subject agreement markers. However, it then functions as the finite predicate of the sentence, e.g. *men öören–mišaan men* [1SG study–CVB 1SG] 'I am still studying' (Isxakov & Pal'mbax 1961: 336ff). Rassadin (Ras: 180) mentions a converb –*GIšA* in Tofa with a meaning of 'until' that supposedly takes possessive suffixes and Dative case-marking, e.g. *kel–giše–vis–ke mïnda olïr* [come–CVB–1PL–DAT here sit[IMP.2SG]] 'sit here until we come!' (Ras: 180). However, this too takes its origin in a participle and takes the possessive suffix in the usual possessive slot for NPs, i.e. before the Dative suffix; furthermore, nowadays it is not used and often not even understood anymore (Gregory Anderson, pers. comm.). It is thus quite clear that the person-marked canonical converbs in Sakha are very distinctive within the Turkic language family.

Interestingly, although she says that there are few differences between Dolgan and Sakha in the converbs, Ubrjatova (Ubr: 162-166) mentions predicative person-marking only for the Perfective Converb –*An* in Dolgan, not for the Imperfective, Purposive, or Privative Converbs. Whether this is an indication that the other converbs do not take subject agreement suffixes in Dolgan is, however, unclear – the lack of information on this in the grammar may just be an oversight of Ubrjatova's, or perhaps these forms occur too rarely for her to have noticed them.

3.5.4 Converbs and switch-reference in Tungusic languages

As described in detail by Igor Nedjalkov (1995), Evenki has a very elaborate system of converbs: he lists 16 different converbal suffixes belonging to different syntactic and semantic types (I. Nedjalkov 1995: 445). These differ in whether in complex sentences they can be used only with coreferential subjects (SS), exclusively with non-coreferential subjects (DS), or variably with either coreferential or non-coreferential subjects (VS). SS converbs do not take any personal subject agreement markers, with the exception of the plural suffix –*l* (118a). DS converbs obligatorily agree in person and number with their subject, which is non-coreferential with the subject of the main clause. The agreement markers used for this purpose are identical to nominal possessive suffixes. Varying subject converbs take possessive person markers when they are non-coreferential with the subject of the main clause, and reflexive possessive suffixes (–*vi* for singular subject, –*vAr* for plural) when they are coreferential. Compare for instance the SS use of the VS Simultaneous Converb –*ŋesi* (118b) with its DS use (118c):

(118a)

ju–la–ver	*eme–**mi–l***	*jep–čo–tin*
house–LOC–PREFL.PL	come–TEMP.CVB–PL	eat–PST–3PL

'Having come home they ate.'

(118b)

Turu–du	*bi–**ŋesi–vi***	*tara–ve*	*sa–ča–v*
Tura–DAT	be–SIM.CVB–PREFL	that–DEF.ACC	know–PST–POSS.1SG

'I knew that when I was/lived in Tura.'

(118c)

Turu–du	*bi–**ŋesi–n***	*tara–ve*	*sa–ča–v*
Tura–DAT	be–SIM.CVB–POSS.3SG	that–DEF.ACC	know–PST–POSS.1SG

'I knew that when he/she was/lived in Tura.'

[I. Nedjalkov 1995: 445f]

With SS converbs, subject agreement marking is unnecessary, since these converbs occur exclusively with coreferential subjects. With DS converbs, subject agreement helps to keep track of the subject of the subordinate clause, since this differs from the subject of the main clause. The same holds true for the VS converbs, where person marking in addition helps to distinguish between coreferential and non-coreferential use of the converb. This system clearly constitutes a system of switch-reference, as defined by Haiman and Munro (Haiman & Munro 1983: ix): "Canonical switch-reference is an inflectional category of the verb, which indicates whether or not its subject is identical with the subject of some other verb." Thus, in Evenki the subject agreement suffixes on converbs have a very clear syntactic purpose, namely to track the referent of subjects in complex sentences.

The converbal system in Èven is similar to that in Evenki, with SS converbs ocurring only in constructions with a coreferential subject with the main clause, while DS and VS converbs occur in non-coreferential constructions (Novikova 1980: 94ff). The four converbs participating in referent tracking by marking the person of the subject of the subordinate clause are cognate to converbs found in Evenki; these are the Conditional-Temporal Converb *–rək*, the Purposive Converb *–dA*, the Simultaneous Converb *–ŋsi*, and the Posterior Converb *–dle*. However, in contrast to Evenki, the latter attaches only to the negative verb *e–*, e.g. *hin **e–dle–s** em–re e–le bi–ji–m* [2SG NEG–POS.CVB–2SG come–CONNEG this–LOC be–FUT–1SG] 'I shall stay here until you come' (Mal: 18).

Like Evenki and Èven, Udihe has a system of syntactic reference tracking functioning through SS, DS and VS converbs. The variable subject converbs are the Purposive Converb, e.g. *bu ə–**ləgə–u** ŋua əniŋə ŋimaŋku–wə ŋimasi–ə–ni* [1PL NEG–PURP–1PL.EXCL sleep mother tale–ACC tell–PST–3SG] 'Our mother told the tale so we would not sleep' (N/T: 237) and the Imperfective Converb (corresponding to the Simultaneous Converb in Evenki and Èven in function and partly in form, with a suffix *–ŋiə*) (N/T: 236f).

Three converbs with different subject or variable subject use are found in both Northern and Amur Tungusic languages. These are the Purposive Converb (variable subject) and the Conditional-Temporal converb (DS), which Sunik reconstructs to Proto-Tungusic (Sunik 1962: 167f, 171), as well as the Simultaneous Converb (variable subject). This latter is not found in all Tungusic languages, but is found in languages belonging to both the Northern Tungusic and Amur Tungusic branches (Sunik 1962: 260). This indicates that the system of switch reference found in Evenki and Èven is a very old and inherited feature. However, a number of converbal suffixes found in Evenki and Èven are not shared between the two languages, indicating a fair amount of independent innovation in this area of the grammar.

3.5.5 Converbs in Mongolic languages

In most Mongolic languages, not even finite verbs take subject agreement markers (Sanžeev 1964: 82, 83f; PopWM: 91). An exception is made by Kalmyk, Dagur, Buryat, Moghol, Monguor and Bonan (Weiers 1977: 313). Amongst these languages, however, Monguor and Bonan stand out in that they do not distinguish between three persons, but only between first and non-first; furthermore, the verbal suffixes do not distinguish number, and finally the suffixes are not derived from the personal pronouns (Weiers 1977: 321). In Kalmyk, Dagur, Buryat and Moghol the verbal person-marking suffixes are derived from personal pronouns that are postposed after the verb; these are obligatory in Dagur, Buryat, and Moghol, but optional in Kalmyk (Weiers 1977: 321: 313, 320). Postposed personal pronouns to mark subject agreement were optional in the written sources from the 13th to the beginning of the 18th century, after which period they were replaced entirely by a lack of person agreement in most Mongolic languages/dialects (Weiers 1977: 312f, 320). The disappearance of optional subject agreement marking that took place in the written Mongolic sources at the beginning of the 18th century may have been stimulated through the translation of Buddhist texts, the originals of which were written in isolating languages (Weiers 1977: 323). Similarly, the development of obligatory person-marking in Moghol, Dagur and Buryat could be attributed to contact influence; thus, Moghol (which is highly endangered) is spoken in Afghanistan in Indo-Iranian surroundings, where it has undergone substantial structural changes through contact influence (Weiers 1977: 312; 2003: 248). Dagur and Buryat, on the other hand, although they are not spoken in geographical proximity, are both in contact with Evenki. On the other hand, since in older Mongolic texts subject agreement marking on verbs is optionally present, it may also

be that these few dialects/languages just took the person marking that was incipient in Mongolic as a whole to a further stage (Uwe Bläsing, pers. comm.).

Even if the development of obligatory subject agreement marking on finite verbs in Buryat and Dagur may not be attributable to Evenki influence, I feel that Evenki contact has clearly influenced the converbal system of these languages: converbs in both Buryat and Dagur attach possessive person markers to reference the subject of the subordinate clause (Skr: 118; Tsu: 146). This clearly parallels the Evenki system, while it is not found in most of the other Mongolic languages, both historical and modern. For example, in Khalkha neither finite verbs, nor converbs take any person or number marking (K/Ts: 113, 156).

In Buryat, the Final, Intentional and Comparative Converbs occur only in SS constructions and do not take person marking (Skribnik 1988: 143; Skr: 117). The Modal, Imperfective and Perfective Converbs occur mainly in SS constructions, e.g. *ežii morin deere–hee harabšala–**n** udaan xara–na* [mother horse.OBL top–ABL screen.eyes.with.hand–MOD.CVB long.time look–DUR] 'the mother looks into the distance from horseback, screening her eyes with her hand' (Skribnik 1988: 145; Skr: 116). The Modal Converb can take reflexive possessive person marking (Skr: 116, Table 5.8), while the other two converbs do not take person marking. The remaining converbs take possessive or reflexive-possessive subject-agreement markers; the possessive person markers are used in DS constructions, and the reflexive-possessive person markers are used in SS constructions (PopB: 70), e.g. in a DS construction: *tende xüre–že ošo–**tor–nay** dayn baldaan duuha–xa yohotoi* [there reach–IPF.CVB go–TERM.CVB–POSS.1PL enemy.OBL ?? end–FUTPT probably] 'by the time we get there the war will surely be over'; in a SS construction: *Butedmaa teren–iiyi tani–**xalaar–aa** bayarla–sha–ba* [B. that.OBL–ACC recognize–SUCC.CVB–PREFL be.glad–INTS–TERM] 'recognizing him, Butedmaa was glad' (Skr: 116f). The Conditional and Abtemporal Converbs can occur in both DS and SS constructions; in the former, they take possessive suffixes to mark agreement with the dependent subject, while in SS constructions they remain unmarked (Skribnik 1988: 152). The Terminative, Contemporal, and Successive Converbs take possessive marking in DS constructions and reflexive possessive marking in SS constructions (Skribnik 1988: 149). It thus appears that the Buryat converb system functions in a manner very similar to that found in Evenki, with converbs functioning predominantly or solely as SS markers remaining unmarked, while converbs that occur in DS constructions take possessive suffixes to express their subject; some converbs (classifiable as VS converbs; Skribnik 1988: 142) take reflexive-possessive suffixes when their subject is coreferential with the subject of the main clause.

While it is difficult to judge whether the Dagur converbs function in a comparable way to the Evenki and Buryat system, Tsumagari (Tsu: 146) does say that there are two types of converbs in Dagur, of which one type (which he calls 'simple') are used with coreferential subjects, while the other type (called 'clausal') can occur with non-coreferential subjects. "In different-subject constructions the clausal converbs, like many quasiconverbs, can contain personal marking by the possessive suffixes." (Tsu: 146). Two examples indicate that there are at least parallels to the Tungusic system in the way clausal converbs mark the subject agreement in DS constructions and in SS constructions: DS: *geri–d–ee kucir–***gweeteer–miny**[3] *xwar war–j eurkee–seng* [house–DAT–PREFL ??–SUCC.CVB–POSS.1SG rain enter?–IPF.CVB ??–PFV] 'soon after I came home, it began to rain'; SS: *saw–**oojaar–aa** wantaa tali–seng* [sit–PROG.CVB–PREFL sleep put–PFV] 'while he was sitting, he fell asleep' (Tsu: 147). Thus, the possessive suffix is used in the DS construction, and the reflexive possessive suffix is used in the SS construction.

In Kalmyk, the Terminative and the Successive Converbs can take possessive and reflexive possessive suffixes to express agreement with the dependent subject (Bläsing 2003: 244). In Spoken Oirat, too (though not in Written Oirat), the Terminative Converb can take possessive person marking to express the (non-coreferential) subject (Birtalan 2003: 225). Interestingly, in Oirat and Kalmyk the suffix of the Terminative Converb (with a meaning of 'until', and also 'while' in Kalmyk) is *–tl*, derived from *–tAlA*; this is very similar in both form, function and meaning to the Evenki converb suffix *–dAlA* and the Ėven converb suffix *–dle*, which also occur in variable subject constructions and denote posterior situations, with a meaning of 'until' (I. Nedjalkov 1995: 452f; Mal: 18). Janhunen reconstructs this converb to Proto-Mongolic (Janhunen 2003d: 25), so that it might represent a Mongolic copy in Northern Tungusic.

3.5.6 Converbs in Siberian languages

Kolyma Yukaghir has both SS and DS converbs, but no variable subject converbs. There are two DS converbs, one with a general temporal meaning, the other with a conditional meaning. These DS converbs derive their origin from Locative case forms of the Action Nominal. They make a person distinction between 1/2, 3SG and 3PL, i.e. there is no distinction in number or person for the first and

---

[3] However, it appears that the suffix *–gweeteer* is at least historically derived from the future participle (Tsu: 145); it is thus not entirely clear whether this can be classified amongst the canonical converbs, or whether it should rather be regarded as a quasi-converb.

second person. There are six different SS converbs that do not take any person marking: the Imperfective, Perfective, Iterative, Conditional, Privative, and Connective Converbs (Mas: 158ff). The Connective Converb takes a suffix identical to the Comitative case suffix –ńit, while the Privative Converb is formally identical to the nominal Privative form (Mas: 165). The DS converb with a general temporal meaning functions as a switch-reference marker in clause chains, alternating with different SS converbs (Mas: 370), while the DS converb with conditional meaning alternates with the SS Conditional converb in conditional chains (Mas: 374f). Thus, the Yukaghir system is similar to the Evenki system by having converbal forms that track the referent of each (co)subordinate subject in a chain of clauses; it differs, however, in that the person agreement of the DS converbs is rather restricted and is not performed by possessive suffixes (which are lacking in Yukaghir, cf. section 3.2.1.4), and also in that there are no VS converbs.

Chukchi has three converbs, none of which takes person marking. Although they preferentially occur in SS constructions, this is not obligatory; the subject reference of a non-coreferential subordinate clause can only be inferred from the context (Dnn: 240, 245f). In Itelmen, several non-finite verb forms (called Infinitive I through VI by Georg & Volodin) are found with different functions, such as infinitive or participial; none of these are comparable to converbs in the languages discussed so far. Only the Infinitive I can take subject agreement suffixes, albeit with restricted person-number complement; these turn it into a finite form (G/V: 167, 170f).

Siberian Inupik Eskimo has eight different converbs, seven of which take both intransitive subject agreement and transitive subject-object agreement suffixes, while one takes only intransitive subject agreement. The transitive subject-object agreement markers are identical to the transitive person markers found on finite verbs (with a few slight differences), while the intransitive subject agreement markers are more or less identical to the possessive markers, with the exception of the dual number forms. Nearly all of the converbs can occur in both SS and DS clauses, taking the same subject agreement markers in both cases. There is just one exception: the 3rd person agreement markers differ between coreferential and non-coreferential uses, i.e. when the 3rd person subject of the converb clause is coreferential with the subject of the main clause, one subject agreement suffix is used, while a different subject agreement suffix is used when the 3rd person subject of the converb clause is non-coreferential with the subject of the main clause (be that a 1st, 2nd or 3rd person subject). Converbs are used in adverbial and subordinate clauses, but also as the content-carrying verb in analytical forms with an auxiliary conveying the grammatical function (Men: 142-162).

In Nivkh, most finite verbs do not take any subject agreement marking (Mat: 21). There are a large number of different converb forms that function to combine clauses as well as occurring in combination with auxiliaries. Twelve of these converbs are invariant, i.e. they do not take any subject agreement markers; while two converbs stand out in the verbal paradigm in that they have two different suffixes according to the number and person of the converbal subject. However, these person marking suffixes are rather minimal: one suffix (–*t* for the 'General/Manner Converb', –*tot* for the 'Anterior/Temporal Converb') marks subject agreement with the 1st person, 2PL and 3PL, while a second suffix (–*r* for the 'General/Manner Converb', –*ror* for the 'Anterior/Temporal Converb') marks agreement with 2SG and 3SG (Mat: 23f; Grz: 55). The converbs showing a form of subject agreement can occur in SS (119a) and DS (119b, c) constructions; in the latter case, the converb additionally takes the Causative suffix –*g(u)* (which does not add a meaning of actual causation). Unexpectedly, the converb agrees not with the subject of the subordinate clause, but with the subject of the main clause (119b, c; Grz: 55; Ekaterina Gruzdeva, pers. comm.). The invariant converbs can occur in DS constructions as well; without, however, indicating the switch reference (119d; Mat: 32).

(119a)

ki	*xiz–roř*	*kiɣvuř*	*yup–t*
footwear	put.on–CVB.2/3SG	lace	tie.up–FIN

'After putting on (his) shoes, he tied up (his) laces.'

(119b)

ńi	*vi–g–roř*	*pʰeɣrdoχ*		*nudoχ*	*pʰur–ya*
1SG	go–CAUS–CVB.2/3SG	PTL		whatever	say–IMP

'After I leave, say whatever (you like)!'

[Grz: 55]

(119c)

ńi	*pʰřə–g–r*		*ezmu–d*
1SG	come–CAUS–CVB.2/3SG		rejoice–FIN

'He was happy that I came.' (lit. 'He rejoiced letting me come.')

[Mat: 32]

(119d)

či	*ń–ro–laχ*	*(ńi)*	*če–rχ*	*ŋixə–d–ra*
2SG	1SG–help–CVB:for	(1SG)	2SG–ALL	thank–FIN–FOC

'I thank you for helping me.'

[Mat: 33]

Ket does not have specialized converbal forms. The Perlative case of person-marked verbs that additionally may carry the past tense suffix functions as a sort of adverb, as do some of the infinitive forms. However, this part of Ket verbal grammar has not been studied in sufficient detail, so that it is not known what the syntactic functions of the person marking are (Andrej Nefedov, pers. comm.).

Nganasan has only two converbal forms; one is invariant and does not take any subject agreement suffixes, while the other takes Genitive case suffixes of the possessive declension. The invariant converbal form has a very broad meaning, expressing both anteriority and simultaneity as well as functioning as an adverbial modifier. Although Tereščenko does not say this explicitly, from the examples it appears that this form is restricted to SS constructions (Ter: 274ff). The converbal form taking person-marking has a predominantly conditional as well as a temporal meaning; this appears to be used in both SS and DS constructions (Ter: 277ff). Furthermore, the supine suffix in Nganasan takes Genitive case suffixes of the possessive declension; from the examples given, it appears to be restricted to SS constructions (Ter: 272ff).

Mansi has two verb forms that can express the predicate of adverbial clauses; these take possessive person marking. Rombandeeva (Rmb: 147) classifies them as converbs and says they are homonymous with participles, while Riese (Rse: 65ff) calls them participles, pointing out, however, that since "… some of the participles […] are often used as adverbials in sentences there is no sharp dividing line between them and the gerunds" (Rse: 65). Since the verb forms classified by Riese as participles occur widely in attributive function, and have converbal function in conjunction with case-marking or with postpositions, they appear to be classifiable as quasi-converbs rather than as participles or converbs. One verb form functions as an adverbial of manner and as such does not take any person-marking; however, it can also function as a temporal adverbial when marked by possessive suffixes and/or case suffixes or postpositions (Rse: 69). Unfortunately, since Riese gives only examples for adverbial clauses without the accompanying main clauses for all (quasi-)converbs, it is not clear whether the possessive-marking is used specifically in DS constructions.

In Khanty as well, there exists an invariant converb and two participles that can function as quasi-converbs. The converb can function as the subordinate predicate of adverbial clauses in SS constructions, and it also combines with other verbs to express resultative aspect (Nik: 46). The participles can also function as the predicate of adverbial clauses; in this case they are followed by a postposition or take Locative case-marking (Nik: 47). Participles in this function can take optional subject agreement markers that agree with the topicalized subject. These can be

added in SS constructions as well, since here the embedded subject is generally dropped, e.g. *o:pe:–m xo:ll–ə–**t–al**–na nu:ms–ə–s* [sister–1SG cry–E–PTCP–3SG–LOC think–E–PST.3SG] 'My sister was thinking while crying' (Nik: 48).

As can be seen from the preceding discussion, Turkic languages do not take person-marking on canonical converbs, but use possessive suffixes to mark subject agreement on quasi-converbs derived from participles (cf. section 3.5.3). Most Mongolic languages do not even mark subject agreement on finite verbs, let alone on converbs. However, Buryat and Dagur constitute a marked exception, as they use pronominal person marking on finite verbs and possessive suffixes to mark agreement with the subordinate subject on converbs. The converbs that can take subject agreement suffixes make a distinction between possessive suffixes in DS constructions and reflexive possessive suffixes in SS constructions. Kalmyk and Spoken Oirat show person marking on a very limited number of converbs (two in Kalmyk, one in Oirat; cf. section 3.5.5). Tungusic languages use person-marked converbs to mark the subject of the subordinate clause (cf. section 3.5.4). Here, converbs that occur only in SS constructions remain unmarked, except for some that take plural agreement suffixes; those that occur only in DS constructions take possessive person markers, while converbs that occur both in DS and SS constructions take possessive suffixes in the former and reflexive possessive suffixes in the latter. The source of the use of possessive suffixes to mark subordinate subject agreement in Buryat and Dagur may well be due to Evenki influence.

Ket and Itelmen do not seem to have specialized converbal forms, while the three converbs found in Chukchi are always invariant, even when used in DS constructions (cf. section 3.5.6). Mansi and Khanty appear to have only quasi-converbs taking person agreement suffixes, while the invariant converb occurs only in SS constructions. In Nganasan, an invariant converb appears to occur only in SS constructions, while a converb marked with Genitive suffixes of the possessive declension to express subject agreement occurs in both SS and DS constructions. Both Yukaghir and Nivkh have several converbal forms, of which the majority take no person suffixes; however, in both languages there are two converbs that take a limited set of subject agreement markers. In Yukaghir, these two person-marked converbs occur only in DS constructions, alternating with the invariant SS converbs to track syntactic reference; in Nivkh, the person-marked converbs can occur in both SS and DS constructions (cf. section 3.5.6). Finally, in Siberian Inupik Eskimo there are several converbs that take both intransitive (subject) agreement and transitive (subject-object) agreement suffixes. The intransitive agreement suffixes are basically identical to possessive suffixes, while the transitive agreement suffixes are basically identical to the subject agreement markers found on finite verbs. Converbs occur in

both SS and DS constructions without making a difference in the person marking (with the exception of the third person).

It thus becomes clear that not very many Siberian languages permit person marking on canonical converbs. Furthermore, possessive person marking to express the subject of the (co-)subordinate clause is restricted to the Tungusic languages, to Buryat and Dagur (and very marginally Kalmyk and Spoken Oirat), to one converb of Nganasan, and to the intransitive subject agreement markers in Siberian Inupik Eskimo.

3.5.7 The origins of Sakha person-marked converbs

From the above, it has become very clear that Sakha differs from its Turkic relatives by permitting person-marking on canonical converbs. This feature is quite rare amongst Siberian languages as a whole; however, it constitutes a characteristic feature of the Tungusic language family. It is therefore not surprising that Evenki contact influence has been proposed to account for this feature in Sakha. However, there are two weighty arguments against Evenki influence playing any role in the development of person-marked converbs in Sakha.

Firstly, in Evenki DS converbs take possessive person marking, while nearly all Sakha converbs take predicative person marking. If Sakha speakers had started marking subject agreement on converbs under Evenki influence, one might have expected them to use possessive person marking as well. Possessive subject agreement markers are used frequently in Sakha TAM forms – for instance, the Recent Past, Imperfect, and Future tense and the Assertive mood all take possessive person marking. Therefore, converbs marked with possessive person suffixes would not clash with the general structure of verbal paradigms in this language. Phonologically, too, there is no reason to prefer predicative person markers over possessive ones: *kel–en–im* [come–PF.CVB–POSS.1SG] is as acceptable a phonological sequence as *kel–em–min* [go–PF.CVB–PRED.1SG]; cf. *kergen–im* [spouse–POSS.1SG]. Lastly, possessive subject agreement markers would actually be more 'efficient' than predicative ones, since with the help of possessive markers it would be possible to mark 3SG subject agreement as well; as it is, converbs agreeing with 3SG subjects are indistinguishable from unmarked converbs, since the 3SG predicative marker is zero.

Of course, it may have been the case that speakers of Sakha were not aware of the difference between possessive person marking and predicative person marking when copying person marking on converbs from Evenki, but that they rather just copied the fact of marking subject agreement as such (Bernard Comrie,

pers. comm.). However, in both Sakha and Evenki the suffixes that mark agreement with the nominal possessor in possessive constructions are fully identical to the subject agreement suffixes of certain tense forms, as exemplified in Table 3.18. In Sakha, the verbal subject agreement marking is identical to the nominal possessive suffixes in the Remote Past, the Imperfect, and the Future, and in the Assertive mood. In the Recent Past, the third person suffixes diverge a bit: if one assumes the Recent Past tense suffix to be –TI (the form it has for the first and second person), one would expect the 3SG form to be –TItA and the 3PL form to be –TIlArA; instead, they are –TA and –TIlAr, respectively. In Evenki, the Past Indefinite, Past Iterative, and Future Categorical tense-aspect forms take subject agreement markers that fully coincide with the nominal possessive suffixes, while the present tense of the auxiliary bi 'to be' (bi–si) takes possessive subject agreement markers only in the 1PL and 2PL, while in the optative mood the 1PLincl and 3PL take divergent suffixes: –p and –n, respectively (Ned: 260, 263).

Table 3.18: Nominal possessive suffixes and verbal suffixes of possessive origin in Sakha and Evenki

	Sakha		Evenki	
	Nominal ('horse')	Verbal ('find', Imperfective)	Nominal ('house'; Ned: 143)	Verbal ('find', Past Indefinite; Ned: 260)
1SG	at–ïm	bul–ar–ïm	ju–v	baka–ča–v
2SG	at–ïŋ	bul–ar–ïŋ	ju–s	baka–ča–s
3SG	at–a	bul–ar–a	ju–n	baka–ča–n
1PL(ex)	ap–pït (at–BIt)	bul–ar–bït	ju–vun	baka–ča–vun
1PL(in)			ju–t	baka–ča–t
2PL	ak–kït (at–GIt)	bul–ar–gït	ju–sun	baka–ča–sun
3PL	at–tara (at–LArA)	bul–al–lara	ju–tïn	baka–ča–tïn

The other set of subject agreement markers found on most tense and mood forms in Sakha and in Evenki differ in form from the possessive markers (cf. Table 3.19; see also Table 3.11 in section 3.3.1.1). In Sakha, they are identical to the subject agreement marking on nominal predicates, while in Evenki they appear to be unifunctional: they are not derivable from personal pronouns (Ned: 200f, 259), and Evenki does not have subject agreement marking on nominal predicates (Ned: 59). The 1PL and 2PL possessive and predicative suffixes in Sakha are homonymous, while in Evenki the 3SG possessive suffix is homonymous to the 'non-possessive' verbal subject agreement marker. In Sakha, the Present Tense, the Resultative Past,

the Episodic Past, the Necessative I, the Voluntative-Potential, the Conditional I, the Habitual, and the Presumptive take predicative subject agreement suffixes. In Evenki, the subject agreement suffixes that are not identical to possessive suffixes occur with the Present Tense, Non-future, and two different future tenses (Ned: 259). In addition, the present tense of the auxiliary *bi* 'to be' (*bi–si*) takes non-possessive subject agreement markers in the singular number and 3PL (in this case no suffix; cf. Table 3.19).

Table 3.19: Verbal agreement markers not derived from possessive suffixes in Sakha and Evenki

	Sakha		Evenki
	nominal predicate ('teacher')	verbal ('find', conditional I)	verbal ('find', future; Ned: 259)
1SG	*uču:tal–bïn*	*bul–lar–bïn*	*baka–ja–m*
2SG	*uču:tal–gïn*	*bul–lar–gïn*	*baka–ja–nni*
3SG	*uču:tal*	*bul–lar*	*baka–ja–n*
1PL(excl)	*uču:tal–lar–bït*	*bul–lar–bït*	*baka–jara–v*
1PL(incl)			*baka–ja–p*
2PL	*uču:tal–lar–gït*	*bul–lar–gït*	*baka–ja–s*
3PL	*uču:tal–lar*	*bul–lal–lar*	*baka–jara*

In contrast to Sakha and Evenki, in Turkic languages the nominal possessive suffixes and the verbal subject agreement suffixes do not overlap fully, but differ in the third person and in the first person plural. Thus, the 3SG verbal suffix is generally zero, and for the third person plural it is often just the plural suffix *–LAr*, or zero as well. The possessive suffix for the third person singular, however, is generally *–(s)I*; for the third person plural it is often identical to the third singular, or else *–LArI* (i.e. the plural suffix followed by the 3SG possessive suffix). The 1PL verbal agreement suffix is generally *–k*, as compared to *–(I)mIz* or *–(I)bIz* in nominal possessive constructions (G/K: 88ff; Clk: 214; Somfai Kara 2002: 38f; Imr: Tableau P in 828f, 1777, 1783; Bdr: 641, 659-709; And: 25; A/H: 39). Furthermore, in the Turkic languages other than Sakha, the subject agreement suffixes that are derived from the nominal possessive suffixes occur only with the recent past *–DI* and with the conditional *–sA*, while in Sakha they take a much more important place in the verbal system, occurring in the Recent Past, the Remote Past, the Imperfect, the Future, and in the Assertive mood. Admittedly, the Remote Past, the Imperfect, and the Future are all formed with participles (past, present and future participle, respectively), which may explain the use of possessive subject agreement markers.

On the other hand, the Resultative past and the Habitual mood are also derived from participles, and yet both take predicative subject agreement suffixes. Thus, in my opinion the fact that the nominal possessive suffixes are so clearly identical to the possessive subject agreement markers in both Sakha and Evenki, and the prominence of possessive person markers in Sakha indicates that Sakha speakers may well have been able to distinguish which suffixes are being used by Evenki speakers to mark agreement on converbs with the subordinate subject.

Secondly, the Evenki person-marked converbs fulfill a very specific syntactic function, since they occur within the Evenki switch-reference system, marking subjects that are non-coreferential with the subject of the main clause (cf. section 3.5.4). Sakha person-marked converbs, on the contrary, occur predominantly in (co-) subordinate clauses that are coreferential with the main clause subject (cf. section 3.5.2), while in the system of clausal reference tracking it is case- and person-marked participles that fulfill the role of DS markers. That it would be in theory possible to make use of person-marked converbs to track clausal subject reference is shown by the immediate-precedence converb –*A:t*, which is the only Sakha converb functioning in this way: here, the converb takes no person-marking when it is coreferential with the subject of the main clause, while in non-coreferential clauses it takes Possessive-Accusative suffixes plus the postposition *kïtta*. Compare examples (113a) and (113b), repeated here for convenience as (120a, b):

(120a)

min	*jieber*	*kele:t*	*küöspün*
min	jie–BAr	kel–A:t	küös–BIn
1SG	house–DAT.1SG	come–IMM.CVB	pot–ACC.1SG

küöstü:bün
küös–LA:–A–BIn
pot–VR–IPF.CVB–PRED.1SG

'As soon as I come home I start cooking.'   [translation, Ver]

(120b)

jieber	*kele:ppin*	*kïtta*	*iỹem*
jie–BAr	kel–A:t–BIn	kïtïn–A	iỹe–(I)m
house–DAT.1SG	come–IMM.CVB–ACC.1SG	join–IPF.CVB	mother–POSS.1SG

ebieti	*belemne:bitinen*	*barar*
ebiet–(n)I	belem–LA:–BIt–(t)InAn	bar–Ar
lunch–ACC	ready–VR–PSTPT–INS.3SG	INCP–PRSPT

'As soon as I come home, my mother starts cooking lunch.'

[translation, Tat]

It is notable that in this case the converb takes possessive subject agreement markers, not predicative ones; i.e. this one case of a Sakha person-marked converb coincides surprisingly well with the Evenki system. That languages can develop person agreement marking on converbs under Evenki influence is shown by Buryat and Dagur; in both of these languages, the subject agreement markers are possessive and reflexive possessive suffixes that appear to perform the same function as Evenki person-marked converbs, namely to keep track of the subordinate subject in complex clauses (cf. section 3.5.4). Since the Sakha Immediate Precedence Converb functions in the same way as the Evenki converb system, one might postulate Evenki influence in its development, similar to the influence that led to the development of possessive-marked converbs in Buryat and Dagur. This assumption is further strengthened by the fact that in Evenki there is a (nearly obsolete) VS converb that used to express immediate anteriority, e.g. *dolbo Amarča asini–**ktava**–n amin–in ju–la–n i–re–n* [night A. fall.asleep–IMM.CVB–3SG father–POSS.3SG house–LOC–3SG enter–NFUT–3SG] 'in the evening as soon as Amarcha fell asleep, his father entered his house' (I. Nedjalkov 1995: 449).

Although it is therefore fairly clear that Evenki contact influence was not the source of Sakha person marking on converbs, it is still possible that this feature is due to influence from some other language. However, as was discussed above, northern Eurasian languages that permit subject agreement marking on canonical converbs mostly do so with the help of possessive suffixes, and thus cannot be considered the source of person-marking on Sakha converbs any more than Evenki. The only exception amongst the languages examined here is Siberian Yupik Eskimo, where the transitive subject-object agreement markers basically correspond to the subject-object agreement suffixes of the finite verb. However, since speakers of Sakha and speakers of Siberian Yupik Eskimo were never in any longterm contact it is rather unlikely that this could be the source of influence, either.

Thus, it would appear that the use of person-marked converbs is a Sakha innovation, and not the result of Evenki influence, contrary to the opinion of Johanson and Ubrjatova (Johanson 2001: 1732; Ubrjatova 1976: 45), nor the result of contact influence from any other language. The only instance where Evenki influence may have played a role is in the development of the person-marking to track subject reference in the case of the Immediate Precedence Converb, since this differs from the other Sakha converbs in both form and function, while it mirrors the form and function of the Evenki person-marked converbs to track subject reference. Furtheremore, there exists an Evenki VS converb with immediate precedence meaning.

4 SUBSTANCE COPIES AND PHONOLOGICAL INFLUENCE IN SAKHA

Although the focus of this study is on possible schematic copies in Sakha, a discussion of language contact influence would not be complete without a consideration of the substance copies and the phonological influence found in the recipient language. This chapter therefore gives a brief overview over the number and kinds of substance copies and some phonological changes found in Sakha, based mainly on previous studies by other authors (cf. section 1.3).

4.1 Substance copies and phonological influence from Mongolic

The absolute number of lexical substance copies from Mongolic in Sakha is claimed to be very high, between 2,000 (Popov 1986: 8) and 2,500 (Rassadin 1980: 65). The very comprehensive Sakha-Russian dictionary compiled by Pekarskij ([1907-1930] 1958-1959) contains approximately 6,200 lexical roots (out of a total of more than 25,000 lexical items; Popov 1986: 7); thus, if one takes the number of lexical roots as the base against which to estimate the proportion of substance copies in the language, 30-40% of the Sakha roots might have been copied from Mongolic. Rassadin (1980: 92), however, criticizes this high number and derives a frequency of Mongolic substance copies of only 10% by taking the number of lexical items covered in the *Jakutsko-Russkij Slovar'* (edited by Slepcov 1972), namely approximately 25,000, as the basis for his calculations. However, this dictionary contains a large number of derivations, e.g. *üle* '(the) work', *üle–le:* [work–VR] 'to work', *üle–le–n* [work–VR–REFL], *üle–le–s* [work–VR–REC], *üle–le–t* [work–VR–CAUS], *üle–le–t–i:* [work–VR–CAUS–NR], *üle–le:χ* [work–PROP] 'labour-intensive; busy', and *üle–hit* [work–AGNR] 'worker' (Slepcov 1972: 452f). Since the estimates of number of substance copies are based on individual roots (i.e. *üle* is counted as one copy from Mongolic), and the derivations from these roots are not counted, using the total number of entries in this dictionary as the basis for calculations leads to an underestimate of the proportion of substance copies from Mongolic. On the other hand, a survey of approximately 1,500 lexical meanings (compiled for the 'Loanword Typology Project' at the MPI for Evolutionary Anthropology in Leipzig) counts only 187 items probably or clearly copied from Mongolic (Pakendorf & Novgorodov in preparation). This, too, amounts to only 13% of substance copies from Mongolic in the Sakha lexicon. The discrepancy between the estimates of Mongolic substance copies based on the lexical roots in Pekarskij and those derived from the 'Loanword Typology' meaning list can be at least partially explained by the fact that a number of the substance copies from Mongolic are descriptive verbs,

such as *amčïɣ* 'have deeply sunken lips', or *belgey* 'be swollen (of face)' (Kałużyński 1962: 130, 131). This kind of descriptive verb is not included in the meaning list of the Loanword Typology Project, and therefore a whole domain of copied words is excluded.

Leaving aside the issue of the proportion of substance copies from Mongolic in the Sakha lexicon, it is interesting to note that these copies are not restricted to cultural items, which are generally considered to be easily copied, but include a large number of items from the so-called basic lexicon as well. Thus, we find a number of kinship and body part terms copied from Mongolic, as well as verbs, both of the descriptive kind mentioned above, but also with a less specialized meaning, such as *orguy* 'boil' (from Mongolic *orgi* 'spurt, gush forth, boil up'), *teniy* 'stretch' (from Mongolic *teni* 'unbend, become straight, stretch'), or *ergiy* 'turn around' (from Mongolic *ergi* 'turn or move around, revolve'; Kałużyński 1962: 148, 55, 57).

From the types of substance copies it would seem that the Mongols had a socially dominant role (cf. Kałużyński 1962: 120), being in a position to prohibit actions, enforce laws, make decisions, and also having concubines and servants. Thus, amongst others we find the following copies or derivatives of copies of Mongolic words in Sakha: From the domain of law and order: *ba:* 'force, fine; accusation' (from Mongolic *baɣa* 'fine, penalty'; Kałużyński 1962: 131, Lessing 1995: 67), *kuolu* 'order, command; tradition, rule, law' (copied from Mongolic *xauli* 'usage, custom; law, rule, regulation; codex; punishment'; Kałużyński 1962: 36; Lessing 1995: 946), and *buruy* 'guilt' (copied from Mongolic *buruɣu* 'error, mistake, guilt'; Kałużyński 1962: 38). From the domain of social relations we find: *ojolu:n* 'concubine' (probably derived from Mongolic *oju* 'to kiss'; Kałużyński 1962: 148, Lessing 1995: 626), *jehel* 'servant, messenger'; probably derived from Mongolic *jese* 'put in order, arrange, prepare', *jeselge* 'something prepared, supplies, preparation'; Kałużyński 1962: 139, Lessing 1995: 1047), *simeχsin* 'old woman, old servant' (copied from Mongolic *šibegčin* 'maid servant'; Kałużyński 1962: 41), and *noχo:* 'perjorative form of address for a boy or man' (copied from Mongolic *noxay* 'dog'; Lessing 1995: 592; cf. example (67a) in section 3.3). However, Antonov (1971: 121) suggests that the large amount of Mongolic copies referring to law and order and to poor people, servants, and the like is an indication that the Mongolic-speaking groups were the socially subordinate ones: "This poor, working part of the population introduced into the Yakut language Mongolian words for the poor layers of society, the secondary wives, concubines, servants, for the children which it

looked after, and a significant number of terms of common law, being the most frequently convicted and punished party." (translation mine[1]).

Furthermore, the Sakha have adopted a large part of their livestock terminology, including food products derived from milk, from the Mongolic-speaking tribe or tribes they were in contact with; thus we find *süöhü* 'livestock' (copied from Mongolic *jögeri* 'property, household goods, chattle; possessions, pack-animal'; Kałużyński 1962: 35), *dal* 'corral' (from Mongolic *dal* 'roof, shelter, barn'; Kałużyński 1962: 44), *meččiy* 'graze' (from Mongolic *belči* 'graze'; Kałużyński 1962: 40), *süögey* 'cream' (from Mongolic *jögekey* 'cream'; Kałużyński 1962: 35), and *umda:n* 'buttermilk with water' (from Mongolic *umdaɣan* 'a drink'; Kałużyński 1962: 19).

Interestingly, a number of terms for body parts and illnesses are of Mongolic origin; for example *kieli* 'womb' (copied from Mongolic *kegeli* 'womb, pregnancy, belly'; Kałużyński 1962: 29), *bïlčarχay* 'gland' (copied from Mongolic *bulčirxa*y 'gland' (Kałużyński 1962: 19), *soɣuo* 'goitre' (from *saxaɣu* 'glanders, farcy'; Kałużyński 1962: 31), *kïjïk* 'epidemic, plague' (copied from Mongolic *kijig* 'epidemic, plague'; Kałużyński 1962: 26), *omuru:n* 'infection of the mouth' (from Mongolic *amarau* 'pustules in the mouth, cankers'; Kałużyński 1962: 148, Lessing 1995: 36), *χamsïk* 'plague' (copied from *ɣamčiy* 'illness, misfortune'; Kałużyński 1962: 41), and *jaŋ* 'epidemic' (copied from the Mongolic adjective *yaŋ* 'foreign', which occurs in the phrase *yaŋ yara* 'syphilis, syphilitic sore'; Kałużyński 1962: 50, Lessing 1995: 427). On the one hand, this could be an indication that the Mongolic-speaking groups the Sakha were in contact with had a higher degree of anatomical and medical knowledge than the Sakha themselves. However, another explanation is that the original Sakha names for illnesses were taboo, as they are nowadays, so that copied terms were adopted to avoid having to use the indigenous names. This assumption is supported by the fact that some of the terms for illnesses are copied from Buryat and thus stem from a later time period, possibly due to the necessity of having to renew taboo lexicon. Thus, *albaχ* 'epidemic, infectious disease' appears to be copied from Buryat *alban* 'pox', and *kumakï/kumaχa* 'fever' appears to be copied from Buryat *χumχa:* 'fever' (Kałużyński 1962: 144, 145).

Apart from lexical copies, Sakha has also copied a number of Mongolic suffixes, some of which have become productive in the recipient language. These suffixes were probably initially copied together with the Mongolic stems and only

---

[1] Original: "Эта бедная, трудящаяся часть населения внесла в якутский язык монгольские слова о бедных слоях населения, второстепенных женах, наложницах, челяди, о детях, за которыми она ухаживала и значительное количество терминов обычного права, как наиболее часто судимая и наказуемая сторона."

later became segmented and productive in their own right. For example, Sakha copied the verb *χata:* 'hammer in (a nail)' from Mongolic *qada* 'hammer in', and it also copied the derived noun *χata:hïn* 'nail, bolt' from Mongolic *qadaγasun* 'nail'. Later, this suffix *–A:hIn* (Mongolic *–γasun*) became productive in Sakha; thus, we find *kurda:hïn* 'circle' from *kur–da:* [belt–VR] 'put on belt' (from the Turkic root *kur* 'belt'), and *argumenna:hïn* 'argumentation' from *argumen–na:* [argument–VR] 'to argue' (from the Russian copy *argument* 'argument'; Kałużyński 1962: 71). According to Kałużyński (1962: 70), the highly productive habitual suffix *–A:ččI* is copied from Mongolic *–γači*; however, Ubrjatova (1972: 584) and Širobokova (1977: 115) suggest that it is inherited, because it is found in other Turkic languages (e.g. Yenissey Kirghiz) as well. Other copied productive nominalizing suffixes are *–GAy* (copied from Mongolic *–GAy*; Kałużyński 1962: 76f), *–lGA(n)* (copied from Mongolic *–lγa(n)*; Kałużyński 1962: 85–*mńI* (from Mongolic *–mji*; Kałużyński 1962: 90), *–mtAGAy* (copied from the identical Mongolic form; Kałużyński 1962: 92), and *–mtIA* (copied from Mongolic *–mtaγu*), which might have been copied independently of any lexical forms, as it is not found with any Mongolic copies (Kałużyński 1962: 93). A large number of suffixes that derive onomatopoetic and descriptive verbs were copied from Mongolic (Kałużyński 1962: 97ff). Furthermore, Kałużyński suggests that the suffix of the Immediate Precedence converb *–A:t* was copied from Mongolic *–γad*, and that the suffix *–(č)čI*, which derives adverbs from verbs, was copied from the Mongolic imperfective converb *–ču/–ju* (Kałużyński 1962: 112f). Finally, in Sakha a number of nouns ending in *–n* and denoting humans have a plural in *–ttAr*, e.g. *toyon* 'master', *toyottor* 'masters', *oyu:n* 'shaman', *oyu:ttar* 'shamans' (cf. footnote 10 in section 3.4.3.1). It has been suggested that this plural is in actual fact a double plural marking, consisting of the copied Mongolic plural suffix *–t* (which is mainly added to nouns ending in *–n*) followed by the Turkic plural *–LAr*. Since two nouns at the beginning of the century occurred without the additional Turkic plural suffix in a collective meaning (*toyot* 'gentlemen' and *χotut* 'ladies'; Xaritonov 1947: 102), it is clear that the Mongolic plural suffix was initially copied by itself and only later was reinforced by the Turkic plural suffix (Xaritonov 1947: 102f; Kałużyński 1962: 116).

Through the introduction of the Mongolic substance copies, the Sakha language has undergone some phonological changes: [a] in first syllables in Mongolic copies remains [a], while it changed to [ï] in Turkic roots, and the voiced velar stop and fricative in word-final position, which are lost in Turkic roots, are retained as unvoiced velar stop and fricative in Mongolic copies. Initial [č] remains [č] in Mongolic and, later on, Russian copies, while it changed to [s] in Turkic roots; initial [s], which was lost in Turkic roots, is retained in Mongolic and Russian

copies, and initial [y] in Mongolic and Russian copies changes to [j], while in Turkic roots it changed to [s] (cf. section 3.1). In Mongolic copies, word-initial [j] is retained, while this is very rare and only secondary in Turkic roots, and Sakha words that were copied from Mongolic have word-initial [n], which is lacking in native words in Turkic languages in this position (Kałużyński 1962: 15, 27f, 39, 40, 47, 54; Johanson 1998b: 106). However, these changes of the Sakha phonological system were all introduced through the transfer of Mongolic substance copies. There is no evidence of phonological change taking place independently of the copies; furthermore, a number of consonant clusters occurring in copies transferred from Mongolic that may have posed difficulties for speakers of Sakha to pronounce, were not adopted, but assimilated (Kałużyński 1962: 58ff), e.g. *–ld–* > *–ll–* in *bolloχ* 'hill' (copied from Mongolic *boldaγ* 'hill'; Kałużyński 1962: 61). This is an indication that the transfer of substance copies took place through recipient language agentivity, that is, through individuals for whom Sakha was the linguistically dominant language, rather than that a number of speakers of Mongolic shifted to Sakha.

4.2 Substance copies and phonological influence from Evenki

Contrary to what we find for the Sakha-Mongolic contact situation, the Sakha-Evenki contact situation led to a relatively small amount of substance copies being transferred from Evenki to Sakha. Thus, Romanova et al. (1975: 158ff) list 35 lexical copies from Evenki in the literary Sakha language, and a further 97 in the northwestern and southern dialects of Sakha; however, as pointed out by Popov (1986: 58) this is not a complete list of lexical copies from Evenki to be found in the dialectal lexicon. Popov (1986: 58) counts approximately 250 words copied from Tungusic languages amongst the 6,200 lexical roots in the dictionary compiled by Pekarskij ([1907-1930] 1958-1959), i.e. approximately 4%; however, the majority of these are restricted to certain dialects. Amongst the approximately 8,500 items included in the *Dialektologičeskij slovar' jakutskogo jazyka* (Afanas'ev et al. 1976), nearly 300 words are copies from Evenki (i.e. approximately 3.5%), 31 have a model in both Evenki and Ėven, only 17 are copied solely from Ėven, and only four are copied from Yukaghir.

Amongst the 1,500 meanings from the 'Loanword Typology Project', only 13 are probably or clearly copied from Evenki, not even 1% (Pakendorf & Novgorodov in preparation). These are mainly words denoting natural phenomena, e.g. *χočo* 'valley' (copied from Evenki *kočo* 'bend in a river'; Romanova et al. 1975: 158), *jü:kte* 'spring' (copied from Evenki *yu:kte/ju:kte*; Cincius 1975: 350), and

čï:ča:χ 'little bird' (copied from *čïčaka:n* 'sparrow'; Cincius 1977: 401), as well as some cultural items specific to the northern climate and the hunting and reindeer-herding way of life typical of Evenks, e.g. *untu:* '(reindeer) fur boots' (copied from Evenki[2] *unta* 'footwear, fur boots'; Cincius 1977: 275), *ü:te:n* 'hut' (copied from *uten/u:te:n* 'hunter's cabin'; Cincius 1977: 295), and *mamïkta* 'lasso' (copied from *ma:βukta* 'lasso'; Cincius 1975: 520).

The copies from Evenki are predominantly nouns (Popov 1986: 60); some exceptions are *soŋo:* 'weep, cry bitterly', which was copied from Evenki *soŋo* 'cry' (Pekarskij [1926] 1959: 2281; Popov 1986: 66), and *argïy* 'knock heavily on something, drum' (copied from *argi* 'flow noisily (of rivers)'; Popov 1986: 66). The lexical domains from which the copies from Evenki come are mainly names for different animals, terms denoting natural phenomena, hunting implements, some items of clothing, and, not surprisingly, a large number of terms from the domain of reindeer herding and breeding; these, however, are mainly restricted to dialects whose speakers have switched to reindeer herding as a mode of subsistence (Antonov 1971: 64; Romanova et al. 1975: 158ff; Kałużyński [1982] 1995: 225ff; Popov 1986: 61ff). However, one reindeer herding term is found in the literary language as well: *u:čaχ* 'reindeer used for riding' (copied from *u:čak* 'reindeer for riding' found in the southern and southeasterly dialects of Evenki, or possibly from *ugu:čak* with the same meaning found in the dialects spoken in Yakutia; Myreeva 2004: 660f). The kinds of substance copies from Evenki in Sakha are what one would expect for a language whose speakers migrated into a very different environment, with different flora and fauna and a somewhat rougher climate than what they were used to.

Only two suffixes were copied from Evenki into the Sakha standard language: the emphatic *–kA:n* and the diminutive *–čA:n* (Kałużyński 1962: 80; Romanova et al. 1975: 157f), both of which are very productive, e.g. *en uol–ča:n–ïŋ* [2SG boy–DIM–POSS.2SG] 'your (little, dear) boy' [IvaP381]; *beye–tin kurduk–ka:n* [self–ACC.3SG like–EMPH] 'just like him' [Afny88]. Furthermore, the Sakha dialects spoken in the Bulun and Žigansk district have copied the Evenki augmentative suffix *–nja* (Romanova et al. 1975: 158).

Romanova et al (1975: 145ff) list a number of phonological changes that they claim the northern, and especially the northwestern, dialects have undergone due to shift of Evenki speakers to Sakha. These are most notably the occurrence of long

---

[2] However, Romanova et al. (1975: 159) claim that *untu:* was copied not directly from Evenki, but from Russian, which copied it from Evenki. This claim is not supported by Anikin (2003: 634), who only mentions that Dolgan copied *u:nči:k* 'short fur boots decorated with beads' from Russian *untik*, the diminutive form of *unty*.

vowels where the standard language has short vowels (e.g. *bar–a:m–mïn* instead of *bar–am–mïn* [go–PF.CVB–PRED.1SG]) (p. 147), the defricativization of the velar fricatives, e.g. the pronunciation of *χanna* 'where' as *kanna* (p. 150f), or *kïhalga* instead of standard Sakha *kïhalγa* 'necessity', since they claim that Evenki lacks a uvular fricative[3] (p. 152f), and the lack of consonant assimilation in clusters, e.g. *harsïn* instead of *hassïn* 'tomorrow' (which is the non-literary, but common pronunciation in the central districts), since in Evenki such consonant clusters do not undergo assimilation (p. 154). They furthermore point out that in copies from Evenki the front high unrounded vowel does not participate in vowel harmony, but occurs with back vowels as well, e.g. *aχtami:* '7-8 year old male reindeer' (Romanova et al. 1975: 147). Similarly, *banji:t* 'bandit', which was copied from Russian *bandit*, lacks vowel harmony. They also claim that the loss of the voiced uvular fricative *γ* in intervocalic position in the Olëkma dialect is due to Evenki influence, since this is frequently found in the 'shushing' subdialects of the southern dialect cluster (p. 153). However, judging from my data, this loss of intervocalic *γ* is widespread in all dialects of Sakha.

Judging from the narratives I recorded in the Olenëk district, the defricativization of the voiceless velar fricative does occur here, but very sporadically and infrequently, e.g. *tukarï* instead of *tuχarï* 'during, throughout' [Afny56, 57]; the very irregular occurrence of this phenomenon was pointed out by Romanova et al. (1975: 152) themselves. Similarly, there are very sporadic occurrences of long vowels instead of short ones, e.g. the person-marked perfective converbs of 'go' *bara:mmïn* instead of *barammïn* (1SG) or *bara:ŋïn* instead of *baraŋïn* (2SG). But these features are so sporadic that it is hard to classify them as substrate influence, which would lead to a more regular restructuring of the recipient language's phonological system. The only feature mentioned by Romanova et al. as constituting substrate influence from Evenki that I can find with some regularity in my data from the Olenëk district is the lack of consonant assimilation; thus, my older consultants in this district pronounced the words *tahïrja* 'outside', *χarčï* 'money', *harsïn* 'tomorrow', *barsa:r* 'go.REC.DSTIMP.2SG', and *ilje* 'carrying' as they are written here, while in the central, Suntar and Verxojansk districts they are pronounced *tahïjja*, *χaččï*, *sassïn/hassïn*, *bassa:r*, and *ijje* (although this last is pronounced with a palatalized *l* in the Verxojansk district, *ill'e*). Thus, there may have been some language shift by Evenki speakers to Sakha in northwestern Yakutia, which would be in agreement with the shift from cattle- and horse-breeding to reindeer-herding in this area. There is little phonological evidence for Evenki substrate influence in the Sakha language as a whole, however.

---

[3] However, the voiced uvular fricative *γ* occurs as an allophone of the stop *g* in intervocalic position (Ned: 321).

4.3 Differences in the origin of gender-specific lexical domains?

In a previous study (Pakendorf et al. 2003) I had concluded that the Sakha had undergone substantial Evenk admixture in the maternal line. This led me to suggest (Pakendorf 2001: 141) that sex-biased gene-flow might lead to a difference in origin of words from gender-specific lexical domains. That is, if the Sakha ancestors who immigrated from southern Siberia had in their new area of settlement married predominantly Evenki women, but not Evenki men, the female-specific lexical domains (those pertaining to child-bearing and –raising, cooking, and various household activities) might contain a number of Evenki copies, while lexical domains dealing with specifically male activities such as hunting or warfare, might be predominantly of inherited Turkic origin.

A look at different words from the 'Loanword Typology' chapters on 'mankind: sex, age, family relationship', 'body parts', 'cooking and utensils', 'arts and crafts', and 'warfare and hunting' (Pakendorf & Novgorodov in preparation) as well as at a few lexical items elicited separately in the field shows no difference between the female-specific and the male-specific lexical items. In both the female- and the male-specific domains, we find a number of Mongolic copies, as well as a large number of Russian copies in the domain of 'cooking and utensils' and 'warfare and hunting', while the majority of words appear to be inherited from Turkic. Thus, if such sex-biased gene flow took place, it appears not to have led to the introduction of their native words by the inmarrying Evenki women.

For most of the items a semantic shift is also not apparent, as far as I can judge at the moment. However, there is one exception: Sakha has several synonyms for 'pregnant', one of which is *ïaraχan*, which literally means 'heavy'. In the Turkic languages for which I was able to check this (Karakalpak, Kazakh, Turkish, Turkmen, Chuvash, and Tuvan), there is no extension of meaning from 'heavy' to 'pregnant'. In Evenki, however, *urge* means both 'heavy' and 'pregnant' (Boldyrev 2000: 22, 451); this semantic extension of the word for 'heavy' is also found in the Indigirka and Yukaghir dialects of Ėven (Cincius & Rišes 1952: 28, 628). Unfortunately, the *Russko-Nanajskij Slovar'* (Onenko 1986: 17) does not include an entry for 'pregnant', so that I am unable to judge whether this semantic extension is found in other Tungusic languages as well. If the parallel meaning extension was not due to independent internal innovations in Evenki and Sakha, the shared extension of 'heavy' to denote 'pregnant' might be an indication of contact between the two languages. However, since I am not sure how widespread this is in the Tungusic languages, it is hard to judge the direction of this potential contact influence.

5 DISCUSSION AND CONCLUSIONS

5.1 The linguistic results reviewed

In the very extensive third chapter I have discussed the origin and development of several features from phonology, morphosyntax and syntax in which Sakha differs from the other Turkic languages, most of which had previously been suggested as being the result of contact influence. These features are the shift of [s] to [h], several changes in the Sakha case system, the development of the Distant Future Imperative, the domain of possessive marking in Sakha, and the optional subject agreement on converbs. It should be pointed out at the outset of the discussion of these results that the number of features investigated here is limited, chosen as they were on account of their obvious difference from other Turkic languages. It can therefore not be excluded that further investigations of other traits of Sakha (for instance in the realm of the TAM system) may discover further evidence for contact influence that may shift the conclusions somewhat.

Of the features analyzed here, three have been shown to be probably due to internal innovations in Sakha that happened independently of any contact influence: these are (entirely uncontroversially) the development of the Comparative case, as well as (contrary to previous suggestions) the shift of [s] to [h] and the optional use of subject agreement markers on canonical converbs. The extension of the Dative case to include locative functions could be shown to be due to Mongolic influence (section 3.2.2.6). This result is quite straightforward, since both the Mongolic languages and Sakha have only one case to fulfill the three functions of marking recipients, the goal of motion, and stative location. Most Siberian and Turkic languages, in contrast, have a separate case to mark stative location, although very often allative and dative functions are conflated (cf. Tables 3a and 3b). Furthermore, initial Mongolic influence might have played a role in the development of the partitive meaning of the Locative case (cf. section 3.2.3.5) and in the retention of the distinction between a Comitative and Instrumental case (cf. section 3.2.4.6), though both of these features may equally well be the result of language-internal development.

Evenki contact influence could be shown for five features: the loss of the Genitive case in Sakha (section 3.2.1.5), the development of the indefinite accusative meaning of the Partitive case (section 3.2.3.9), the retention of the distinction between the Comitative and Instrumental case (section 3.2.4.6), the development of the Distant Future Imperative (and possibly even the development of the Purposive Converb; section 3.3.7), and the pragmatic uses of the possessive suffixes as well as the use of a derivational suffix to mark terms for kin and friends

that are not in a relationship with the speaker (section 3.4.2.2, 3.4.4, and 3.4.5). Furthermore, it is possible that Evenki contact influence played a role in the development of the switch-reference function of the Immediate Precedence Converb −*A:t* (section 3.5.6).

Of the five schematic copies from Evenki, two have good diagnostic value, because they are quite rare in Siberia and even world-wide. These are the development of the indefinite accusative meaning of the Partitive case in Sakha, which is shared only between Evenki and Sakha (and Dolgan), and the Distant Future Imperative, which is a relatively rare feature on a world-wide scale. The distinction between Instrumental and Comitative is cross-linguistically very frequent, so that this feature alone does not carry much diagnostic weight; however, in conjunction with the other evidence for contact influence from Evenki, and given the formal similarity between the Evenki suffixes and the Sakha suffixes in the possessive declension, it can justifiably be assigned to Evenki contact influence as well.

With regards to the development of the indefinite accusative meaning of the Partitive case, assigning the direction of the contact influence is not straightforward, since Evenki did not inherit this case marker or this function from the Tungusic languages. However, following Heath's method of internal reconstruction (Heath 1978: 23, 74f; cf. section 1.4.2) it appears more likely that the Indefinite Accusative case in Evenki is older than the indefinite accusative function of the Partitive in Sakha, since the Evenki case has more functions and occurs in more contexts than the Sakha case (cf. section 3.2.3.9). Similarly, the arguments for Evenki contact influence on the retention of the Comitative-Instrumental distinction in Sakha are somewhat tenuous, since such a distinction is very common cross-linguistically, both world-wide and in northern Eurasia, making an internal development quite plausible. Furthermore, although contact appears likely from the similarity in form between the two Evenki Comitative case suffixes −*nun* and −*nAn* and the Sakha variants of the Comitative suffix in the possessive declension −*nI:n* and −*nA:n*, Evenki once again stands alone among the Tungusic languages (even the Northern Tungusic languages) in having the suffix −*nAn*. However, the Evenki suffix −*nAn* has a very specialized function of marking joint actions with relatives or close friends (i.e. members of the class of inalienably possessed persons), while the Sakha variants appear to be synonyms. The Evenki functional specialization is in agreement with the overall Tungusic system of formally distinguishing inalienably from alienably possessed items, while the Comitative case is the only case of Sakha to have two variants in the possessive declension. This makes Evenki influence on Sakha more plausible than the other way round.

The changes undergone by Sakha under contact with Mongolic and Evenki were both system-altering and system-preserving (*qua* Aikhenvald 2003a: 2). Thus, the extension of the Dative case under Mongolic influence led to the loss of the separate Locative case, while the development of a Partitive out of the Locative case led to the development of a new case with new functions. The extension of the Partitive case to a meaning of indefinite accusative, however, did not alter the system, although it extended the category of direct object marking by a third possibility. The retention of the Instrumental-Comitative distinction was similarly system-preserving, while the loss of the Genitive case, the development of the Distant Future Imperative, and the development of the pragmatic uses of the second singular possessive suffix all altered the system.

From the above it is clear that we are dealing with two different contact situations in Sakha prehistory: contact with Mongolic-speaking groups and contact with speakers of Evenki. These two contact situations were obviously of an entirely different nature, since Sakha has adopted a large number of substance copies (lexical items including basic words such as kinship and body part terms, as well as suffixes) from Mongolic (cf. section 4.1), while the adoption of schematic copies appears to have been relatively weak. The contact with Evenki, on the other hand, led to the transfer of schematic copies without concomitant transfer of a significant number of substance copies (cf. section 4.2). However, there is the possibility of yet another layer of contact influence in Sakha, as will be discussed in the following section.

5.2 Samoyedic substrate in Sakha?

In section 3.4, I discussed the possibility that the extensive possessive marking of kinship and body part terms, and the 'non-possessive' uses of the possessive suffixes, especially 2SG, in Sakha may be due to Samoyedic substrate influence (cf. especially sections 3.4.1 and 3.4.4). Such Samoyedic contact influence has been postulated by Stachowski (1998) for the development of the 'non-possessive' use of the 2SG possessive suffix in Dolgan. The similarity in function of the possessive suffixes between Nganasan and Sakha are quite striking: in both languages, kinship terms and body parts take extensive possessive marking; and in both languages the 2SG possessive suffix marks anaphoric reference and has a pragmatic function. This similarity of use of the possessive suffixes is much greater between Sakha and Nganasan than between Sakha and Evenki, since Evenki, like the other Tungusic languages, has a formal means of distinguishing alienably from inalienably possessed body parts. Furthermore, an unspecified Samoyedic substrate ("uralische Spuren") in Sakha has been suggested by Skribnik (2004: 159).

In addition, as mentioned in section 1.1.1.2 and 3.4.5, the Sakha stand out amongst the populations of Eurasia by their very high frequency of Y-chromosomal haplogroup N-Tat C (Pakendorf et al. 2006). This haplogroup is found in high frequency in Finno-Ugric and in some Samoyedic populations (Lahermo et al. 1999, Karafet et al. 2002), but is in low frequency in Turkic-speaking groups (Karafet et al. 2002; Zerjal et al. 2002), suggesting possible prehistoric male gene flow between Samoyedic groups and Sakha. However, Selkups and Nganasans have no or only very low frequencies of Tat C (Karafet et al. 2002); thus, male-mediated gene flow in Sakha from Nganasans in particular is rather unlikely. Yet the possibility exists that there was female-mediated gene flow, or that the Samoyedic substrate consisted of a group speaking a language related to Nganasan, but with a different haplogroup composition (the Forest Nenets, for example, have nearly 52% of haplogroup N-Tat C; Karafet et al. 2002). Although approximately half of the mtDNA HVR1 sequence types found in Nganasans, Nenets, and Enets (Derbeneva et al. 2002; Goltsova et al. 2005) are shared with Sakha, they are also shared with other groups, especially with Tuvans. This makes the hypothesis of a Nganasan or Samoyedic genetic substrate being absorbed by the Sakha after their migration to the north somewhat implausible; however, to be able to come to a more confident conclusion proper phylogenetic analyses of both mtDNA sequences and Y-chromosomal STRs would be necessary. However, if a group of Samoyedic-speakers had shifted to Sakha, leading to Samoyedic influence in the extended use of the Sakha possessive suffixes, one would expect to find linguistic evidence for contact influence in more features of the recipient language. I will therefore review the features analyzed in chapter 3 with the specific view of detecting potential Samoyedic substrate influence.

Nganasan has both an /s/ and a /h/ in its phonemic inventory; the /h/ in word-initial position developed out of Proto-Samoyedic initial *p– (cf. section 3.1.4). In Forest Nenets, Proto-Samoyedic word-initial /s/ developed into /h/, i.e. this sound change is identical to that taking place in Sakha. This fact is quite intriguing in the light of the high frequency of Tat C in Forest Nenets. However, as discussed in section 3.1.7, the sound change in Sakha is so recent that it is highly unlikely to have been the result of substrate influence; this holds even more for a hypothetical substrate for which no historical evidence of contact exists, such as would be the case for Samoyedic. Thus, the coincidental development of /s/ to /h/ in Forest Nenets and in Sakha cannot be interpreted as evidence for contact influence, and the conclusion that this was an independent development in Sakha still stands.

Next I analysed different features of the Sakha case system, starting with the loss of the Genitive case. As discussed in section 3.2.1.4, a genitive case is reconstructed for the Samoyedic languages, and although Nganasan has lost the case

ending, this was still documented by Castrén in the mid-19th century. Thus, for this feature Samoyedic contact influence in Sakha cannot be postulated. As to the extension of the Dative case in Sakha to include locative functions, this, too, cannot be traced to Samoyedic influence, since the Samoyedic languages have a separate case to mark stative location, while recipients and goals of motion are marked by one case (cf. section 3.2.2.6). Regarding the indefinite accusative meaning of the Sakha Partitive case, this is shared only between Sakha (and Dolgan) and Evenki amongst the Siberian languages, making contact between these languages very likely. Like most of the languages of Eurasia, in the Samoyedic languages direct objects are marked by the Nominative and Accusative case; the difference in case-marking does not, however, necessarily mark a distinction in referentiality of the direct object (cf. section 3.2.3.5). In Nganasan, some singular nouns appear with a special oblique stem when they are indefinite direct objects. However, since this concerns only a small part of the nouns, and since this is also the form that these nouns take when they are possessors in possessive noun phrases, it cannot be compared to the Sakha Partitive or the Evenki Indefinite Accusative case. Likewise, the fact that in Samoyedic languages direct objects following upon imperatives take a different case than in the Indicative mood (Nominative instead of Accusative) is too different from the Sakha object-marking following upon imperatives to be explainable as language contact. As to the distinction between the Comitative and Instrumental case which was retained in Sakha, Samoyedic languages make such a distinction as well. However, since this is a very widespread trait of Siberian languages, it cannot be taken as strong proof for contact influence of a Samoyedic language on Sakha (cf. section 3.2.4.5), while the assumption of Evenki contact influence is strengthened by a similarity in form between the Evenki and the Sakha case endings (cf. section 3.2.4.6). Finally, regarding the origins of the Comparative case in Sakha, like most of the Siberian languages, Nganasan marks the standard of comparison with the Ablative case, thus differing from Sakha. Although Vakh Khanty has a separate Comparative case to mark the standard of comparison, this alone provides too scanty evidence for an Ob-Ugric (or even deeper in time, Uralic) substrate, so that the development of a separate Comparative case in Sakha, Vakh Khanty, and in Nivkh must be attributed to independent innovations (cf. section 3.2.5.3).

Nganasan does have a Distant Future Imperative, in addition to an Immediate Future Imperative (cf. section 3.3.5). This might be interpreted as indicative of contact influence from Nganasan in Sakha; however, Nganasan stands alone among the Samoyedic languages in having this feature. Evenki, on the other hand, shares the form of its Distant Future Imperative with the other Northern Tungusic

languages, and the category (though not the formal means of expressing it), with one branch of the Amur Tungusic languages (cf. 3.3.4). Thus, it is more plausible that Evenki, and not Nganasan, influence led to the development of the Distant Future Imperative in Sakha. Furthermore, it appears that Evenki influence may have played a role in the development of the Nganasan Distant Future Imperative as well, since Evenki and Ėven are at the core of a Siberian area encompassing unrelated languages that share this feature (cf. Figure 3.2).

The final feature I examined in Sakha was the frequent use of subject-agreement marking on canonical converbs. In Sakha, this is done with predicative person markers (cf. section 3.5.1) and has pragmatic rather than syntactic functions (cf. section 3.5.2). In Nganasan, there are only two canonical converbs, one of which takes the Genitive case suffix of the possessive declension and occurs in SS and DS constructions; this is obviously very different from what we find in Sakha. Therefore, the conclusion still holds that the person-marking on canonical converbs in Sakha is an independent innovation (cf. section 3.5.7).

It thus becomes clear that of the features examined here, Samoyedic influence can be postulated with reasonable conviction only for the extended use of possessive suffixes in Sakha, both to mark inalienably possessed objects and to highlight pragmatically salient or important discourse participants. However, although the similarity of Sakha and Nganasan in this respect is striking, the pragmatic use of the Sakha possessive suffixes does resemble similar uses of the Tungusic possessive suffixes with additional marking by the Alienable Possession suffix (cf. section 3.2.5.2). Furthermore, in Sakha kinship terms and nouns denoting friends must take a derivational suffix when they are to be used without possessive marking. This suffix appears to be very similar in function and origin to derivational suffixes in the Tungusic languages (cf. section 3.4.5). Thus, there is some indication that the extension of use of the possessive marking in Sakha may have been due to Evenki, rather than Samoyedic, contact influence. There are further arguments against the influence of a Samoyedic substrate in this one feature: other linguistic traits that have undergone changes in Sakha can be shown to be due to Evenki contact influence, while this is the only feature for which Samoyedic contact influence can be postulated. Furthermore, there is historical evidence for contact between speakers of Sakha and speakers of Evenki, while this is lacking for the putative Samoyedic-Sakha contact scenario. As Thomason (2001: 93) points out: "The source language must be shown to be, or to have been, in contact with the receiving language, and the contact has to be intimate enough to make structural interference possible." Such close contact between speakers of Sakha and speakers of Samoyedic languages cannot be demonstrated, making the inference of Samoyedic contact influence in Sakha highly implausible.

5.3 Different kinds of contact situation in Sakha prehistory

There is thus evidence of two different kinds of contact that the ancestors of the Sakha were engaged in: contact with speakers of Mongolic, and contact with speakers of Evenki.

5.3.1 Contact with speakers of Mongolic

As was discussed in chapter 4, the kinds of substance copies that were adopted into Sakha from Mongolic provide an indication that the contact situation between the ancestors of the Sakha and the neighbouring Mongolic tribe(s) was socioculturally and politically unbalanced, with the Mongolic-speaking group dominating the Sakha ancestors (cf. Kałużyński 1962: 120). This has led to a large number of substance copies entering the Sakha language, but not very many schematic copies being made: as mentioned in section 5.1, I was able to find good evidence only for the extension of the Sakha Dative case to include locative functions. There might possibly have been some influence from Mongolic on the change of the old Locative case to a partitive function (although this could also be due to internal innovation), and initially on the retention of the Instrumental-Comitative distinction; although this, too, could well be due to language-internal development.

On the 'borrowing scale' proposed by Thomason & Kaufman (1991: 74ff) the Sakha-Mongolic contact situation falls into stage 3 ('more intense contact') of contact situations in which the recipient language is maintained, i.e. according to their classification this represents an intermediate stage of language contact situation comparable to that between some dialects of Nahuatl (the recipient language) and Spanish (the model language), or the situation of immigrant languages in the USA and the model language English (Thomason & Kaufman 1991: 80, 81f). These are both instances in which the politically and socially dominant, more prestigious language has been exerting its influence over the subordinate recipient language.

The fact that the Sakha ancestors adopted such a large number of lexical copies from Mongolic suggests that their speech community must have been quite exocentric, i.e. open to accepting the linguistic norms of others (Andersen 1988: 72). However, although they adopted such a large number of substance copies, there is as yet little evidence for a concomitant amount of schematic copying from Mongolic into Sakha. According to Winford (2005: 376f, following Van Coetsem 1988), substance copies are transferred predominantly in recipient-language agentivity, although some schematic copies can be transferred in such a situation as well. This

would imply that the ancestors of the Sakha were for the most part linguistically dominant in their ingroup language Sakha, and not in the Mongolic language of their socially dominant neighbours; this is further supported by the phonological adaptation of consonant clusters in the copies, as discussed in section 4.1. We can thus conclude that the ancestors of the Sakha were exocentric, but linguistically dominant in their ingroup language. This implies that they formed a relatively closed community with few network ties to the Mongolic neighbours, i.e. only a few individuals in the Sakha community were bilingual in Mongolic. This finding might appear to be contradicted by the number of suffixes copied from Mongolic, since bound morphemes are often claimed to be more difficult to copy than free morphemes. However, the copied suffixes entered the Sakha language together with a number of copied lexical items. Due to the agglutinative nature of both Mongolic and Sakha, the separate meanings of these suffixes are easily discernible and the suffixes are easily segmentable, enabling their use with native roots. Thus, the typological similarity of Mongolic and Sakha facilitated the copying of suffixes even without wide-spread bilingualism in Mongolic.

According to Ross (2003: 193), lexical substance copies are an "…intrinsic part of contact-induced change only when that change is catastrophic. Otherwise, lexical borrowing is not a necessary condition or concomitant of contact-induced change." In Ross' view, the copying of lexical items is indicative of 'culture contact' rather than language contact. This kind of 'culture contact' may have been the driving force behind the adoption of terms dealing with cattle breeding and the use of milk products, since it is likely that the Sakha adopted both the techniques and the names for them from the Mongols. Furthermore, lexical items from the domain of law and order may well have been copied together with the concepts themselves, if the Mongols imposed their political system on the Sakha. However, given the relatively large amount of lexical substance copies from Mongolic in Sakha, including some quite basic terms and a number of verbs, as well as the number of copied suffixes that have become productive, it is unlikely that the contact situation involved only 'culture contact'.

As mentioned above, from the types of substance copies it is clear that the Sakha were socially, culturally, and politically dominated by their Mongolic-speaking neighbours. Given the social dominance of the latter, one might speculate that those speakers of Sakha who knew enough of the socially dominant outgroup language made "emblematic forays into the second language" (Brody 1987: 509) by copying Mongolic lexical items as a means of acquiring some of the prestige associated with the dominant group. This has been suggested as the reason for the large amount of copied particles found in the native languages of Mesoamerica: "A

sprinkling of these words through a narrative lends a Spanish feel to the discourse. Even where most of the words and constructions are Mayan, a few Spanish particles will give the flavor of the prestige language." (Brody 1987: 510; cf. Stolz & Stolz 1996: 110f).

The group of Sakha individuals with an active command of Mongolic can only have been small; had a larger proportion of the Sakha speech community been bilingual in Mongolic, one would expect to find more schematic copies from Mongolic in Sakha than I do here. Furthermore, I would hypothesize that the group of Sakha with at least a passive knowledge of Mongolic would have been influential within the speech community, since the lexical substance copies introduced by them spread widely and covered all aspects of life. However, it is somewhat unexpected that an exocentric speech community dominated by a group that was perceived as having higher prestige did not attain a degree of bilingualism that would have enabled the transfer of more schematic copies than the few found in this study. This may either be an indication that the ancestors of the Sakha moved out of the sphere of Mongolic dominance relatively quickly, so that the period of contact was relatively short (this, however, is contradicted by Kałużyński's estimate of a 'longer' period of contact between the 12th/13th and 15th/16th century AD; Kałużyński 1962: 119, 122), or possibly the prestige associated with the Mongolic language was attained simply by sprinkling one's speech with a liberal helping of Mongolic substance copies, obtainable through a passive knowledge of the language, without the need of being actively bilingual (cf. Stolz & Stolz 1996: 110f). On the other hand, since the number of features analyzed here is not exhaustive, perhaps further investigations might uncover more evidence of Mongolic influence on the structure of Sakha and so change the conclusion that only a small number of Sakha were bilinguals.

5.3.2 Contact with speakers of Evenki

Contrary to what we find for the Sakha-Mongolic contact situation, the contact between Sakha speakers and Evenki speakers led to a relatively small amount of substance copies (approximately one tenth of the number of Mongolic copies) being transferred from Evenki to Sakha (cf. section 4.2). On the other hand, there is evidence of a larger number of schematic copies from Evenki than from Mongolic in the language. While a much larger amount of substance copies is found

in the dialectal lexicon than in the standard language[1], the schematic copies are found in the language as a whole, and are not restricted to specific dialects. This implies that the schematic copies were transferred at an early period after the Sakha migration to the north, while a number of the substance copies were transferred at a later stage, after the Sakha expanded over the area they currently inhabit.

The lexical items that were copied from Evenki into Sakha are such as would be expected when a speech community migrates into a new environment: words denoting local fauna and flora and some cultural items that accompany a change in climate and life-style. This may be an indication that Evenki did not have the social prestige that Mongolic had, so that language mixing noticeable to linguistically naïve speakers (i.e. the transfer of lexical copies easily detected as 'foreign') was eschewed (cf. Thomason & Kaufman 1991: 117: "[…] if the language of a shifting population did **not** contribute lexicon to the target language, other than a few words for local natural and cultural items, then we can conclude that the shifting population did not enjoy much social or political prestige.", emphasis theirs).

The number and kind of schematic copies from Evenki in Sakha are indicative of some degree of model-language agentivity, i.e. for some speakers of Sakha, Evenki would have been the linguistically dominant language. As outlined in section 1.4.3, such model-language agentivity is only possible if a considerable number of the Sakha ancestors were dominant in Evenki. This would be expected if a group of Evenki speakers shifted to Sakha, for example; however, rapid language shift is expected to lead to phonological influence from the model language in the recipient language, because the shifting speakers do not have the time to master the new phonological system (Thomason & Kaufman 1991: 39, 119f). However, in Sakha, Evenki phonological influence appears to be marginal at most, and restricted to the northwestern dialect (at least judging from my data), where it has not led to a regular restructuring of the phonology (cf. section 4.2). The only exception may be the loss of intervocalic γ, which might be due to Evenki influence (cf. section 5.5). However, this is very tenuous and needs more investigation.

In the absence of phonological change, what we are dealing with is a case of incipient metatypy (Ross 1996: 182). According to Ross, in some instances speakers of a language may use an outgroup language more frequently, so that they are actually more proficient in the outgroup language than in their ingroup language, with the latter serving mainly as an emblem of identity. In such cases, the kind of

---

[1] Of course, given the very recent standardization of Sakha, the distinction between the 'dialectal lexicon' and the 'standard language' is not really meaningful. What it mostly reflects is a lower proportion of Evenki substance copies in the central dialects, since these formed the basis for the standard language.

changes occur that I have found in Sakha: a restructuring of the recipient language under the influence of the model language, but with no concomitant phonological influence, and not much lexical substance copies.

The scale of restructuring of Sakha does not appear to be comparable to that documented by Ross (1996) on Karkar Island; this may be an indication that the period of model-language agentivity, i.e the period during which the Sakha were bilingual in Evenki, was relatively short. Of course, as mentioned above, this study has focussed on a limited number of features to analyze, and it may be that further investigations will turn up more schematic copies from Evenki. Additionally, given the overall similar structure of Sakha and Evenki, metatypy could not lead to such deep changes as those documented by Ross on Karkar Island, where two languages with a very different structure were in contact.

As mentioned above, the number and kind of schematic copies are indicative of model-language agentivity, i.e. proficiency of Sakha speakers in Evenki, while the number and kind of substance copies are indicative of a low prestige of the model language. This leads me to the intriguing conclusion that a number of individuals in the Sakha community must have been active bilinguals in Evenki, even though Evenki was not accorded a prestigious status. Such bilingualism could be explained by intensive social interaction, for example through barter, or through a large number of mixed marriages in a relatively cohesive network structure, i.e. in a situation where families of mixed ethnic origins were able to remain in contact and so maintain both languages. Whether such mixed marriages truly took place is a matter of investigation of molecular anthropology, and I turn here to a brief overview over the results of the genetic studies.

5.4 The genetic results

In order to facilitate comprehension of the following sections, I have included a brief introduction to Molecular Anthropology in Appendix 1. The following is merely intended to provide the reader with the necessary information to be able to understand the subsequent discussion; for details of the genetic results, readers are referred to the original articles in which they were published (Pakendorf et al. 2006; Pakendorf et al. 2007).

5.4.1 mtDNA analyses

From the mtDNA analyses it becomes clear that the Sakha are a very homogenous population, notwithstanding their large area of settlement. The pairwise Fst values[2] between Sakha from the central districts, the Vilyuy river, and the Verxojansk district are all not significantly different from each other, indicating that until very recently, at least, there were no geographical barriers to intermarriage. This is in good accordance with the historical records that document a very recent expansion of the Sakha over the huge area they occupy nowadays, originating from a fairly restricted area of settlement in the Lena-Amga-Aldan area. Likewise, there is no difference between individuals coming from the central districts who speak dialects characterized by *okan'e* and those whose dialects are characterized by *akan'e* (cf. section 1.1.1.1). Thus, there is no genetic evidence that these dialectal differences arose through differential substrate influence, at least not from the maternal side.

Interestingly, the group of Sakha-speaking Evenks from the Olenёk district are indistinguishable from the Sakha groups, their reindeer-herding culture and self-identity as Evenks notwithstanding. This corresponds to the observations of Gurvič (1977: 23f), who documented an origin from central Yakutia for the major clans inhabiting this district. This is also in good agreement with the linguistic data that do not show any substantial differences between the speakers from the Olenёk district and the other districts; even as regards such features as word-initial [h] and the 'non-possessive' use of the second person singular possessive suffix the Olenёk district is not any different from the Suntar district.

The Sakha are genetically closest to Central Asian groups (Mongols, Buryats, Kazakh, and Kirghiz; average Fst = 0.029), followed by the South Siberian Turkic groups (average Fst = 0.039), while they are genetically relatively distant from Evenks and Ėvens (average Fst = 0.062). This is an indication of a predominantly South Siberian origin of the Sakha mtDNA lineages, which is in agreement with the archaeological and ethnographic data postulating a southern origin (cf. section 1.1.1.2). This southern origin is confirmed by the AMOVA analyses, and is also apparent in the MDS plot based on Fst values between 17 Eurasian populations (Figure 4 in Pakendorf et al. 2006): here, the Sakha group closer to Central Asian groups and the South Siberian Khakas and Altai than they do to their closest geographic neighbours, the Evenks, Ėvens, and Yukaghirs. These results appear to

---

[2] Technically, what was calculated here was not Fst but Φst, a similar measure which takes the molecular distance between sequences into account.

indicate a lack of admixture of the Sakha with the indigenous groups after they came into contact with them.

However, there are clear affinities of the Sakha and Sakha-speaking Evenks with the Western Ėvens from the Ėveno-Bytantaj district: the Western Ėvens share nearly all their mtDNA haplotypes with the Sakha (cf. Table 3 in Pakendorf et al. 2007), and in the MDS plot based on several Siberian populations they cluster closely with the Sakha and the Sakha-speaking Evenks (cf. Figure 2 in Pakendorf et al. 2007), and not with other Ėven or Evenk subgroups. This is indicative of Sakha admixture in the Western Ėvens.

The AMOVA analyses also demonstrate some affinity of the Sakha with the Evenks, Ėvens, and Yukaghirs as well as with the South Siberian Turkic groups. The reason for this becomes apparent from the network based on haplogroup C sequences (Figure 2a in Pakendorf et al. 2006, Figure 3 in Pakendorf et al. 2007): here, it is evident that the South Siberian Turkic groups, the Evenks, Ėvens, and Sakha share the three major sequence haplotypes. This may be indicative of a shared maternal genepool of these populations, most likely before the migration of the Northern Tungusic-speaking groups and the Sakha to the north, when all of these populations were still residing in South Siberia. However, because of this widespread haplotype sharing, it is impossible to exclude further admixture between Sakha and the Northern Tungusic groups in Yakutia: if such admixture had involved women carrying the shared haplogroup C haplotypes (which, given their relatively high frequency in the groups concerned, is not unlikely), this would not be detectable.

From the network of the haplogroup D sequences (Figure 3a in Pakendorf et al. 2006) it becomes clear that a part of the Sakha maternal ancestors underwent a founder event, which can very approximately be dated to 1,300 ± 800 years. This founder event is evident from the very high frequency of haplogroup D5a in the Sakha (nearly 16%), which is much higher than that found in other populations; and most of the Sakha individuals belonging to this haplogroup share a single sequence type. Interestingly, the direct ancestor of this sequence type is a Buryat haplotype; this, together with the very small average Fst between Sakha and Central Asian groups may be an indication that there was some Mongolic admixture in Sakha in the maternal line.

Thus, the mtDNA results show that the Sakha maternal lineages are clearly of south Siberian origin. There is no conclusive evidence for admixture with the indigenous groups after their migration to the Lena river; however, a small amount of admixture in the maternal line cannot be excluded.

316

5.4.2 Y-chromosomal analyses

The most striking feature of Sakha Y-chromosomal diversity is the lack of it: on average 94% of the men carry the Tat C mutation (cf. section 1.1.1.2 and Table 6 in Pakendorf et al. 2006). Furthermore, similar to the mtDNA results, the Y-chromosomal analyses demonstrate that the Sakha are a very homogenous population; the frequency of Tat C is not significantly different between individuals from the central districts, the Vilyuy river, and the Verxojansk district. However, in contrast to the mtDNA results, the Sakha-speaking Evenks from the Olenëk district have a significantly lower frequency of Tat C, although this is still by far the most frequent haplogroup in these individuals (73%).

The Evenks from the Podkamennaja (Stony) Tunguska lack Tat C completely; they are characterized by very high frequencies of haplogroup C-M86 and N-P43 (cf. Table 5 in Pakendorf et al. 2007). The Evenks from southern Yakutia (Iengra) and the geographically close Nyukzha river also have high frequencies of haplogroup C-M86; but these groups do show approximately 20% of N-Tat C (cf. Table 5 in Pakendorf et al. 2007). However, the sample size of the Iengra Evenks is too small to allow any firm conclusions about their genetic affinities. As to the Ėvens, the Central and Eastern Ėvens have high frequencies of C-M86 as well; the Central Ėvens furthermore have high frequencies of N-P43. The Western Ėvens from the Ėveno-Bytantaj district (which is adjacent to the Verxojansk district), however, stand out in a striking fashion amongst the Tungusic-speaking groups by having an extremely high frequency of N-Tat C (91%, close to the value in Sakha).

The very high frequency of Tat C clearly distinguishes the Sakha, the Sakha-speaking Evenks, and the Western Ėvens from other Siberian and Eurasian populations (cf. Figure 6 in Pakendorf et al. 2007, and Appendix 2). A close look at the genetic variation within individuals carrying the Tat C mutation, however, reveals a number of interesting points: First of all, the Sakha and Sakha-speaking Evenks share nearly all of their STR-haplotypes, indicating a large amount of Sakha male admixture in the Sakha-speaking Evenks. Conversely, in accordance with the mtDNA results and the ethno-historical data, one could say that the Sakha-speaking Evenks are clearly a group of Sakha who experienced some male gene flow from Evenks and switched to a reindeer-herding and hunting way of life, without, however, giving up their language. Secondly, the Sakha have clearly undergone a substantial founder event – the majority of Sakha individuals carrying Tat C have one of only three closely related haplotypes (cf. Figure 7 in Pakendorf et al. 2006). This founder event can be dated to approximately 900 ± 440 years before present, i.e. it is in reasonably good agreement with the archaeological data that point to a

migration north in the 13th or 14th century. Thirdly, although the Western Ėvens have clearly undergone some Sakha male admixture (as can be seen from the fact that they, too, share the dominant Sakha haplotypes), half of the Western Ėven Y-chromosomes carrying Tat C are not related to Sakha Y-chromosomes, but consitute a single haplotype falling into the part of the network which contains individuals from other Eurasian populations (Tuvans, Yukaghirs, Buryats, Finno-Ugric-speaking groups; only the Ėven, Tuvan, and Yukaghir haplotypes are shown in Figure 7 in Pakendorf et al. 2007).

Thus, although there is evidence for admixture from Sakha into Western Ėvens, there is no evidence for admixture from the indigenous groups into Sakha in the paternal line. This means that overall the genetic analyses show no conclusive evidence of admixture from indigenous populations. Although some admixture from Evenks and Ėvens in the maternal lineage cannot be excluded, the lack of admixture in the paternal line makes it clear that language and identity shift of whole groups of indigenous populations cannot have played a major role in the prehistory of the Sakha people.

5.5 Bringing the linguistic and the genetic evidence together

Taken together, the linguistic and genetic results are very interesting. On the one hand, there is evidence of intense linguistic contact with Mongolic-speaking groups which can be dated both genetically and linguistically to the period of the Mongol Empire: The expansion of the Sakha Y-STR haplotypes on the background of haplogroup N-Tat C falls around 1100 AD using the pedigree-based mutation rate, with 95% confidence intervals from 700 AD to 1500 AD. Similarly, the expansion of mtDNA haplogroup D5a falls around 700 AD, with 95% confidence intervals from 100 BC to 1500 AD; thus, if both of these expansions were brought about by the same event, there is evidence of a founder event (e.g. the migration north) taking place in the period between 700 and 1500 AD. The nature of the substance copies from Mongolic indicates that the Mongolic-speaking neighbours of the Sakha were socially and politically dominant. This makes it very likely that the ancestors of the Sakha were in contact with Mongolic-speaking clans or tribes belonging to the Mongol Empire, since this is the period when the Mongols dominated vast areas of Eurasia. This period falls into the range of age estimates for the founder event based on the genetic data: The Mongol Empire lasted less than 200 years, from 1206 (although Chinggis Khan's military successes began in 1197; Kämpfe 1986: 185) until the second half of the 14th century, with the individual khanates in different regions being defeated between 1357 and 1380 (Trauzettel

1986: 254f; Weiers 1986a: 297; Weiers 1986c: 343; Weiers 1986b: 366). Thus, it seems that the Sakha copied a large number of lexical items from Mongolic (2,000-2,500 items, between 10% and 30 or 40% of the Sakha lexicon) in a relatively short time, contrary to Kałużyński's claims (1962: 119, 122; cf. section 5.3.1). During this time they remained linguistically dominant in Sakha, copying mainly lexical items from their socioculturally and politically dominant neighbours, perhaps for reasons of prestige. The relatively short period of Sakha-Mongolic contact may be the reason why not more Sakha individuals attained proficiency in Mongolic, as evinced by the small number of schematic copies (cf. section 5.3.1).

Interestingly, the social and political dominance of the Mongolic-speaking neighbours of the Sakha notwithstanding, there appears to have been hardly any male-mediated genetic admixture from Mongols into the Sakha ancestral population. Mongols and their close genetic relatives, Kalmyks, are characterized by very high frequencies of Y-chromosomal haplogroup C (51 and 61%, respectively) and by very low frequencies of Y-haplogroup N-Tat C (2 and 1%, respectively; Karafet et al. 2002; Nasidze et al. 2005). Only 2.5% of the Sakha men, however, carry Y-haplogroup C, while the majority of Sakha men belong to haplogroup N-Tat C (94%). Of course, since both the Mongols and the Sakha have patrilocal marriage patterns (Jankowiak 1994: 475; Balzer 1994: 406), one would not necessarily expect intermarriage of Mongol men with Sakha women to lead to the introduction of Mongolic Y-chromosomes into the Sakha population. Since the Sakha brides or concubines would have been expected to settle with their Mongolic husbands, the Mongolic Y-chromosomes would have remained in the Mongolic community[3]. However, in this context it is very interesting that the Tuvans, who are similarly patrilocal (Leighton & Bicheldei 1994: 374), did experience some male admixture from Mongols. This is shown by the presence of four Tuvan individuals (7%) with haplogroup C-M86 and a duplication of the STR locus DYS19 (cf. Pakendorf et al. 2007); this duplication on this background has been found in Kalmyks (Nasidze et al. 2005). Thus, the lack of substantial male admixture from Mongols in Sakha may be an indication that the interactions of the Sakha with their dominant Mongolic neighbours were not very frequent and not very intimate, further supporting the assumption based on the linguistic data that the Sakha speech community as a whole was a closed community (*qua* Andersen 1988).

If admixture with the dominant Mongolic groups had taken place, one would rather expect to see it in the maternal line, since Mongolic brides would have settled

---

[3] Rape is, of course, a different matter, and the lack of Mongolic male admixture in Sakha indicates that there was relatively little rape of Sakha women by Mongol men (or if it did take place, it did not lead to the birth of sons).

with their Sakha in-laws. The clustering of the Sakha together with other South Siberian and Central Asian groups in the MDS analysis based on mtDNA sequences (cf. Figure 4 in Pakendorf et al. 2006) may be an indication that there was an exchange of brides between these groups. For example, the most frequent mtDNA sequence type found in the Sakha (where it has a frequency of nearly 13%) may possibly have originated from Mongolic female admixture, since the most direct ancestor of this haplotype is a Buryat sequence (cf. Figure 3a in Pakendorf et al. 2006). However, it is plausible that this admixture concerned only a single Mongolic woman (or a small group of women related in the maternal line), since a single sequence type expanded in the Sakha. This is in agreement with the linguistic data: if a large number of Mongolic-speaking women had married into the Sakha community, one would expect to find more schematic copies from Mongolic, since these inmarrying women would have been linguistically dominant in the Mongolic language. Model-language agentivity, however, is expected to lead to the introduction of schematic copies into the recipient language. Thus, the relative paucity of schematic copies from Mongolic in Sakha indicates that this kind of intermarriage cannot have been very frequent, possibly due to the higher social standing of the Mongolic community: women more frequently marry up the social ladder than down (e.g. Wooding et al. 2004).

On the other hand, the linguistic data indicate that the contact with the Evenks was of a different kind, leading to linguistic dominance of Sakha individuals in Evenki; yet in the genetic data there is no conclusive evidence for a high rate of intermarriage between the two groups. This is a very intriguing result, since one would expect the kind of intense linguistic contact postulated here to occur in a situation of intermarriage, similar to that found by Aikhenvald in the Vaupés river area (Aikhenvald 1996 and elsewhere). There are several possible explanations for this rather surprising lack of correlation of the genetic and the linguistic data: 1) intermarriage with Evenk women did take place, but cannot be detected with the methods used here, 2) the Sakha were initially dependent on the indigenous groups for the local knowledge necessary for survival in the harsh northern climate and therefore used Evenki as a frequent outgroup language, and 3) during its period of use as a lingua franca in the region, Sakha underwent a process of koineization that led to the structural changes analysed here.

As regards intermarriage with Evenk women, female-mediated gene flow from Evenks cannot be excluded, as mentioned in section 5.4.1, since there is widespread sharing of haplogroup C mtDNA sequences between Evenks and Sakha. Since these sequences are shared with South Siberian Turkic groups as well, they indicate a common maternal gene pool of Turkic- and Tungusic-speaking groups in

the period when all were still settled in South Siberia, prior to the migration to the north, first of the Northern Tungusic groups and later of the Sakha. However, given this sharing of sequence types, later admixture with individuals carrying these types will not be detectable with mtDNA analyses alone. However, the Sakha are genetically closer to South Siberian and Central Asian groups than to the Evenks and Ėvens, indicating that the female-biased intermarriage with Evenks may not have been very extensive. Unfortunately, all conclusions regarding the degree of Evenk admixture in Sakha are complicated by the fact that Evenks appear to have undergone genetic drift, making an analysis of their genetic affinities difficult. Thus, although there is no conclusive evidence for intermarriage with Evenk women, this cannot be excluded; possibly the number of schematic copies from Evenki in the Sakha language, which are indicative of model-language agentivity, are proof of such intermarriage.

With respect to Sakha dependence on their neighbours, it is hard to imagine nowadays that the Sakha, who are one of the largest and most dominant aboriginal ethnic groups in Siberia, could once have been dependent on the indigenous hunting and reindeer-herding groups. However, the Y-chromosomal analyses provide good evidence for a severe founder effect in the paternal prehistory of the Sakha population, with only a very small group of related men migrating to the north. Initially, the Sakha immigrants may well have been in a very vulnerable position, until they had acquired the necessary local knowledge to cope with the new environment, since their cattle depend entirely on hay during the very long winters. This vulnerability would have been exacerbated by their small group size. Although the climate in Siberia was generally warmer in the 9th to 14th centuries and there is evidence for a relatively warm climate in the Suntar Khayata mountain range in eastern Yakutia at 660 ± 90 years BP, i.e. in the 13th to 15th centuries, there were at least two cold periods in the 11th and 13th century, and after the 15th century the climate cooled off noticeably (Solomina & Alverson 2004: 6, 10f). Thus, the ancestors of the Sakha may well have had to cope with periods of difficulty; furthermore, the climate in northeastern Siberia can be expected to have been harsher than that in southern Siberia even in relatively benign periods. Thus, it is possible that the small group of Sakha immigrants interacted closely with the neighbouring hunting groups and used Evenki as their outgroup language.

Of course, the two hypotheses outlined above are not mutually exclusive: the initial small size of the immigrating group of Sakha ancestors and their dependence on their Evenki-speaking neighbours may have facilitated the intermarriage with Evenks in the female line. This may have been further enhanced by the fact that the Sakha were exogamous and patrilocal, meaning that brides would have had to come

from a different clan than the groom and reside with the groom's family after marriage. This would have led to a specifically female influx into the Sakha population – as indicated by the higher genetic diversity in the mtDNA than in the Y-chromosome (cf. Table 3, 6 and 7 in Pakendorf et al. 2006).

Regarding the possibility of koineization, Sakha was in use as lingua franca in a wide region of northeastern Siberia during the 19th and 20th century (Wurm 1996b: 976; Maslova & Vaxtin 1996: 999), and this may have held for earlier periods as well. If the Sakha themselves were monolingual in Sakha, but if their speech community was open, with a lot of communicative ties to other groups who used Sakha as their outgroup language, the Sakha speakers may have adapted to the speech of their communication partners in a process of koineization, adopting the structural features imposed by these on their variant of Sakha. This scenario would explain why we find linguistic evidence of close social interaction with Evenks without concomitant physical interaction. However, since Sakha was used as a lingua franca by several different ethnic groups, not just Evenks, one would expect the process of koineization to have introduced features from different languages, not just Evenki. However, although features such as the loss of the Genitive case or the development of the Distant Future Imperative may have been introduced by Ėven-speakers as well, the development of the indefinite accusative function of the Partitive case and the low-vowel variant of the Comitative case suffix in the possessive declension, link Sakha specifically to Evenki, making the hypothesis of a lingua franca-mediated koineization process rather unlikely. Furthermore, according to Ross (2003: 193), such a koineization process would be expected to lead to a simplification of the language; however, compared to the other Turkic languages, the Sakha case system has been expanded rather than simplified, as has the Imperative mood. In addition, at least at the end of the 19th century there were far more Sakha than members of other ethnic groups in Siberia: the census of 1897 counted 53,200 Evenks (apparently including Ėvens), 948 Yukaghirs, and 227,400 Sakha. For the other ethnic groups data appear only to be available from the 1926 census onwards; they then numbered in their hundreds or thousands (Evstigneev 2003: 75, 77, 80, 135, 136, 140, 143). This shows that the Sakha numbered approximately four times more than the other ethnic groups taken together; this numerical dominance makes it rather unlikely that their language would have been influenced to a detectable degree by second-language speakers belonging to much smaller ethnic groups. Thus, the explanation of koineization leading to the changes in Sakha appears rather unlikely.

A very interesting issue is that of language shift of Evenks to Sakha. It has been suggested that whole groups of Evenks shifted to Sakha and became

assimilated by the latter both culturally and linguistically (Seroševskij [1896] 1993: 230f; Dolgix 1960: 369, 461, 486; Tugolukov 1985: 220). If this had truly been the case, one would expect to find signs of genetic admixture from both men and women in the Sakha; likewise, one would expect to find an indication of this shift in the phonological system of the Sakha language (Thomason & Kaufman 1991: 39, 119f; Ross 2003: 193). As shown by the results of this study, there is no sign of Evenki male admixture in the Sakha as a whole, since the Evenks (and other Northern Tungusic groups; Karafet et al. 2002) are characterized by high frequencies of Y-chromosomal haplogroup C-M86, while the Sakha have predominantly Y-haplogroup N-Tat C. This is an indication that there cannot have been large-scale shift in language and identity of entire Evenk communities; if language shift took place at all, these Sakha-speaking Evenks must have retained their Evenk self-identity and would thus have been excluded from my genetic analyses. Of course, it is possible that there were small groups of Evenki-speakers who shifted to Sakha who had a large and detectable effect on the language as a whole, while their effect on the Sakha genepool was either negligible or not picked up in this study[4]. This might have been the case if such groups of linguistically influential Evenks were wiped out by epidemics after they had brought about the documented changes in the structure of Sakha, thus not leaving genetic traces of their influence; or if their genetic influence had remained relatively localized. However, given the relatively good genetic coverage of the central Yakutian districts, which represent the core of the later expansion, it is rather unlikely that major genetic lineages were missed in this study. Furthermore, it is somewhat hard to imagine a situation in which a group of shifting speakers has so much prestige that their way of speaking influences the entire language they are shifting to, while this prestige does not enable them to make a substantial contribution to the genepool of the influenced population. There also does not appear to be much evidence for a restructuring of the Sakha phonological system due to Evenki influence in most of the dialectal groups analysed here, corroborating the genetic evidence that large-scale language shift did not take place. There is a very slight possibility that the loss of intervocalic γ observed in Sakha (e.g. in fast speech phenomena, but also in the standard variation of the Indicative Future suffixes in the singular) may be due to Evenki substrate influence, as suggested by Romanova et al. (1975: 153) for the Olëkma dialect of Sakha. In Evenki, γ is frequently dropped in intervocalic position (where it occurs as an allophone of /g/), e.g. *ugi/uvi/ui* 'the upper part' (Ned: 321), and such a fricativization and subsequent deletion of the uvular voiced stop in

---

[4] I thank Christfried Naumann for bringing this possibility to my attention, as well as pointing out to me that temperatures were warmer in the Middle Ages.

intervocalic position is also found in Udihe (N/T: 57). However, more investigation is necessary to elucidate the range of this sound change in Sakha, as well as the time of its development, in order to evaluate the claim of substrate influence. If it should be substantiated, it might provide evidence for Ross' (2003: 193) suggestion that "…where phonological change accompanies metatypy, it is the result of some separate factor such as the frequent introduction of 'foreign' spouses into the group…", – of course, only if the possibility of female Evenk admixture in the Sakha should also be substantiated.

However, in the Olenëk district, where the immigrating Sakha (Gurvič 1977: 13, 23f) switched from a cattle- and horse-breeding lifestyle to nomadic hunting and reindeer-herding, we find both some evidence of male-mediated gene flow from Evenks as well as some indication of phonological influence from Evenki, i.e. the lack of consonant assimilation characteristic of spoken Sakha (cf. section 4.2). Thus for this district I can postulate that the change in lifestyle was mediated by intermarriage with Evenk men who adopted the Sakha language of the incoming groups. It is rather interesting that here we find male Evenk admixture in the incoming group of Sakha, although the Evenks, like the Sakha, were patrilocal (Fondahl 1994: 122). This may be due to the social system having been severely disrupted by the radical depopulation brought about by smallpox epidemics (Dolgix 1960: 446; Gurvič 1977: 10). Although the inmarrying Evenks appear to have exerted some phonological influence on their newly-adopted language, the Sakha language itself was transmitted by the Sakha mothers.

The combination of genetic and linguistic studies has thus shown that although there may have been female admixture with Mongolic-speaking groups, the Sakha community as a whole was predominantly Sakha-speaking. This led to the adoption of substance copies by the recipient-language dominant Sakha, with not much transfer of schematic copies. On the other hand, although the evidence for genetic admixture with indigenous groups is very scant, the linguistic data show evidence of model-language agentivity, with individuals who were dominant in Evenki introducing schematic copies into Sakha. This can either be explained by some Evenk female admixture in the Sakha population, or by Evenki being used as a frequent outgroup language by the Sakha community. Such frequent intergroup communication may have been necessitated by the small size of the immigrant Sakha, who initially were quite vulnerable in the new environment and therefore dependent on contact with the indigenous neighbours.

5.6 Some theoretical considerations

Thomason & Kaufman (1991: 50, 74ff) suggest that in cases of language maintenance, progressively more substance copies are introduced from the model language, with schematic copies following in ever more intense contact situations. This prediction is borne out by the evidence of Mongolic contact influence in Sakha: here, a large number of substance copies have been made from Mongolic, with a relative paucity of schematic copies. On the other hand, Thomason & Kaufman (1991: 50, 121ff) suggest that cases in which relatively few substance copies were introduced at the same time as a number of schematic copies were made are indicative of language shift. This prediction is revised by Thomason (2003: 692), who claims that such cases are indicative of imperfect learning of a target language (the recipient language) by a group of people. This prediction, however, cannot be borne out by the current study: although there is evidence for more schematic copies from Evenki in Sakha than from Mongolic, and although the number of substance copies from Evenki is rather small, there is no conclusive evidence of language shift. Furthermore, the amount of phonological influence from Evenki in Sakha as a whole appears minor as compared to the amount of structural influence, a fact which contradicts both the assumption of language shift and that of imperfect learning. Thus, some other factor must have been at play here, which emphasizes the fact that language contact situations cannot simply be dichotomized into cases of 'borrowing' or 'language maintenance' vs. 'language shift' or 'imperfect learning' (cf. section 1.2.7).

Furthermore, there is an interesting inverse correlation between the number of substance copies and the amount of schematic copies found in Sakha. Thus, Sakha has copied the largest number of lexical items in the 'Loanword Typology' database from Russian, followed by a large number of copies from Mongolic, and a relatively small number of copies from Evenki (Pakendorf & Novgorodov, in preparation). Conversely, the amount of structural influence from Russian in the spoken language of Sakha living in rural settlements appears negligible (though this statement is based solely on superficial observation, and not on any actual analysis), while there has been some detectable influence from Mongolic. The strongest structural effect on Sakha can be traced to Evenki. This observation is supportive of Ross' suggestion that the introduction of substance copies is a cultural factor and not necessarily indicative of language contact (Ross 1996: 209f; Ross 2003: 193).

5.7 Conclusions and outlook

This study aimed at evaluating the extent to which the Sakha came into contact with the indigenous populations of the area in which they are currently settled, both from a physical and from a sociocultural perspective (cf. section 1.4.1). The results have shown that the sociocultural contact was quite intense, especially with the Evenks, since the schematic copies from Evenki found in Sakha are indicative of model-language agentivity. In contrast, the physical contact appears to have been less intense, since the evidence of admixture with Evenks is rather weak. This opens up the possibility that the Sakha may have been in social contact with their Evenk neighbours without concomitant intermarriage.

A further aim of this study was to evaluate previous claims of Evenki contact influence in Sakha. Through very detailed analyses of the linguistic features under investigation not only in Sakha and Evenki, but also in other Tungusic, Mongolic and Siberian languages, I was able to show that regardless of the superficial similarity between Sakha and Evenki in the change of [s] to [h] and in person-marking on canonical converbs, these features cannot be attributed to contact influence, but must have arisen through independent innovations in Sakha.

Last, but definitely not least, I aimed at evaluating the feasibility and advantages of combined genetic and linguistic studies of prehistoric population contact. This raises the obvious question of whether I could have come to the same conclusions that I have reached using only one of the fields investigated here, i.e. if I had focussed solely on the genetic analyses of Sakha prehistory, or solely on the linguistic evidence for contact influence. The answer to that question is a clear 'no': although the genetic results are not as conclusive as may have been desired, the combination of both kinds of analysis has permitted me a broader perspective of Sakha prehistory than would have been possible with only one line of investigation.

Had I performed only a molecular anthropological study, I would not have been able to show that the Sakha and the Evenks must have been in intense contact, since the evidence for admixture is rather weak in the maternal line, and completely absent in the paternal line. However, the fact that some of the changes that have taken place in the Sakha language as compared to its Turkic relatives are due to model-language agentivity, i.e. the dominant bilingualism of Sakha speakers in Evenki, provides a very strong indication that the contact cannot have been casual. On the other hand, had I studied only the linguistic outcome of the contact, I might have concluded that there was substrate influence from Evenki in Sakha, in accordance with the historical literature claiming the assimilation of groups of Evenks by Sakha. The genetic results, however, show clearly that large-scale language shift from entire Evenk communities to Sakha cannot have taken place,

and that the contact between Sakha and Evenks must therefore have been of a different nature. In addition, the genetic results provide clear evidence of a strong founder event in the Sakha paternal lineage – thus, it is clear that the group of Sakha ancestors who migrated to the north must have been very small. This raises the possibility that the Sakha were initially dependent on their Evenk neighbours for survival, providing a possible explanation for why the Sakha may have been in close social contact with Evenks without necessarily intermarrying with them.

It is clear that the combination of genetic and linguistic analyses is a very fruitful approach to population prehistory; however, in the current study it was hampered somewhat by the lack of genetic differentiation of the Sakha and the indigenous populations of Yakutia. This combined approach therefore promises to be even more fruitful in areas of the world where groups of very different origin have come into contact, such as Melanesia, where estimates of genetic admixture should be more straightforward. But even in an area like Siberia, there are further analyses that can be undertaken that may provide more insights into prehistoric population contact. This is first and foremost the inclusion of a large amount of autosomal markers into the analyses. After all, the mtDNA and the Y-chromosome represent only a tiny fragment of the entire human genome, and although their advantages in studies of population prehistory are undeniable, their small size is a limiting factor. Utilizing the large amount of information inherent in the autosomes will permit more fine-scaled analyses of genetic admixture of Sakha and Evenks.

Furthermore, although this study has provided some very interesting perspectives on Sakha prehistory, the linguistic analyses have been far from exhaustive. It is clear that an investigation of more features is necessary to obtain an accurate estimate of the degree of schematic copies from Mongolic as well as Evenki in Sakha.

In the following I give a very brief overview of some of the main issues involved in the use of genetic data to elucidate human prehistory. This will hopefully facilitate understanding the discussion as well as the genetic results of this study, which were published elsewhere (Pakendorf et al. 2006, 2007). For a more detailed description the reader is referred to the text book by Jobling, Hurles & Tyler-Smith (2004).

Molecular anthropology is the branch of biological anthropology that uses molecular genetic methods to study the origin, relationships, history, and migration patterns of human populations. Like other physical anthropologists, molecular anthropologists study the differences (polymorphisms) between humans. These polymorphisms exist at the molecular (protein and DNA, see below) level as much as at the physical level such as hair or skin colour, etc. Assuming an ultimately single origin of modern humans, similarities in the genetic make-up of individuals or populations may indicate a shared history and therefore relationship, whereas differences can indicate separate histories over a time frame specified by the genetic system studied.

1 The human genome

In animals and plants, the genetic information needed to build cells and run metabolic processes is stored in large molecules called deoxyribonucleic acid (DNA). DNA is present in two regions of the cell: the nucleus and the mitochondria. The DNA molecules present in the cell's nucleus are so large that at certain times during the cell cycle they can be seen under an ordinary light microscope as linear structures: the chromosomes. (That is, each chromosome is a DNA molecule.) Humans have 46 chromosomes in each cell (the exception being germ cells, i.e. ova and sperm, which contain only half the number of chromosomes), of which 44 form 22 pairs of homologous chromosomes. It used to be thought that these homologous chromosomes are identical in that they contain the same genes in the same linear order, and that they differ from each other only in the specific allele present at each locus. However, recent investigations have shown that even phenotypically normal, healthy humans differ by large-scale rearrangements of DNA such as inversions, insertions, and deletions (Iafrate et al. 2004; Tuzun et al. 2005; Conrad et al. 2006, *inter alia*). The 22 pairs of homologous chromosomes are termed autosomes, and are numbered 1 through 22. The remaining two chromosomes are the sex chromosomes, which in women form a homologous pair, whereas in men they are not homologous.

The sex chromosomes present in women are called X-chromosomes, while men have one X and one Y-chromosome. Every individual receives one set of 22 autosomes plus a sex chromosome from his or her father, and one set of 22 autosomes plus an X-chromosome from his or her mother.

In addition to the chromosomes present in the nucleus of the cell, DNA molecules are also found in the cell's energy-producing bodies, the mitochondria. This mitochondrial DNA (mtDNA) differs from the nuclear DNA in several ways. Firstly, it is much shorter (on average, nuclear chromosomes are approximately 7,800 times longer than mtDNA) and it is circular, instead of linear. Secondly, whereas the nuclear DNA is present as different pairs of homologous chromosomes, mtDNA is present in practically identical[1] copies. Furthermore, the number of molecules differs between nuclear DNA and mtDNA: there are only 23 different pairs of chromosomes (i.e. 46 molecules altogether) in each cell's nucleus, while there are hundreds of mtDNA molecules distributed within the several mitochondria in each cell. Lastly, mtDNA is passed on to the next generation only by women, i.e. it is inherited only in the maternal line (see below).

2 DNA in the germ line

It may be advisable to define some terms at this point: gene, locus and allele. A gene is a stretch of DNA that contains the information necessary to synthesize a protein (although occasionally the product encoded by a gene is not a protein, but a regulatory molecule called RNA). In general, genes consist of alternating sequences of coding DNA that can be translated into the sequence of amino acids making up the protein, and non-coding DNA, which is removed before DNA synthesis. A locus is a particular stretch of DNA and/or the product encoded by such a stretch of DNA. Since molecular anthropologists are interested in variation, the loci they study are polymorphic, i.e. at least two variants can be detected in human populations. Each of these variants is an allele at that specific locus.

In sexually reproducing species such as humans, germ cells (sperm and ova) differ from the other cells of the body in the number of chromosomes they carry. In normal cells, which carry two homologous sets of 23 chromosomes, the entire genome is present in duplicate – normal cells are therefore said to be diploid. In contrast, germ cells carry only one set of chromosomes, and are therefore said to be

---

[1] The mtDNA molecules within one individual may differ from each other in single nucleotides. This is termed heteroplasmy, and it arises through new mutations occurring on one molecule which gets replicated and can so spread within the cell. If an ovum contains heteroplasmic mtDNA, the arising offspring may be heteroplasmic.

haploid. One very important feature of germ cell production is recombination. Before the cell division, the two homologous chromosomes of each pair align. The two molecules physically cross each other, and sometimes during this process of crossing-over they break at the contact points. Frequently, in the fusion of such breakpoints two pieces from the maternal and paternal chromosome, respectively, are joined, resulting in an exchange of DNA stretches between two homologous chromosomes. Recombination shuffles the alleles inherited from the mother and those inherited from the father, resulting in a new mix of alleles on each chromosome. In addition, the resulting germ cell receives a random mix of chromosomes inherited from the mother and chromosomes inherited from the father. This process ensures enormous variation among an individual's offspring.

There are, however, exceptions to this biparental inheritance of DNA with recombination. First of all, mtDNA does not undergo recombination; that is, the genes carried on the mtDNA do not get shuffled (Ingman et al. 2000; Jorde and Bamshad 2000; Kumar et al. 2000; Behar et al. 2007). Furthermore, sperm carry very little mtDNA and the little that does enter the ovum at conception apparently gets eliminated within a few days (cf. references in Pakendorf & Stoneking 2005). This means that although every individual carries hundreds of mtDNA molecules in his or her cells, these are inherited solely from the mother[2].

On the other hand, the Y-chromosome in men's cells has no homologous counterpart. During cell division and germ-cell production the tips of the X and Y-chromosome align and recombine, but the major part of the Y-chromosome (termed NRPY – non-recombining portion of the Y-chromosome) does not undergo recombination (Jobling & Tyler-Smith 2003). In contrast to mtDNA which is inherited only through the mother, the Y-chromosome is passed on only from fathers to sons.

---

[2] A recent case study of a patient with exercise intolerance actually found that 90% of the mtDNA in his muscle tissue was inherited from his father, not his mother (Schwartz & Vissing 2002). However, since the other tissue samples tested (skin, blood, and hair) contained only maternal mtDNA, this seems to have been due to an accidental non-elimination of the paternal mitochondria during fertilization, which then ended up in certain cell lines of the embryo. A later study reported that the mtDNA molecules of maternal and paternal origin in this patient's muscle tissue had undergone recombination (Kraytsberg et al. 2004), demonstrating that this is possible. However, since in general the mtDNA molecules within a single individual are of uniparental origin, recombination between these will have no effects, and the non-recombining maternal inheritance of mtDNA can still be assumed for purposes of population studies.

The differences in inheritance between autosomes, sex chromosomes, and mtDNA lead to differences in the way polymorphisms on these molecules will behave in a population. In effect, when a man and a woman reproduce, there are four copies of every autosome that have a chance of being passed on to the next generation, as against three copies of the X-chromosome (two in the woman, one in the man), and only one copy each of the Y-chromosome and the mtDNA (which are passed on only from the man and the woman, respectively). This means that there are four times fewer Y-chromosomes and mtDNA molecules than autosomes participating in reproduction, and thereby in the propagation of particular polymorphisms. Furthermore, not every individual in a population participates in reproduction: some have passed their reproductive phase, some are as yet sexually immature, and not every sexually mature individual will find a mate or be able to reproduce. This means that the actual number of reproducing individuals, which is termed effective population size (N_e), is smaller than the census population size. One can therefore say that the effective population size of the Y-chromosome and the mtDNA is four times smaller than that of the autosomes, and three times smaller than that of the X-chromosome (Jobling, Hurles & Tyler-Smith 2004: 134). This has implications for population genetic analysis, since chance events and population movements have a stronger effect on smaller populations than on larger ones. Chance events will therefore effect the variation on the Y-chromosome and the mtDNA more than autosomal variation.

3 mtDNA

One of the most frequently studied stretches of DNA in molecular anthropology is the mitochondrial DNA (mtDNA). This molecule has a number of advantages for studies of population history (for a more detailed account see Pakendorf & Stoneking 2005). One of these is its high frequency in the cell, making studies possible even when only minute quantities of DNA are available. A further important advantage of mtDNA is that, although there are multiple copies of individual molecules present in every cell, these are all identical, i.e. one has to deal with only one sequence of nucleotides. (That is, mtDNA is in effect haploid as opposed to the diploid nuclear genome, where every autosomal gene is present in two copies; see above). Furthermore, as mentioned above, mtDNA does not undergo recombination, as does autosomal DNA, so that mutations remain in the context in which they arise (cf. section 5 below); this allows one to reconstruct genealogies of mutations. A further useful feature is that mtDNA is inherited solely in the maternal line, as described above. Therefore, studies of mtDNA variation illuminate

specifically the maternal history of a population, and, as recent studies have shown, this can differ from the paternal history (Passarino et al. 1998; Helgason et al. 2000a, b; Oota et al. 2001; Goodacre et al. 2005; Nasidze et al. 2006). Also, the smaller effective population size of the mtDNA as opposed to autosomal DNA (see above) means that processes affecting the genetic variation at this locus, such as population migrations or strong reductions in population size, will have much stronger effects and will therefore be more easily detectable.

There are two different approaches to study mtDNA variation. One is to establish the actual sequence of bases in a particular stretch of the molecule. Very often, molecular anthropologists focus on sequences of the so-called hypervariable region I (HVR I) of mtDNA. This is a 360 basepair (bp) long stretch of non-coding DNA that has a high rate of mutation. Sequencing this region means that every single mutation, and therefore even minute differences between individuals and/or populations, can be detected. Alternatively, molecular anthropologists study known polymorphisms dispersed over the whole mtDNA molecule. Although this approach does not detect every polymorphic site, it has the advantage of covering the entire mtDNA molecule, instead of focussing on one small stretch of DNA. With the advent of relatively cheap sequencing technology it is becoming more feasible to combine both approaches by sequencing the entire mtDNA molecule in several individuals and even whole populations (Ingman et al. 2000; Herrnstadt et al. 2002; Reidla et al. 2003; Fraumene et al. 2006; Olivieri et al. 2006), thus combining the advantages of both approaches.

4 The Y-chromosome

The counterpart of mtDNA is the Y-chromosome, which has been the focus of molecular anthropological studies in recent years (cf. Jobling & Tyler-Smith 2003 for a more detailed overview). The Y-chromosome is inherited only in the paternal line, so that it illuminates the history of a population's men. Since the Y-chromosome is present in only one copy in the cell nuclei of men (who have one X- and one Y-chromosome), it is haploid, like mtDNA. The major portion of the Y-chromosome also does not undergo recombination, and therefore mutations can be traced back in time. Studies of Y-chromosomal variation in humans focus mainly on two kinds of polymorphisms: STRs and SNPs. STRs (short tandem repeats; also termed microsatellites) are stretches of DNA consisting of repeated short sequences of DNA, for example dinucleotide repeats such as $(CA)_n$, or the tetranucleotide repeat $(CAGT)_n$. They vary in the number of repeats present in different individuals – for example, some individuals may contain 16 copies of a particular repeat

sequence, while others may have 17, 18, or more copies. STRs have a high mutation rate and can therefore be of importance for fine-grained studies of population history.

As the name suggests, SNPs (single nucleotide polymorphisms) are polymorphisms at which a single base in a certain stretch of DNA has been substituted for another. The mutation rate at these sites is so low that it is assumed that a specific substitution has occurred only once in human prehistory. This implies that all individuals (in the case of Y-chromosomal SNPs, all men) carrying a substitution at a specific site are genetically related. Therefore, SNPs are a powerful tool for establishing relationships of human populations and following paths of migration.

5 Haplogroups and haplotypes

Mutations arising on a DNA molecule in any individual in a population may spread through the population from generation to generation by reproduction. Since mtDNA and the Y-chromosome do not undergo recombination, as do autosomes, a mutation on these types of DNA will remain in its original context within the molecule. (In autosomes, such a mutation might well be passed to a different homologous chromosome, with a different DNA context, through recombination.) With time, further mutations may arise on a mtDNA molecule or Y-chromosome carrying the original mutation, leading to several mutations being associated with each other. This is called a haplotype. A particular haplotype is defined by the researchers conducting the study in order to describe their results. Sometimes, a haplotype is found in only one individual, whereas other haplotypes may be found in several individuals.

A haplogroup, on the other hand, is a group of related haplotypes that is defined by researchers on account of a recurrent association of specific mutations ('shared innovations'). A haplogroup is an indicator of genetic relationship, since it is assumed that the mutations defining it arose in a linear manner at different points in time on certain DNA molecules, and then spread through the population by reproduction of these molecules. Languages and language families are possible linguistic analogies to haplotypes and haplogroups, respectively: a language is defined by having a certain set of features, and a language family is a group of genealogically related languages that share certain, but not all, of the features of the individual languages.

The nomenclature of haplogroups in mtDNA has developed over time, with related groups of haplotypes being classed together as they were found in studies of

populations from around the world. Most mtDNA haplogroups are defined by SNPs in the coding region of mtDNA, but some can also be distinguished on the basis of certain HVR1 sequence types (haplotypes). Since the definition of these haplogroups proceeded in an *ad hoc* manner, the nomenclature lacks any coherent logic. Haplogroups A through G are found in Asia; of these, A to D also occur in the New World, where they make up the vast majority of mtDNA types found. Haplogroups H, I, J and K are found in Europe, with H and I present in western Asia as well. All the mtDNA variation found in Africa has been classified as (super)haplogroup L, with numbered subscripts differentiating between subgroups (L_0–L_6). Haplogroups T, V, W, and X are European haplogroups; interestingly, haplogroup X is also found in the indigenous populations of the New World. Recently, this haplogroup was also found in individuals from the Altay mountains in south Siberia, providing the first link between haplogroup X in Europe and in the Americas (Derenko et al. 2001). Finally, haplogroups Y and Z are found in restricted populations of northeastern Siberia and the Far East.

Researchers studying Y-chromosomal variation also distinguish haplogroups based on SNPs. Since the mutation rate of these SNPs is so low that they have arisen only once in human history, they are ideal markers for defining groups of related individuals: every man carrying a particular SNP is genetically related to every other man carrying this SNP. Of course, some SNPs are very widespread and therefore not of much use for studies of specific populations, whereas others are limited to certain geographical areas and can therefore provide information on the relationship of populations in these areas. One of these localized SNPs is the so-called "Tat" marker mentioned in sections 1.1.1.2 and 1.1.3.2, which is very widespread in populations of northern Eurasia and practically absent or very marginal elsewhere (Zerjal et al. 1997; Karafet et al. 1999; Lahermo et al. 1999; Rootsi et al. 2007). This may have arisen in northern China and spread throughout Siberia to Europe (Rootsi et al. 2007). After a period of nomenclatural confusion the Y-chromosomal haplogroup nomenclature has been unified (YCC 2002). Haplogroups of importance in Siberia are C, N and Q, while haplogroup O is widespread in Southeast and East Asia. Haplogroup J is found in Europeans, while haplogroup R is widespread in Europeans and Central Asians (Jobling & Tyler-Smith 2003).

6 Relationships between populations

The genetic distance between individuals can be measured directly by counting the number of mutations that have accumulated since they shared a common ancestor. However, if new mutations arising on diverging lineages of human populations were the only means of analyzing the prehistory of these populations, the time depth that could be investigated by molecular genetic methods would be far too early to be of interest to biological anthropologists. Fossil evidence points to an origin of anatomically modern humans in Africa between 200,000 to 100,000 years before present. In the fast-evolving HVR I of mtDNA a new mutation arises on average every 15,000 to 20,000 years. Therefore, only five to ten new mutations will have arisen on each lineage in the HVR I since the origin of modern humans – which implies that current-day populations would be too similar to permit any meaningful study of their past history. Questions of recent population history, such as the migration of the Sakha from South Siberia to the middle reaches of the Lena, could not be answered at all. The reason that molecular anthropologists continue to make claims about such recent events is that new mutations are not the only source of inferences about the past. Other random factors can shape the patterns of genetic variation much more strongly.

One important factor is the fate of the preexisting genetic variation in the ancestral population. If a population starts to split, each daughter population will receive some of the variation already present. Depending on the geographic and cultural circumstances involved, the daughter populations may each receive a slightly different subset of the initial variation. Furthermore, some of the mutations present initially in the daughter populations may disappear through random factors – individuals carrying these polymorphisms might not find a mate and therefore will not reproduce, resulting in their genetic lineage dying with them, or men might have only daughters, or women might have only sons, resulting in the termination of their Y-chromosomal and mtDNA lineages, respectively. These random processes affecting the genetic variation in a population are termed genetic drift. Genetic drift has a stronger effect in a small population than in a large one, so that demographic factors such as population size play a role in shaping genetic variation, too. For example, an allele with a frequency of 1% would be present in 1,000 individuals in a population of 100,000, whereas in a population of 1,000 this allele would be present in only ten individuals. It is more probable that ten people die young, or do not find a partner, or have children who die young, than this happening in 1,000 individuals (cf. Jobling, Hurles & Tyler-Smith 2004: 131ff).

An extreme case of genetic drift is the so-called founder event. In this event, a small group of individuals leaves the ancestral population and migrates to a

different location, for example an island. This founding group will carry only a small subset of the variation present in the entire ancestral population, and this subset will thus be present at high frequency in the resulting new population. A very good example of a founder event is the colonization of the New World, where four of the seven mtDNA haplogroups defined in Asia are present in very high frequency in the indigenous populations, whereas the other three are completely absent (Torroni et al. 1993, Stone & Stoneking 1998). Similarly, should a large number of the ancestral population rapidly die off, the survivors will also retain only a subset of the initial variation. This event is termed a bottleneck, and it, too, can lead to differences between populations that share a common ancestor. Identifying whether the prehistoric event was a bottleneck or a founder event is not easy, since the result seen in the current-day population is the same: reduced genetic variation that represents a subset of the variation present in the presumed ancestral population (Jobling, Hurles & Tyler-Smith 2004: 132f). To distinguish between a founder event and a bottleneck, data from other fields, especially archaeology, are needed.

7 Methods of inferring population relationships: Fst, MDS, and AMOVA

Population relationships can be inferred from molecular genetic data by estimating genetic distances between individuals or populations. One measure that is frequently used as an estimate of genetic distance is Fst (Jobling, Hurles & Tyler-Smith 2004: 168). Fst is the proportion of the total genetic variation[3] that is due to differences between populations. Fst is estimated by subtracting the genetic variation that exists within individual populations from the genetic variation found in the combined populations. If two populations do not differ genetically, the amount of variation present within each of them will make up all or most of the genetic variation present within the total sample, and therefore the Fst value will be close to zero (cf. the non-significant Fst value of 0.02 between Central Ėvens and Western Ėvens in Table 3 of Pakendorf et al. 2007). If, however, populations are genetically very different, the Fst value will be significantly larger than zero (even though most of the variation within the total sample will still be due to the variation existing within them; cf. the significant Fst value of 0.21 between the Iengra Evenks and the Stony Tunguska Evenks (STE) in Table 3 of Pakendorf et al. 2007). The significance of Fst values can be estimated by permutation, which is of great

---

[3] Technically, Fst is the proportion of the variance that is due to differences between populations. However, for the purposes of this introduction, genetic variance and genetic variation are equivalent.

advantage in data interpretation, since it is allows one to distinguish small values that are only due to chance from small values that are truly indicative of shared history. In order to test the significance of Fst values by permutation, all the haplotypes of the total sample are pooled, and then randomly allocated to "populations" equivalent in size to the original populations in the comparison. The Fst value is calculated for this random sample of "populations", and then the procedure is repeated; in general, to assess the significance of Fst values one performs 10,000 permutations. At the end of the procedure, the real, observed Fst values can be compared with the randomly generated ones to assess whether they, too, are due to chance or whether they are significantly larger than these random values. The measure of Fst was first developed by population geneticists for classical allele frequency data, but it has since been expanded to take account of the new data being generated nowadays. Measures related to Fst are Rst, which is adapted specifically to the singular mode of mutation of STRs, and Φst, which is specifically adapted to molecular data – however, this latter is often called Fst in the literature, such as in this thesis and in the two articles in which the genetic results were published (Pakendorf et al. 2006, 2007).

Fst values are usually presented in tables (e.g. Table 3 in Pakendorf et al. 2007) and are often depicted either as trees or with the help of multidimensional scaling analysis (MDS). Since the Fst values estimate the genetic distances between populations, such trees and MDS plots are a way of depicting the relationships between the populations (cf. Figure 4 in Pakendorf et al. 2006, and Figure 2 in Pakendorf et al. 2007). MDS is a multivariate method that attempts to arrange the objects of study (e.g. population samples of mtDNA sequences) in geographical space in such a way that the distances between the objects are reproduced as accurately as possible (STATISTICA for Windows 1995: 3237). The advantages of MDS plots over trees is that trees will always join all the populations in a sample, whether or not this reflects the true genetic relationships. This can falsely create the impression of a genetic relationship where none exists. MDS plots are not constrained in this manner, and isolated populations will remain at a distance from all the other populations in the sample without being arbitrarily linked to any of them. Furthermore, trees are in effect unidimensional, while MDS analyses utilize more of the information present in the data by extracting several dimensions (although it is impossible to visualize more than three dimensions at once).

An alternative method to calculating genetic distances or Fst values and visualizing them in trees or MDS plots is to analyze the relationship of the actual sequences or STR haplotypes by constructing networks among them (cf. Figure 2a and 3a in Pakendorf et al. 2006). Since in a network every sequence or haplotype is

linked to those sequences or haplotypes that are most closely related to it, all possible trees are shown (Jobling, Hurles & Tyler-Smith 2004: 175ff).

A method that is conceptually based on Fst analysis is AMOVA (Analysis of Molecular Variance), which allows one to examine the effect of external factors (such as geography or language) on the genetic differences between populations. AMOVA estimates the proportion of variation at different levels of population groupings. One level examined is the genetic variation between populations within a group of populations defined by the researcher, based for example on linguistic affiliation or geographic source of the populations concerned. The second level is the genetic variation that exists between two or more such defined groups, while the third level is the variation present within the individual populations of the sample. If the groups defined by the researcher fit well with the underlying genetic variation, the genetic variation among the populations **within** the groups will be small, while the genetic variation **between** the different groups will be large. For example, one could group a sample of populations according to their linguistic affiliation; if linguistic relationship reflected genetic relationship, the populations grouped together under each language family label should be genetically homogenous. In such a case, the proportion of variance among the populations within each group (language family) would be small, while the proportion of variance between the individual groups (language families) would be large. If, on the other hand, linguistic relationship should NOT reflect genetic relationship, for instance because of prehistoric language shift, then the proportion of variance within the language family groups would be large, while the proportion of variance between the groups would be relatively small (cf. Table 5 in Pakendorf et al. 2006, where the proportion of variance within the linguistically defined groups is 4.06%, as opposed to only 1.94% of the variance being due to between-group variation).

338

APPENDIX 2:

Figure showing MDS plot based on Y-chromosomal SNPs in Eurasian populations;
not shown in Pakendorf et al. (2007). The data for most of the populations included
in the figure come from Karafet et al. (2002).
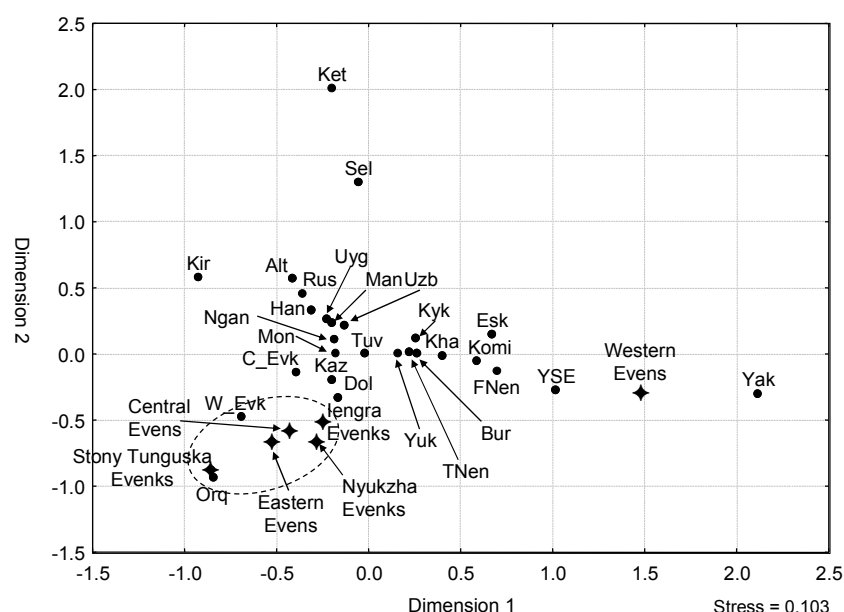
Alt = Altay; Bur = Buryats; C_Evk = Chinese Evenks; Dol = Dolgans; Esk =
Eskimos; FNen = Forest Nenets; Han = Han Chinese; Kaz = Kazakhs; Ket = Kets;
Kha = Khanty; Kir = Kirghiz; Komi = Komi; Kyk = Koryaks; Man = Manchu; Mon
= Mongols; Ngan = Nganasan; Orq = Oroqen; Sel = Selkup; TNen = Tundra Nenets;
Tuv = Tuvans; Uyg = Uyghurs; Uzb = Uzbek; W_Evk = Western Evenks; Yak =
Sakha; YSE = Sakha-speaking Evenks; Yuk = Yukaghirs.

The dotted line circles the cluster of groups speaking Northern Tungusic languages.
The Western Èvens are distinct from all the other Northern Tungusic groups, since
they have exceedingly high frequencies of haplogroup N-TatC. Interestingly, the
only other Tungusic group included in the analysis, the Manchu, do not cluster with
the Northern Tungusic groups, but are very close to Uyghurs, Uzbeks, and Han
Chinese, while the Dolgans do not cluster with the Sakha, but are close to the
Northern Tungusic cluster.

340

APPENDIX 3: Sakha case forms in the possessive declension

	Possessor					
	1SG	**2SG**	**3SG**	**1PL**	**2PL**	**3PL**
NOM	–(I)m	–(I)ŋ	–(t)A	–BIt	–GIt	–LArA
DAT	–BAr	–GAr	–(t)IgAr/ –Ar	–BItIgAr	–GItIgAr	–LArIgAr
ACC	–BIn	–GIn	–(t)In	–BItIn	–GItIn	–LArIn
PART	–BInA	–GInA	–(t)InA	–BItInA	–GItInA	–LArInA
ABL	–BIttAn	–GIttAn	–(t)IttAn	–BItIttAn	–GItIttAn	–LArIttAn
INSTR	–BInAn	–GInAn	–(t)InAn	–BItInAn	–GItInAn	–LArInAn
COMIT	–BInA:n/ –BInI:n	–GInA:n/ –GInI:n	–(t)InA:n/ –(t)InI:n	–BItInA:n/ –BItInI:n	–GItInA:n/ –GItInI:n	–LArInA:n/ –LArInI:n
COMPAR	BInA:γAr	GInA:γAr	(t)InA:γAr	BItInA:γAr	GItInA:γAr	LArInA:γAr

In the simple declension, the case forms are as follows:
NOM: unmarked
DAT: –*GA*
ACC: –*(n)I*
PART: –*TA*
ABL: –*(t)tAn*
INSTR: –*(I)nAn*
COMIT: –*LI:n*
COMPAR: –*TA:γAr*

The possessive formants for the Singular differ from the Nominative possessive suffixes, although the –*m* and –*B* of the 1SG and the –*ŋ* and –*G* of the 2SG are clearly related.

The following possessive elements can be postulated for the possessive declension:

1SG:	–*BI*	1PL:	–*BIt*
2SG:	–*GI*	2PL:	–*GIt*
3SG:	–*(t)I*	3PL:	–*LAr*

The following case suffixes can be postulated for the possessive declension:

ACC: *–(I)n*
PART: *–(I)nA*
DAT: *–(I)gAr* (with the 1SG and 2SG being contracted forms)
ABL: *–(I)ttAn*
INST: *–(I)nAn*
COM: *–(I)nI:n.–(I)nA:n*
COMP: *–(I)nA:γAr*

The epenthetic high vowel is added after the plural possessive formants, which end in consonants.

A common feature of the case suffixes of the possessive declension is the initial *–n–*; in the Accusative and Instrumental this is found in the simple declension as well, but in the Partitive, Comitative and Comparative this is a characteristic feature of the declension. The Dative is the only case for which the suffix of the possessive declension cannot be straightforwardly derived from the simple declension; the 1SG and 2SG forms of this case are also the most portmanteau-like.

APPENDIX 4: Case suffixes in the Tungusic languages

	Evenki	Èven	Negidal	Nanay	Ulča	Orok	Udihe	Oroč	Manchu
NOM	-ø	-ø	-ø	-ø	-ø	-ø	-ø	-ø	-ø
GEN									i/ni
ACC	-va	-β/-bu	-wa	-va	-va	-va	-va	-va	be
INDF.ACC	-(y)a		-(y)a						
DAT	-du	-du	-du	-do	-du	-du	-du	-du	de
ALL	-tki	-tki	-tki	-či	-tu	-tai/-taki	-tigi	-ti	
LOC	-la:/-dula:	-la/-dula:	-la/-dula	-la/-dola	-la/-dula	-la/-dula	-la/-dula	-la/-dula	
PROL	-li:/-duli:	-li/-duli	-li/-duli	-ki	-ki	-ki	-li/-dili	-li/-duli	
ALL-LOC	-kla	-kla							
ALL-PROL	-kli	-kli							
ABL	-duk	-duk	-dukkaj			-du/		-dui	
ELAT	-git	-gič		-diadi	-jiji	-dukti/-jedu	-digi	-jiji	ci
INS	-t/-di	-č/-ń	-ji	-di	-ji	-ji	-zi	-ži	
COM	-nu:n	-ńun				-ndo/-ndu			
DES	-(y)a	-ga		-goa		-do/-du	-na	-na:/-ya:/-la:	

The table is based primarily on the sketches given in the *Jazyki narodov SSSR 5* and the grammars listed in Table 2.3. The assimilatory processes that the suffixes are subject to are not indicated.

REFERENCES

ABONDOLO, DANIEL (ed). 1998. *The Uralic Languages*. London, New York: Routledge. (Reprinted as paperback in 2006).

—. 1998a. Khanty. In: Daniel Abondolo (ed), 358-86.

—. 1998b. Introduction. In: Daniel Abondolo (ed), 1-42.

AFANAS'EV, P.S. 1965. *Govor verxojanskix jakutov*. [The dialect of the Yakuts of Verxojansk]. Yakutsk: Jakutskoe knižnoe izdatel'stvo.

AFANAS'EV, P.S., VORONKIN, M.S. and ALEKSEEV, M.P. 1976. *Dialektologičeskij slovar' jakutskogo jazyka*. [Dialectological dictionary of Yakut]. Moscow: Izdatel'stvo 'Nauka'.

AIKHENVALD, ALEXANDRA Y. 1996. Areal diffusion in northwest Amazonia: The case of Tariana. *Anthropological Linguistics*, 38.73-116.

—. 1999. Areal diffusion and language contact in the Icana-Vaupes basin, northwest Amazonia. *The Amazonian Languages*, ed. by Robert M.W. Dixon and Alexandra Y. Aikhenvald, 384-416. Cambridge: Cambridge University Press.

—. 2002. *Language Contact in Amazonia*. Oxford, New York: Oxford University Press.

—. 2003a. Mechanisms of change in areal diffusion: new morphology and language contact. *J. Linguistics*, 39.1-29.

—. 2003b. Language contact and language change in Amazonia. *Historical Linguistics 2001. Selected Papers from the 15th International Conference on Historical Linguistics, Melbourne, 13-17 August 2001*, ed. by Barry J. Blake, Kate Burridge and Jo Taylor. Amsterdam/Philadelphia: John Benjamins Publishing Company.

ALEKSEENKO, E.A. 1980. Ketskaja problema. [The Ket problem]. In: Il'ja Samojlovič Gurvič (ed), 118-40.

ALEKSEEV, ANATOLIY NIKOLAEVICH. 1996. *Drevnjaja Jakutija - Neolit i epoxa bronzy*. [Ancient Yakutia: The Iron Age and the Medieval Epoch]. Novosibirsk: Izdatel'stvo Instituta Arxeologii i Etnografii SO RAN.

ANDERSEN, HENNING. 1988. Center and periphery: adoption, diffusion, and spread. *Historical Dialectology. Regional and Social*, ed. by Jacek Fisiak, 39-83. Berlin, New York, Amsterdam: Mouton de Gruyter.

ANDERSON, GREGORY D.S. 1998a. Historical aspects of Yakut (Saxa) phonology. *Turkic Languages*, 2.3-31.

—. 1998b. *Xakas*: Languages of the World/Materials. München, Newcastle: LINCOM Europa.

—. 2004. The Languages of Central Siberia. Introduction and Overview. *Languages and Prehistory of Central Siberia*, ed. by Edward J. Vajda, 1-119. Amsterdam, Philadelphia: John Benjamins Publishing Company.

ANDERSON, GREGORY DAVID and HARRISON, K. DAVID. 1999. *Tyvan*. München, Newcastle: LINCOM Europa.

ANIKIN, ALEKSANDR EVGEN'EVIČ. 2003. *Ètimologičeskij slovar' russkix zaimstvovanij v jazykax Sibiri*. [Etymological dictionary of Russian loanwords in the languages of Siberia]. Novosibirsk: 'Nauka'.

ANTONOV, N.K. 1971. *Materialy po istoričeskoj leksike jakutskogo jazyka*. [Materials on the historical lexicon of Yakut]. Yakutsk: Jakutskoe knizhnoe izdatel'stvo.

ARTEM'EV, NIKOLAJ M. 1999a. Dolganskij jazyk: Sovremennoe sostojanie, sootnošenie s jakutskim jazykom, sledy kontaktov, problemy issledovanija. [The Dolgan language: Current state, relationship with the Yakut language, traces of contact, problems of study.] *Issues in Turkic Languages - Description and Language Contact*, ed. by Setsu Fujishiro and Masahiro Shogaito, 43-50. Kyoto: Department of Linguistics, Faculty of Letters, Kyoto University.

—. 1999b. Častnyj padež v dolganskom jazyke v sravnenii s jakutskim i ėvenkijskim jazykami. [The Partitive case in Dolgan in comparison with Yakut and Evenki.] *Issues in Turkic Languages - Description and Language Contact*, ed. by Setsu Fujishiro and Masahiro Shogaito, 105-15. Kyoto: Department of Linguistics, Faculty of Letters, Kyoto University.

ATKNINE, VICTOR. 1997. The Evenki Language from the Yenisei to Sakhalin. *Northern Minority Languages. Problems of Survival. Papers Presented at the Eighteenth Taniguchi International Symposium: Division of Ethnology*, ed. by Hiroshi Shoji and Juha Janhunen, 109-21. Osaka: National Museum of Ethnology.

AVRORIN, VALENTIN ALEKSANDROVIČ. 1956. Predikativno-pritjažatel'nye formy v nanajskom i drugix tunguso-man'čžurskix jazykax. [The predicative-possessive forms in Nanay and other Tungusic languages]. *Voprosy jazykoznanija*, 3.93-100.

—. 1959. *Grammatika nanajskogo jazyka. Tom pervyj*. [A grammar of Nanay. vol. 1]. Moscow, Leningrad: Izdatel'stvo Akademii Nauk SSSR.

—. 1961. *Grammatika nanajskogo jazyka. Tom vtoroj*. [A grammar of Nanay. vol. 2]. Moscow, Leningrad: Izdatel'stvo Akademii Nauk SSSR.

—. 1981. *Sintaksičeskije issledovanija po nanajskomu jazyku*. [Studies of Nanay syntax]. Leningrad: 'Nauka'.

AVRORIN, VALENTIN A. and BOLDYREV, BORIS VASIL'EVIČ. 2001. *Grammatika oročskogo jazyka*. [A grammar of the Oroč language.]. Novosibirsk: Izdatel'stvo SO RAN.

BALZER, MARJORIE MANDELSTAM. 1994. Yakut. In: Paul Friedrich and Norma Diamond (eds), 404-07.

BEISENHERZ, ALBRECHT. 2001. Zum unveräußerlichen Besitz im Tungusischen. In: Winfried Boeder and Gerd Hentschel (eds), 31-47.

BEKKER, ĖRIKA G. 1995. *Morfologija sel'kupskogo jazyka. Južnye dialekty*. [The morphology of Selkup. Southern dialects]. Tomsk: Tomskij gosudarstvennyj pedagogičeskij institut.

BEHAR, D.M., ROSSET, S., BLUE-SMITH, J., BALANOVSKY O., TZUR S., COMAS, D., MITCHELL, R.J., QUINTANA-MURCI, L., TYLER-SMITH, C., WELLS, R.S., THE GENOGRAPHIC CONSORTIUM. 2007. The Genographic Project Public Participation Mitochondrial DNA Database. *PLoS Genetics*, 3.e104

BENZING, JOHANNES. [1942] 1993. Tschuwaschische Forschungen (IV). Die Kasus. *Bolgarisch-tschuwaschische Studien*., ed. by Claus Schönig, 39-88 (originally appeared in ZDMG 96: 421-70). Wiesbaden: Harrassowitz Verlag.

—. 1955. *Lamutische Grammatik*. Wiesbaden: Franz Steiner Verlag GmbH.

BERTA, ÁRPÁD. 1998. Tatar and Bashkir. In: Lars Johanson and Eva Csató (eds), 283-300.

BIRTALAN, AGNES. 2003. Oirat. In: Juha Janhunen (ed), 210-28.

BLACK, LIDIA T. 1988. Peoples of the Amur and Maritime Regions. *Crossroads of Continents: Cultures of Sibiria and Alaska*, ed. by William Fitzhugh and Aron Crowell, 24-31. Kensington, Maryland: Smithsonian Institution Press.

BLAKE, BARRY J. 2001. *Case*. Cambridge: Cambridge University Press.

BLÄSING, UWE. 2003. Kalmuck. In: Juha Janhunen (ed), 229-47.

BODROGLIGETI, ANDRÁS, J.E. 2003. *An Academic Reference Grammar of Modern Literary Uzbek*. München: LINCOM Europa.

BOEDER, WINFRIED and HENTSCHEL, GERD (eds). 2001. *Variierende Markierung von Nominalgruppen in Sprachen unterschiedlichen Typs*. Oldenburg: Bibliotheks- und Informationssystem der Universität Oldenburg.

BOESCHOTEN, HENRIK. 1998. Uzbek. In: Lars Johanson and Eva Csató (eds), 357-78.

BÖHTLINGK, OTTO NIKOLAEVIČ. [1851] 1964. *Über die Sprache der Jakuten*. Uralic and Altaic Series, vol. 35. The Hague: Mouton & Co.

BOLDYREV, BORIS VASIL'EVIČ. 1976. *Kategorija kosvennoj prinadležnosti v tunguso-man'čžurskix jazykax*. [The category of indirect possession in the Tungusic languages.]. Moscow: Izdatel'stvo 'Nauka'.

—. 2000. *Èvenkijsko-russkij slovar'*. [Evenk-Russian dictionary.]. Novosibirsk: Izdatel'stvo SO RAN, Filial "Geo".

BRODY, JILL. 1987. Particles Borrowed from Spanish as Discours Markers in Mayan Languages. *Anthropological Linguistics*, 29.507-21.

BULATOVA, NADEZHDA and GRENOBLE, LENORE. 1999. *Evenki*. Munich: Lincom Europa.

BURAEV, I.D. and L.D. ŠAGDAROV. 2004. Obrazovanie burjatskogo jazyka. [The formation of the Buryat language]. *Burjaty*, ed. by L.L. Abaeva, Žukovskaja, N.L., 227-32. Moscow: Nauka.

ČEREMISINA, MAJJA IVANOVNA. 1977. Deepričastija kak klass form glagola v jazykax raznyx sistem. [Gerunds as a class of verb forms in languages of different types] *Složnoe predloženie v jazykax raznyx sistem. (Sbornik naučnyx trudov)*, ed. by Elizaveta Ivanovna Ubrjatova and Majja Ivanovna Čeremisina, 3-28. Novosibirsk: Sibirskoe Otdelenie AN SSSR - Institut Istorii, Filologii i Filosofii.

CHAMBERS, JACK K. and TRUDGILL, PETER. 1998. *Dialectology*. Cambridge: Cambridge University Press.

CINCIUS, VERA I. (ed.) 1975. *Sravnitel'nyj slovar' tunguso-man'čžurskix jazykov. Materialy k ètimologičeskomu slovarju. Tom I*. [Comparative dictionary of the Tungusic languages. Materials for an etymological dictionary. Volume I]. Moscow: Izdatel'stvo 'Nauka', Leningradskoe otdelenie.

— (ed.) 1977. *Sravnitel'nyj slovar' tunguso-man'čžurskix jazykov. Materialy k ėtimologičeskomu slovarju. Tom II*. [Comparative dictionary of the Tungusic languages. Materials for an etymological dictionary. Volume II]. Leningrad: Izdatel'stvo 'Nauka', Leningradskoe otdelenie.

—. 1982. *Negidal'skij jazyk. Issledovanija i materialy*. Leningrad: Nauka.

CLARK, LARRY. 1998a. *Turkmen Reference Grammar*. Wiesbaden: Harrassowitz Verlag.

—. 1998b. Chuvash. In: Lars Johanson and Eva Csató (eds), 434-52.

CLYNE, MICHAEL. 2003. *Dynamics of language contact: English and immigrant languages*. Cambridge: Cambridge University Press. (cited from Johanson in print).

COMRIE, BERNARD. 1978. Definite direct objects and referent identification. *Pragmatics Microfiche*, 3.1.1-19.

—. 1981. *The Languages of the Soviet Union*. Cambridge: Cambridge University Press.

CONRAD, D. F., ANDREWS, T. D., CARTER, N. P., HURLES, M. E. and PRITCHARD, J. K. 2006. A high-resolution survey of deletion polymorphism in the human genome. *Nature Genetics*, 38.75-81.

CROFT, WILLIAM. 2003. Mixed languages and acts of identity: An evolutionary approach. *The mixed language debate*, ed. by Yaron Matras and Peter Bakker, 41-72. Berlin: Mouton de Gruyter.

DAHL, ÖSTEN. 1985. *Tense and Aspect Systems*. Oxford: Basil Blackwell.

DE REUSE, WILLEM JOSEPH. 1994. *Siberian Yupik Eskimo: The Language and Its Contacts with Chukchi*. Salt Lake City: University of Utah Press.

DELANCEY, SCOTT. 2003. Location and direction in Klamath. *Motion, Direction and Location in Languages. In honor of Zygmunt Frajzyngier*, ed. by Erin Shay and Uwe Seibert, 59-90. Amsterdam, Philadelphia: John Benjamins Publishing Company.

DERBENEVA, O. A., STARIKOVSKAIA, E. B., VOLOD'KO, N. V., WALLACE, D. C. and SUKERNIK, R. I. 2002. [Mitochondrial DNA variation in Kets and Nganasans and the early peoples of Northern Eurasia] *Genetika*, 38.1554-60.

DERENKO, M., MALYARCHUK, B., DENISOVA, G. A., WOZNIAK, M., DAMBUEVA, I., DORZHU, C., LUZINA, F., MISCICKA-SLIWKA, D. and ZAKHAROV, I. 2006. Contrasting patterns of Y-chromosome variation in South Siberian populations from Baikal and Altai-Sayan regions. *Human Genetics*, 118.591-604.

DERENKO, M. V., GRZYBOWSKI, T., MALYARCHUK, B. A., CZARNY, J., MISCICKA-SLIWKA, D. and ZAKHAROV, I. A. 2001. The presence of mitochondrial haplogroup x in Altaians from South Siberia. *American Journal of Human Genetics*, 69.237-41.

DOBRUSHINA, NINA and GOUSSEV, VALENTIN. 2005. Inclusive imperative. *Clusivity. Typology and case studies of the inclusive-exclusive distinction*, ed. by Elena Filimonova, 179-211. Amsterdam, Philadelphia: John Benjamins Publishing Company.

DOERFER, GERHARD. 1964. Klassifikation und Verbreitung der mongolischen Sprachen. *Handbuch der Orientalistik*, ed. by Spuler. B, 35-50. Leiden/Köln: E.J.Brill.

—. 1978. Classification Problems of Tungus. *Beiträge zur Nordasiatischen Kulturgeschichte*, ed. by Gerhard Doerfer and Michael Weiers, 1-26. Wiesbaden: Otto Harrassowitz.

—. 1988. *Grammatik des Chaladsch*. Wiesbaden: Otto Harrassowitz.

DOLGIX, BORIS OSIPOVIČ. 1960. *Rodovoy i plemennoy sostav narodov Sibiri v XVII veke*. [The tribal composition of the peoples of Siberia in the 17th century]. Moscow: Izdatel'stvo Akademii Nauk SSSR.

DORIAN, NANCY. 1993. Internally and externally motivated change in language contact settings: doubts about dichotomy. *Historical linguistics: Problems and perspectives*, ed. by Charles Jones, 131-55. London, New York: Longman.

Drevnetjurkskij slovar'. [Old Turkic dictionary]. 1969. Leningrad: Izdatel'stvo 'Nauka', Leningradskoe otdelenie.

DRYER, MATTHEW. 2005a. Prefixing versus suffixing in inflectional morphology. In: Martin Haspelmath, Matthew Dryer, David Gil and Bernard Comrie (eds), 110-13.

—. 2005b. Position of pronominal possessive affixes. In: Martin Haspelmath, Matthew Dryer, David Gil and Bernard Comrie (eds), 234-37.

—. 2005c. Order of subject, object, and verb. In: Martin Haspelmath, Matthew Dryer, David Gil and Bernard Comrie (eds), 330-33.

DUL'ZON, ANDREJ P. 1968. *Ketskij jazyk*. [The Ket language]. Tomsk: Izdatel'stvo Tomskogo Universiteta.

DUNN, MICHAEL J. 1999. *A Grammar of Chukchi*., Ph.D. thesis. Australian National University.

EFREMOV, NIKOLAJ NIKOLAEVIČ. 1979. O parnosti monosub"ektnyx i raznosub"ektnyx jakutskix temporal'nyx polipredikatvnyx konstrukcij. [On the pairwise occurrence of Yakut same-subject and different-subject temporal polypredicate constructions]. *Infinitnye formy glagola*, ed. by M.I. Čeremisina, 59-74. Novosibirsk: Sibirskoe otdelenie AN SSSR, Institut istorii, filologii i filosofii.

ENFIELD, N.J. 2003. *Linguistic Epidemiology. Semantics and Grammar of Language Contact in Mainland Southeast Asia*. London: Routledge.

ERDAL, MARCEL. 2004. *A Grammar of Old Turkic*. Leiden, Boston: Brill.

ERDÉLYI, ISTVÁN. [1970] 1997. *Selkupisches Wörterverzeichnis. Tas-Dialekt*. Uralic and Altaic Series, vol. 103. Richmond: Curzon Press Ltd.

EVSTIGNEEV, JURIJ ANDREEVIČ. 2003. *Rossijskaja Federacija. Narody i ix podrazdelenija*. [The Russian Federation. The peoples and their subdivisions]. St. Petersburg: Izdatel'stvo S.-Peterburgskogo universiteta.

FEDERAL'NAJA SLUŽBA GOSUDARSTVENNOJ STATISTIKI. 2004. *Nacional'nyj sostav i vladenie jazykami, graždanstvo. Itogi vserossijskoj perepisi naselenija 2002 goda*. [National composition and knowledge of language, citizenship. The results of the All-Russian population census of 2002] Moscow: IIC 'Statistika Rossii'.

FEDOTOV, MICHAIL. R. 1996. *Čuvašskij jazyk*. [Chuvash]. Čeboksary: Izdatel'stvo Čuvašskogo universiteta.

FERGUSON, CHARLES A. 1990. From esses to aitches: identifying pathways of diachronic change. *Studies in Typology and Diachrony. Papers presented to Joseph H. Greenberg on his 75th birthday*, ed. by William Croft, Keith Denning and Suzanne Kemmer, 59-78. Amsterdam/Philadelphia: John Benjamins Publishing Company.

FONDAHL, GAIL. 1994. Evenki (Northern Tungus). In: Paul Friedrich and Norma Diamond (eds), 120-24.

FORSYTH, J. 1992. *A History of the Peoples of Siberia. Russia's North Asian Colony 1581-1990*. Cambridge: Cambridge University Press.

FORTESCUE, MICHAEL. 1984. *West Greenlandic*. London, Sydney, Dover, New Hampshire: Croom Helm.

—. 1998. *Language Relations across Bering Strait: Reappraising the archeological and linguistic evidence*. London/ New York: Cassel.

FRAUMENE, C., BELLE, E. M., CASTRI, L., SANNA, S., MANCOSU, G., COSSO, M., MARRAS, F., BARBUJANI, G., PIRASTU, M. and ANGIUS, A. 2006. High resolution analysis and phylogenetic network construction using complete mtDNA sequences in sardinian genetic isolate. *Molecular Biology and Evolution*, 23.2101-11.

FRAURUD, KARI. 2001. Possessives with extensive use: a source of definite articles? *Dimensions of Possession*, ed. by Irene Baron, Michael Herslund and Soeren Finn. Amsterdam/Philadelphia: John Benjamins Publishing Company.

FRIEDRICH, PAUL and DIAMOND, NORMA (eds). 1994. *Encyclopedia of World Cultures, Volume 6 - Russia and Eurasia/China*, London et al.: G.K. Hall & Company.

GENSLER, ORIN DAVID. 1993. *A Typological Evaluation of Celtic/Hamito-Semitic Syntactic Parallels*, Ph.D. thesis, University of California at Berkeley.

GEORG, STEFAN. 2003. Ordos. In: Juha Janhunen (ed), 193-209.

GEORG, STEFAN and VOLODIN, ALEXANDER P. 1999. *Die itelmenische Sprache*. Wiesbaden: Harrassowitz Verlag.

GOGOLEV, ANATOLIJ IGNAT'EVIČ. 1993. *Jakuty. Problemy ètnogeneza i formirovanija kul'tury*. [The Yakuts. Problems of their ethnogenesis and the formation of their culture.]. Yakutsk: Izdatel'stvo JaGU.

GÖKSEL, ASLI and KERSLAKE, CELIA. 2005. *Turkish: A Comprehensive Grammar*. London: New York: Routledge.

GOLDEN, PETER B. 1998. The Turkic peoples: a historical sketch. In: Lars Johanson and Eva Csató (eds), 16-29.

GOLTSOVA, T. V., OSIPOVA, L. P., ZHADANOV, S. I. and VILLEMS, R. 2005. [The effect of marriage migration on the genetic structure of the Taimyr Nganasan population: genealogical analysis inferred from MtDNA markers] *Genetika*, 41.954-65.

GOODACRE, S., HELGASON, A., NICHOLSON, J., SOUTHAM, L., FERGUSON, L., HICKEY, E., VEGA, E., STEFANSSON, K., WARD, R. and SYKES, B. 2005. Genetic

evidence for a family-based Scandinavian settlement of Shetland and Orkney during the Viking periods. *Heredity*, 95.129-35.

GORDON, RAYMOND G. 2005. *Ethnologue. Languages of the World*. 15th edition. Dallas, TX: SIL.

GORELOVA, LILIYA M. 2002. *Manchu Grammar*. Leiden, Boston, Köln: Brill.

GRACE, GEORGE W. 1996. Regularity of change in what? *The Comparative Method Reviewed. Regularity and Irregularity in Language Change*, ed. by Mark Durie and Malcolm Ross, 157-79. New York, Oxford: Oxford University Press.

GRACHEVA, GALINA N. 1994. Dolgan. In: Paul Friedrich and Norma Diamond (eds), 99-103.

GRANT, ANTHONY P. 2003. Review of Ruth King. The lexical basis of grammatical borrowing: A Prince Edward Island case study. *Word*, 54.251-256.

Grammatika sovremennogo baškirskogo literaturnogo jazyka. [A grammar of the modern Bashkir literary language.]. 1981. Moskva: Izdatel'stvo 'Nauka'.

Grammatika sovremennogo jakutskogo literaturnogo jazyka. Fonetika i morfologija. [A grammar of the modern Yakut literary language. Phonetics and morphology.]. 1982. Moskva: Izdatel'stvo 'Nauka'.

Grammatika turkmenskogo jazyka. Čast' 1: fonetika i morfologija. 1970. Ašxabad: Izdatel'stvo 'Ylym'.

GRANOVETTER, M. 1973. The strength of weak ties. *American Journal of Sociology*, 78.1360-80 (cited from Milroy & Milroy 1985).

GRØNBECH, KAARE. [1936/1979] 1997. *The Structure of the Turkic Languages*. Uralic and Altaic Series, vol. 136. Richmond: Curzon Press Ltd.

GRUZDEVA, EKATERINA. 1998. *Nivkh*. Munich: LINCOM Europa.

GULYA, JÁNOS. [1966] 1997. *Eastern Ostyak Chrestomathy*. Uralic and Altaic Series, vol. 51. Richmond: Curzon Press Ltd.

GUMPERZ, JOHN J. and WILSON, ROBERT. 1971. Convergence and creolization. A case from the Indo-Aryan/Dravidian border in India. *Pidginization and Creolization of Languages. Proceedings of a Conference held at the University of the West Indies Mona, Jamaica, April 1968*, ed. by Dell Hymes, 151-67. Cambridge: Cambridge University Press.

GURVIČ, IL'JA SAMOJLOVIČ 1977. *Kul'tura severnyx jakutov-olenevodov* [The culture of the northern Yakut reindeer herders]. Moscow: Izdatel'stvo 'Nauka'.

— (ed). 1980. *Ètnogenez narodov severa* [The ethnogenesis of the peoples of the north]. Moscow: Izdatel'stvo 'Nauka'.

GURVIČ, IL'JA SAMOJLOVIČ and SIMČENKO, JU.B. 1980. Ètnogenez jukagirov. [The ethnogenesis of the Yukaghirs]. In: Il'ja Samojlovič Gurvič (ed), 141-51.

GUSEV, V.JU. 2005. *Tipologija specializirovannyx glagol'nyx form imperativa* [A typology of specialized imperative verb forms]. Ph.D. thesis, Institut jazykoznanija RAN, Moscow.

HAASE, MARTIN. 1992. *Sprachkontakt und Sprachwandel im Baskenland. Die Einflüsse des Gaskognischen und Französischen auf das Baskische. Hamburg*: Helmut Buske Verlag.

HAIMAN, JOHN and MUNRO, PAMELA. 1983. Introduction. *Switch-reference and Universal Grammar*, ed. by John Haiman and Pamela Munro, ix-xv. Amsterdam, Philadelphia: John Benjamins Publishing Company.

HARRIS, ALICE C. and CAMPBELL, LYLE. 1995. *Historical Syntax in Cross-Linguistic Perspective*. Cambridge: Cambridge University Press.

HASPELMATH, MARTIN. 1995. The converb as a cross-linguistically valid category. In: Martin Haspelmath and Ekkehard König (eds), 1-55.

HASPELMATH, MARTIN and KÖNIG, EKKEHARD (eds). 1995. *Converbs in Cross-Linguistic Perspective*. Berlin, New York: Mouton de Gruyter.

HASPELMATH, MARTIN, DRYER, MATTHEW, GIL, DAVID and COMRIE, BERNARD (eds.) 2005. *The World Atlas of Language Structures*. Oxford: Oxford University Press.

HAUGEN, EINAR. 1950. The Analysis of Linguistic Borrowing. *Language*, 26-2.210-31.

—. 1953. *The Norwegian Language in America: A Study in Bilingual Behaviour. Vol. I: The Bilingual Community; Vol. II: The American Dialects of Norwegian*. Bloomington: Indiana University Press (Reprinted 1969).

HEATH, JEFFREY. 1978. *Linguistic diffusion in Arnhem Land*. Canberra: Australian Institute of Aboriginal Studies.

—. 1981. A case of intensive lexical diffusion: Arnhem land, Australia. *Language,* 57.335-67.

HEINE, BERND. 1997. *Cognitive Foundations of Grammar*. New York, Oxford: Oxford University Press.

HEINE, BERND and KUTEVA, TANIA. 2002. *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.

—. 2003. On contact-induced grammaticalization. *Studies in Language*, 27.529-72.

—. 2005. *Language Contact and Grammatical Change*. Cambridge: Cambridge University Press.

HELGASON, A., SIGURDARDOTTIR, S., GULCHER, J. R., WARD, R. and STEFANSSON, K. 2000a. mtDNA and the origin of the Icelanders: deciphering signals of recent population history. *American Journal of Human Genetics*, 66.999-1016.

HELGASON, A., SIGURDARDOTTIR, S., NICHOLSON, J., SYKES, B., HILL, E. W., BRADLEY, D. G., BOSNES, V., GULCHER, J. R., WARD, R. and STEFANSSON, K. 2000b. Estimating Scandinavian and Gaelic ancestry in the male settlers of Iceland. *American Journal of Human Genetics*, 67.697-717.

HELIMSKI, EUGENE. 1998a. Selkup. In: Daniel Abondolo (ed), 548-79.

—. 1998b. Nganasan. In: Daniel Abondolo (ed), 480-515.

HERRNSTADT, C., ELSON, J. L., FAHY, E., PRESTON, G., TURNBULL, D. M., ANDERSON, C., GHOSH, S. S., OLEFSKY, J. M., BEAL, M. F., DAVIS, R. E. and HOWELL, N. 2002. Reduced-median-network analysis of complete mitochondrial DNA coding-region sequences for the major African, Asian, and European haplogroups. *American Journal of Human Genetics*, 70.1152-71.

HOCK, HANS HENRICH. 1991. *Principles of Historical Linguistics*. Berlin, New York: Mouton de Gruyter.

HONTI, LÁSZLÓ. 1998. ObUgrian. In: Daniel Abondolo (ed), 327-57.

IAFRATE, A. J., FEUK, L., RIVERA, M. N., LISTEWNIK, M. L., DONAHOE, P. K., QI, Y., SCHERER, S. W. and LEE, C. 2004. Detection of large-scale variation in the human genome. *Nature Genetics*, 36.949-51.

IMART, GUY. 1981. *Le Kirghiz. Description d'une langue de littérisation récente.* Aix-en-Provence: Publications de l'Université de Provence.

INGMAN, MAX, KAESSMANN, HENRIK, PÄÄBO, SVANTE and GYLLENSTEN, ULF. 2000. Mitochondrial genome variation and the origin of modern humans. *Nature*, 408.708-13.

ISENGALIEVA, VALENTINA AYTEŠEVNA. 1957. *Služebnye imena i poslelogi v kazaxskom jazyke*. [Auxiliary nouns and postpositions in Kazakh.]. Alma-Ata: Izdatel'stvo Akademii Nauk Kazaxskoj SSR.

ISXAKOV, F.G. and PAL'MBAX, A.A. 1961. *Grammatika tuvinskogo jazyka. Fonetika i morfologija*. [A grammar of Tuvan. Phonetics and morphology]. Moskva: Izdatel'stvo vostočnoj literatury.

JANHUNEN, JUHA. 1990. *Material on Manchurian Khamnigan Mongol*. Helsinki: Castrenianumin toimitteita.

—. 1991. *Material on Manchurian Khamnigan Evenki*. Helsinki: Castrenianumin toimitteita.

—. 1996. *Manchuria. An Ethnic History*. Mémoires de la Société Finno-Ougrienne, vol. 222. Helsinki: The Finno-Ugrian Society.

—. 1998a. Samoyedic. In: Daniel Abondolo (ed), 457-79.

—. 1998b. Ethnicity and language in prehistoric Northeast Asia. *Archaeology and Language II: Correlating archaeological and linguistic hypotheses*, ed. by Roger Blench and Matthew Spriggs, 195-208. London, New York: Routledge.

—. 2003a. Written Mongol. In: Juha Janhunen (ed), 30-56.

— (ed). 2003b. *The Mongolic Languages*. London/ New York: Routledge.

—. 2003c. Khamnigan Mongol. In: Juha Janhunen (ed), 83-101.

—. 2003d. Proto-Mongolic. In: Juha Janhunen (ed), 1-29.

JANKOWIAK, WILLIAM. 1994. Mongols. In: Paul Friedrich and Norma Diamond (eds), 473-76.

Jazyki narodov SSSR, V: Mongol'skie, tunguso-man'čžurskie i paleoaziatskie jazyki. [The languages of the USSR, vol. 5: Mongolic, Tungusic and Paleoasiatic languages]. 1968. Leningrad: Izdatel'stvo 'Nauka', Leningradskoe otdelenie.

JOBLING, M. A. and TYLER-SMITH, C. 2003. The human Y chromosome: an evolutionary marker comes of age. *Nature Reviews Genetics*, 4.598-612.

JOBLING, M.A., M.E. HURLES and TYLER-SMITH., C. 2004. *Human Evolutionary Genetics*: Garland Science, Taylor & Francis Group.

JOCHELSON, VALDIMIR I. [1926] 2005. *Yukagiry i yukagirizirovannie tungusy*. [Yukaghirs and yukaghirized Tungus]. Novosibirsk: Nauka.

JOHANSON, LARS. 1992. Strukturelle Faktoren in türkischen Sprachkontakten. *Sitzungsberichte der Wissenschaftlichen Gesellschaft an der Johann-Wolfgang-Goethe-Universität Frankfurt am Main*, 169-299. Stuttgart: Steiner.

—. 1998a. The structure of Turkic. In: Lars Johanson and Eva Csató (eds), 31-66.

—. 1998b. The history of Turkic. In: Lars Johanson and Eva Csató (eds), 81-125.

354

—. 1999. The dynamics of code-copying in language encounters. *Language encounters across time and space. Studies in language contact*, ed. by Bernt Brendemoen, Elizabeth Lanza and Else Ryen, 37-62. Oslo: Novus forlag.

—. 2001. Vom Alttürkischen zu den modernen Türksprachen. *Language Typology and Language Universals*, ed. by Martin Haspelmath, Ekkehard König, Wulf Oesterreicher and Wolfgang Raible, 1719-42. Berlin, New York: Walter de Gruyter.

—. 2002. Contact-induced change in a code-copying framework. *Language change: the interplay of internal, external and extra-linguistic factors*, ed. by Mari C. Jones; Edith Esch. Berlin, New York.

—. 2006. Turkic language contacts in a typology of code interaction. *Turkic Languages in Contact*, ed. by Henrik Boescheten and Lars Johanson, 4-26. Wiesbaden: Harrassowitz Verlag.

—.in print. Remodeling grammar: Copying, conventionalization, grammaticalization. *Language contact and contact languages.*, ed. by Peter Siemund and Noemi Kintana. Amsterdam/Philadelphia: J. Benjamins.

JOHANSON, LARS and CSATÓ, EVA (eds). 1998. *The Turkic Languages*. London, New York: Routledge. (Reprinted as paperback in 2006).

JORDE, L. B. and BAMSHAD, M. 2000. Questioning evidence for recombination in human mitochondrial DNA. *Science*, 288.1931a.

KAŁUŻYŃSKI, STANISŁAW. 1962. *Mongolische Elemente in der jakutischen Sprache*. Warsaw, 'S-Gravenhage: Panstwowe Wydawnictwo Naukowe, Mouton & Co.

—. 1982 [1995]. Einige tungusische Lehnwörter im Jakutischen. In: Stanisław Kałużyński, *IACUTICA. Prace jakutoznawcz*e, 225-32.

—. [1981] 1995. Ètimologičeskie issledovanija po jakutskomu jazyku. Dvusložnye osnovy (V). [Etymological studies of Yakut. Bisyllabic roots (V).]. In: Stanisław Kałużyński, *IACUTICA. Prace jakutoznawcz*e, 277-88.

—. 1995. *Iacutica. Prace jakutoznawcze*. Warsaw: Wydawnictwo Akademickie DIALOG.

KÄMPFE, HANS-RAINER. 1986. Činggis Khan. In: Michael Weiers (ed), 183-91.

KARAFET, T. M., OSIPOVA, L. P., GUBINA, M. A., POSUKH, O. L., ZEGURA, S. L. and HAMMER, M. F. 2002. High levels of Y-chromosome differentiation among native Siberian populations and the genetic signature of a boreal hunter-gatherer way of life. *Human Biology*, 74.761-89.

KARAFET, T. M., ZEGURA, S. L., POSUKH, O., OSIPOVA, L., BERGEN, A., LONG, J., GOLDMAN, D., KLITZ, W., HARIHARA, S., DE KNIJFF, P., WIEBE, V., GRIFFITHS, R. C., TEMPLETON, A. R. and HAMMER, M. F. 1999. Ancestral Asian source(s) of new world Y-chromosome founder haplotypes. *American Journal of Human Genetics*, 64.817-31.

KING, ROBERT D. 1969. *Historical Linguistics and Generative Grammar*. Englewood Cliffs, New Jersey: Prentice-Hall, Inc.

KONSTANTINOV, IVAN VASIL'EVIČ. [1975] 2003. *Proisxoždenie jakutskogo naroda i ego kul'tury*. [The origins of the Yakut people and their culture]. Yakutsk: Akademija nauk Respubliki Saxa (Jakutija), Institut gumanitarnyx issledovanij.

KONSTANTINOVA, OL'GA ALEKSANDROVNA. 1964. *Ėvenkijskij jazyk*. [The Evenk language]. Moscow, Leningrad: Izdatel'stvo 'Nauka'.

KOPTJEVSKAJA-TAMM, MARIA. 1995. Possessive and Relational Forms in Chukchi. *Double Case. Agreement by Suffixaufnahme*, ed. by Frans Plank. Oxford: Oxford University Press.

—. 2006. The circle that won't come full: two potential isoglosses in the Circum-Baltic area. *Linguistic Areas: Convergence in Historical and Typological Perspective*, ed. by Yaron Matras, April M.S. McMahon and Nigel Vincent, 182-225. New York: Palgrave McMillan.

KORKINA, EVDOKIJA INNOKENT'EVNA. 1970. *Naklonenija glagola v jakutskom jazyke*. [Verbal moods in Yakut]. Moscow: Izdatel'stvo 'Nauka'.

—. 1985. *Deepričastija v jakutskom jazyke*. [Converbs in Yakut]. Novosibirsk: Izdatel'stvo 'Nauka', Sibirskoe otdelenie.

KORNFILT, JAKLIN. 1997. *Turkish*. London, New York: Routledge.

KORTLANDT, FREDERIK. [1998] 2006. Are Mongolian and Tungus genetically related? Acta Orientalia (Academiae Scientiarum Hungaricae) (reprinted as pdf-file on internet: http://www.kortlandt.nl/publications/art156e.pdf), 51.235-37.

—. [2004] 2006. Indo-Uralic and Altaic. http://www.kortlandt.nl/publications/art216e.pdf.

KOŠKAREVA, NATALJA. 2001. The case systems of nouns and pronouns in Khanty. In: Winfried Boeder and Gerd Hentschel (eds), 235-54.

KRAYTSBERG, Y., SCHWARTZ, M., BROWN, T. A., EBRALIDSE, K., KUNZ, W. S., CLAYTON, D. A., VISSING, J. and KHRAPKO, K. 2004. Recombination of human mitochondrial DNA. *Science*, 304.981.

KREJNOVIČ, ERUXIM A. 1958. *Jukagirskij jazyk*. [The Yukaghir language.]. Moskva/Leningrad: Izdatel'stvo akademii nauk SSSR.

KRUEGER, JOHN R. [1961] 1997. *Chuvash Manual*. Uralic and Altaic Series, vol. 7. Richmond: Curzon Press Ltd.

—. 1962 [1997]. *Yakut Manual*. Uralic and Altaic Series, vol. 21. Richmond: Curzon Press Ltd.

KULICK, DON. 1992. *Language Shift and Cultural Reproduction. Socialization, Self, and Syncretism in a Papua New Guinean Village*. Cambridge, New York, Oakleigh: Cambridge University Press.

KULLMANN, RITA and TSERENPIL, D. 2001. *Mongolian Grammar*. Ulaanbaatar: Academy of Sciences, Institute of Language and Literature.

KUMAR, S., HEDRICK, P., DOWLING, T. and STONEKING, M. 2000. Questioning evidence for recombination in human mitochondrial DNA. *Science*, 288.1931a.

KÜNNAP, AGO. 1999. *Kamass*. München, Newcastle: LINCOM Europa.

KURILOV, GAVRIL NIKOLAEVIČ. 2001. *Jukagirsko-russkij slovar'*. [Yukaghir-Russian dictionary.]. Novosibirsk: Nauka.

—. 2005. Predislovie. [Foreword] In: Vladimir Jochelson [1926] 2005, 7-11.

KUZAKOVA, E.A. 1994. *Slovar' Mansi (vostočnyj dialekt)*. [Mansi dictionary (Eastern dialect)]. Moscow: RAN, Institut ėtnologii i antropologii im. N.N. Mikluxo-Maklaja.

KUZMENKOV, EVGENIJ A. 2001. Imperative verb forms in Mongolian. In: Viktor Xrakovskij (ed), 98-105.

KUZNECOVA, A.I, XELIMSKIJ, E.A. and GRUŠKINA, E.V. 1980. *Očerki po sel'kupskomu jazyku*. [Sketches of Selkup]. Moscow: Izdatel'stvo Moskovskogo universiteta.

LAHERMO, P., SAVONTAUS, M. L., SISTONEN, P., BERES, J., DE KNIJFF, P., AULA, P. and SAJANTILA, A. 1999. Y chromosomal polymorphisms reveal founding lineages in the Finns and the Saami. *European Journal of Human Genetics*, 7.447-58.

LE PAGE, ROBERT BROCK AND TABOURET-KELLER, ANDRÉE. 1985. *Acts of identity: Creole-based approaches to language and ethnicity*. Cambridge: Cambridge University Press.

LEIGHTON, RALPH and BICHELDEI, K.A. 1994. Tuvans. In: Paul Friedrich and Norma Diamond (eds), 372-75.

LESSING, FERDINAND D. 1995. *Mongolian-English Dictionary* (3rd reprinting). Bloomington: The Mongolia Society, Inc.

LEVELT, WILLEM J.M. 1992. Accessing words in speech production: Stages, processes and representations. *Cognition*, 42.1-22.

MAAK, RIXARD KARLOVIČ. [1883-1887] 1994. *Viljujskij okrug*. [The Vilyuy district.]. Moscow: a/o 'Jana'.

MAIDEN, MARTIN and ROBUSTELLI, CECILIA. 2000. *A Reference Grammar of Modern Italian*. Chicago: NTC Publishing Group.

MAL'ČUKOV, ANDREJ L'VOVIČ. 1999. *Sintaksis prostogo predloženija v ėvenskom jazyke. Strukturnye i semantičeskie aspekty*. [The syntax of simple clauses in Even. Structural and semantic aspects]. St Petersburg: 'Nauka'.

MALCHUKOV, ANDREI L. 1995. *Even*. München, Newcastle: LINCOM Europa.

—. 2001. Imperative Constructions in Even. In: Viktor Xrakovskij (ed), 160-80.

—. 2006. Yakut interference in North-Tungusic languages. *Turkic Languages in Contact*, ed. by Henrik Boescheten and Lars Johanson, 122-38. Wiesbaden: Harrassowitz Verlag.

MARTIN, SAMUEL E. [1961] 1997. *Dagur Mongolian Grammar, Texts, and Lexicon Based on the Speech of Peter Onon*. Uralic and Altaic Series, vol. 4. Richmond: Curzon Press Ltd.

MASLOVA, ELENA. 2003a. *Tundra Yukaghir*. Munich: LINCOM.

—. 2003b. *A Grammar of Kolyma Yukaghir*. Berlin, New York: Mouton de Gruyter.

MASLOVA, ELENA SERGEEVNA and VAXTIN, NIKOLAJ BORISOVIČ. 1996. The Far North-East of Russia. In: Stephen A. Wurm, Peter Mühlhäusler and Darrell T. Tryon (eds), 999-1001.

MATRAS, YARON. 1998. Utterance modifiers and universals of grammatical borrowing. *Linguistics*, 36.281-331.

—. 2000. How predictable is contact-induced change in grammar? *Time Depth in Historical Linguistics*, ed. by Colin Renfrew, April McMahon and Larry Trask, 563-83. Cambridge: The McDonald Institute for Archaeological Research.

MATTISSEN, JOHANNA. 2003. *Dependent-Head Synthesis in Nivkh: a contribution to a typology of polysynthesis*. Amsterdam/Philadelphia: J. Benjamins.

MENGES, KARL H. 1978. Jakutisch und Tungusisch. *Central Asiatic Journal*, XXII.245-62.

MENOVŠČIKOV, GEORGIJ ALEKSEEVIČ. 1980. J*azyk ėskimosov beringova proliva*. [The language of the Bering Strait Eskimos]. Leningrad: Izdatel'stvo "Nauka", leningradskoe otdelenie.

MERLINGEN, WERIAND. 1977. *Artikulation und Phonematik des H. Mit einem Anhang über Phonematik*. Vienna: Verband der wissenschaftlichen Gesellschaften Österreichs.

MILROY, JAMES. 1997. Internal vs. external motivations for linguistic change. *Multilingua*, 16-4.311-23.

MILROY, JAMES and MILROY, LESLEY. 1985. Linguistic change, social network and speaker innovation. *J. Linguistics*, 21.339-84.

MORAVCSIK, EDITH A. 1978. Language contact. *Universals of Language volume 1. Method & Theory*, ed. by Joseph H. Greenberg, 93-122. Stanford: Stanford University Press.

MURPHY, LAWRENCE WALTER. 1968. *Sosva Vogul Grammar*, Ph.D. thesis, Indiana University.

MYERS-SCOTTON, CAROL. 2002. *Contact Linguistics: Bilingual Encounters and Grammatical Outcomes*. Oxford, New York: Oxford University Press.

MYREEVA, ANNA NIKOLAEVNA. 1962. Očerk tommotskogo govora. [A sketch of the Tommot dialect.]. *Očerki tokkinskogo i tommotskogo dialekto*v, ed. by Agnija Vasil'evna Romanova and Anna Nikolaevna Myreeva, 50-104. Moscow, Leningrad: Izdatel'stvo akademii nauk SSSR.

—. 1964. Očerk govora učurskix ėvenkov. [A sketch of the dialect of the Učur Evenks.]. *Očerki Učurskogo, majskogo i tottinskogo govorov*, ed. by Agnija Vasil'evna Romanova and Anna Nikolaevna Myreeva, 6-74. Moscow, Leningrad: Izdatel'stvo 'Nauka'.

—. 2004. *Evenkijsko-russkij slovar′*. Novosibirsk: Nauka.

NADELJAEV, V.M. 1968. Circumbaikal'skij jazykovoj sojuz. *Issledovanija po fonetike jazykov i dialektov Sibiri*, 3-4 (cited from Skribnik 2004). Novosibirsk.

NASIDZE, I., QUINQUE, D., RAHMANI, M., ALEMOHAMAD, S. A. and STONEKING, M. 2006. Concomitant replacement of language and mtDNA in South Caspian populations of Iran. *Current Biology*, 16.668-73.

NASIDZE, I., QUINQUE, D., DUPANLOUP, I., CORDAUX, R., KOKSHUNOVA, L. and STONEKING, M. 2005. Genetic evidence for the Mongolian ancestry of Kalmyks. *American Journal of Physical Anthropoly*, 128.846-54.

NASIDZE, I., LING, E. Y., QUINQUE, D., DUPANLOUP, I., CORDAUX, R., RYCHKOV, S., NAUMOVA, O., ZHUKOVA, O., SARRAF-ZADEGAN, N., NADERI, G. A., ASGARY, S., SARDAS, S., FARHUD, D. D., SARKISIAN, T., ASADOV, C., KERIMOV, A. and STONEKING, M. 2004. Mitochondrial DNA and Y-chromosome variation in the caucasus. *Annals of Human Genetics*, 68.205-21.

NASILOV, DMITRIJ, ISXAKOVA, XORSHID, SAFAROV, SHAXRIJOR and NEVSKAJA, IRINA. 2001. Imperative sentences in Turkic Languages. In: Viktor Xrakovskij (ed), 181-220.

358

NEDJALKOV, IGOR' V. 1995. Converbs in Evenki. In: Martin Haspelmath and Ekkehard König (eds), 441-63.

—. 1997. *Evenki*. London: Routledge.

NEDJALKOV, VLADIMIR P. 1995. Some typological parameters of converbs. In: Martin Haspelmath and Ekkehard König (eds), 97-136.

NEVSKAYA, IRINA. 2001. Varying marking of spatial arguments in South Siberian Turkic languages: a static dative and a dynamic locative. In: Winfried Boeder and Gerd Hentschel (eds), 293-303.

—. 2005. Inclusive and exclusive in Turkic languages. *Clusivity. Typology and case studies of the inclusive-exclusive distinction*, ed. by Elena Filimonova, 341-58. Amsterdam, Philadelphia: John Benjamins Publishing Company.

NEWMAN, JOHN. 1998. Recipients and 'give' constructions. *The Dative. Volume 2: Theoretical and Contrastive Studies*, ed. by Willy van Langendonck and William van Belle, 1-28. Amsterdam, Philadelphia: John Benjamins Publishing Company.

NICHOLS, JOHANNA and BICKEL, BALTHASAR. 2005. Possessive Classification. In: Martin Haspelmath, Matthew Dryer, David Gil and Bernard Comrie (eds), 242-43, map 59.

NIKOLAEVA, IRINA. 1999. *Ostyak*. Munich: LINCOM Europa.

—. 2003. Possessive affixes in the pragmatic structuring of the utterance: Evidence from Uralic. *International Symposium on Deictic Systems and Quantification in Languages Spoken in Europe and North and Central Asia. Collection of Papers.*, ed. by Bernard Comrie and Pirkko Suihkonen, 130-45. Izhevsk, Leipzig: Udmurt State University, Max-Planck Institute for Evolutionary Anthropology.

—. 2006. *A Historical Dictionary of Yukaghir*. Berlin, New York: Mouton de Gruyter.

NIKOLAEVA, IRINA and TOLSKAYA, MARIA. 2001. *A Grammar of Udihe*. Berlin, New York: Mouton de Gruyter.

NIMAEV, D.D. 2004. Načalo formirovanija ètničeskogo jadra burjat. [The beginning of the formation of the ethnic core of the Buryats]. *Burjaty*, ed. by L.L. Abaeva, Žukovskaja, N.L., 13-25. Moscow: Nauka.

NOVIKOVA, KLAVDIJA A. 1960. *Očerki dialektov evenskogo jazyka*. [Sketches of Even dialects]. Moscow, Leningrad: Izdatel'stvo Akademii Nauk SSSR.

—. 1980. *Očerki dialektov evenskogo jazyka. Glagol, služebnye slova, teksty, glossarij*. [Sketches of Even dialects. Verb, auxiliary words, texts, glossary]. Leningrad: "Nauka".

OKLADNIKOV, A. P. 1955. *Istorija Jakutskoj ASSR. Tom 1: Jakutija do prisoedinenija k russkomu gosudarstvu*. [The History of the Yakut ASSR. Vol 1: Yakutia before Its Incorporation into the Russian State]. Moscow, Leningrad: Izdatel'stvo Akademii Nauk SSSR.

OKSAAR, ELS. 1999. Social networks, communicative acts and the multilingual individual. *Language Change. Advances in Historical Sociolinguistics*, ed. by Ernst Håkon Jahr, 3-19. Berlin, New York: Mouton de Gruyter.

OLIVIERI, A., ACHILLI, A., PALA, M., BATTAGLIA, V., FORNARINO, S., AL-ZAHERY, N., SCOZZARI, R., CRUCIANI, F., BEHAR, D. M., DUGOUJON, J. M., COUDRAY, C., SANTACHIARA-BENERECETTI, A. S., SEMINO, O., BANDELT, H. J. and TORRONI, A. 2006. The mtDNA legacy of the Levantine early Upper Palaeolithic in Africa. *Science*, 314.1767-70.

ONENKO, S.N. 1980. *Nanajsko-Russkij Slovar'*. [Nanay-Russian dictionary]. Moscow: Izdatel'stvo "Russkij jazyk".

OOTA, H., SETTHEETHAM-ISHIDA, W., TIWAWECH, D., ISHIDA, T. and STONEKING, M. 2001. Human mtDNA and Y-chromosome variation is correlated with matrilocal versus patrilocal residence. *Nature Genetics*, 29.20-1.

PAKENDORF, B. and STONEKING, M. 2005. Mitochondrial DNA and human evolution. *Annual Revue of Genomics and Human Genetics*, 6.165-83.

PAKENDORF, B., NOVGORODOV, I. N., OSAKOVSKIJ, V. L., DANILOVA, A. P., PROTOD'JAKONOV, A. P. and STONEKING, M. 2006. Investigating the effects of prehistoric migrations in Siberia: genetic variation and the origins of Yakuts. *Human Genetics*, 120.334-53.

PAKENDORF, B., MORAR, B., TARSKAIA, L. A., KAYSER, M., SOODYALL, H., RODEWALD, A. and STONEKING, M. 2002. Y-chromosomal evidence for a strong reduction in male population size of Yakuts. *Human Genetics*, 110.198-200.

PAKENDORF, BRIGITTE. 2001. Genetic and linguistic perspectives on the prehistory of the Yakuts. *Turkic Languages*, 5.138-43.

PAKENDORF, BRIGITTE and NOVGORODOV, INNOKENTIJ N. in preparation. Loanwords in Sakha. *Loanword Typology*, ed. by Martin Haspelmath and Uri Tadmor.

PAKENDORF, BRIGITTE, NOVGORODOV, INNOKENTIJ N., OSAKOVSKIJ, VLADIMIR L. and STONEKING, MARK. 2007. Mating patterns amongst Siberian reindeer herders: Inferences from mtDNA and Y-chromosomal analyses. *American Journal of Physical Anthropology*, 133.1013-27.

PAKENDORF, BRIGITTE, WIEBE, VICTOR, TARSKAIA, LARISSA A., SPITSYN, VICTOR A., SOODYALL, HIMLA, RODEWALD, ALEXANDER and STONEKING, MARK. 2003. Mitochondrial DNA evidence for admixed origins of central Siberian populations. *American Journal of Physical Anthropology*, 120.211-24.

PANFILOV, VLADIMIR. Z. 1962. *Grammatika nivxskogo jazyka. Čast' 1*. [A grammar of the Nivkh language. Part 1].vol. 1. Moscow: Leningrad: Izdatel'stvo Akademii Nauk SSSR.

PASSARINO, G., SEMINO, O., QUINTANA-MURCI, L., EXCOFFIER, L., HAMMER, M. and SANTACHIARA-BENERECETTI, A. S. 1998. Different genetic components in the Ethiopian population, identified by mtDNA and Y-chromosome polymorphisms. *American Journal of Human Genetics*, 62.420-34.

PEKARSKIJ, EDUARD K. [1907-1930] 1958-1959. *Slovar' jakutskogo jazyka*. [Dictionary of the Yakut language]. Yakutsk: Jakutskij filial Akademii nauk SSSR.

PETROVA, TAISIJA IVANOVNA. 1967. *Jazyk orokov (Ul'ta)*. [The language of the Oroks (Ulta)]. Leningrad: Nauka.

PLANK, FRANS (ed.) 1995. *Double Case. Agreement by Suffixaufnahme*. New York, Oxford: Oxford University Press.

Pis'mennye jazyki mira: Jazyki Rossijskoj Federacii. Sociolingvističeskaja ènciklopedija, kniga 1 [The written languages of the world. Languages of the Russian Federation. Sociolinguistic Encyclopedia, book 1]. 2000. Moscow: Institut jazykoznanija RAN.

Pis'mennye jazyki mira: Jazyki Rossijskoj Federacii. Sociolingvističeskaja ènciklopedija, kniga 2. [The written languages of the world. Languages of the Russian Federation. Sociolinguistic Encyclopedia, book 2]. 2003. Moscow: Academia.

POLINSKAJA, MARIA S. and NEDJALKOV, VLADIMIR P. 1987. Contrasting the Absolutive in Chukchee. *Studies in Ergativity*, ed. by Robert M.W. Dixon, 239-69. Amsterdam, New York, Oxford, Tokyo: Elsevier Science Publishers B.V. (North-Holland).

POPOV, GAVRIIL VASIL'EVIČ. 1986. *Slova "neizvestnogo proisxoždenija" jakutskogo jazyka (sravnitel'no-istoričeskoe issledovanie)*. [The Words of "Unkown Origin" in Yakut (Historical Comparative Analysis)]. Yakutsk: Jakutskoe knižnoe izdatel'stvo.

POPOV, NIKOLAJ ANISIMOVIČ. 2000. *Ėristin. Rasskazy*. [Ėristin. Short stories.]. St. Petersburg: Otdelenie izdatel'stva "Prosveščenie".

POPPE, NICHOLAS. N. [1954] 1991. *Grammar of Written Mongolian*. Wiesbaden: Otto Harrassowitz.

—. 1955. *Introduction to Mongolian Comparative Studies*. Mémoires de la Société Finno-Ougrienne, vol. 110. Helsinki: Suomalais-Ugrilainen Seura.

—. 1959. Das Jakutische. *Philologiae Turcicae Fundamenta*, ed. by Jean Deny, Kaare Gronbech, Helmut Scheel and Zeki Velidi Togan, 671-84. Wiesbaden: Franz Steiner Verlag.

—. 1960. *Buriat Grammar*. Bloomington: Indiana University, The Hague: Mouton & Co.

RADLOFF, WILHELM. 1908. *Die jakutische Sprache in ihrem Verhältnisse zu den Türksprachen*. Mémoires de l'Académie impériale des Sciences de St-Pétersbourg. VIII Série. Classe historico-philologique. St. Petersburg.

RAGIMOV, M.Š. 1966. *Istorija formirovanija naklonenij glagola v azerbajdžanskom jazyke*. [The history of formation of the verbal moods in Azerbaijanian]. Baku (cited from Korkina 1970).

RAMSTEDT, GUSTAV JOHN. [1903] 1968. *Über die Konjugation des Khalkha-Mongolischen*. Osnabrück: Otto Zeller.

RASSADIN, VALENTIN IVANOVIČ. 1971. *Fonetika i leksika tofalarskogo jazyka*. [The phonetics and lexicon of Tofa]. Ulan-Ude: Burjatskoe knižnoe izdatel'stvo.

—. 1978. *Morfologija tofalarskogo jazyka v sravnitel'nom osveščenii*. [The morphology of Tofa in comparative perspective]. Moscow: Izdatel'stvo 'Nauka'.

—. 1980. *Mongolo-burjatskie zaimstvovanija v sibirskix tjurksix jazykax*. [Mongol-Buryat Borrowings in the Turkic Languages of Siberia]. Moscow: Izdatel'stvo 'Nauka'.

RAUN, ALO. 1988. Comparative-historical morphosyntax. *The Uralic Languages. Description, History and Foreign Influences*, ed. by Denis Sinor, 555-71. Leiden, New York, Copenhagen, Cologne: E.J. Brill.

REIDLA, M., KIVISILD, T., METSPALU, E., KALDMA, K., TAMBETS, K., TOLK, H. V., PARIK, J., LOOGVALI, E. L., DERENKO, M., MALYARCHUK, B., BERMISHEVA, M., ZHADANOV, S., PENNARUN, E., GUBINA, M., GOLUBENKO, M., DAMBA, L., FEDOROVA, S., GUSAR, V., GRECHANINA, E., MIKEREZI, I., MOISAN, J. P., CHAVENTRE, A., KHUSNUTDINOVA, E., OSIPOVA, L., STEPANOV, V., VOEVODA, M., ACHILLI, A., RENGO, C., RICKARDS, O., DE STEFANO, G. F., PAPIHA, S., BECKMAN, L., JANICIJEVIC, B., RUDAN, P., ANAGNOU, N., MICHALODIMITRAKIS, E., KOZIEL, S., USANGA, E., GEBERHIWOT, T., HERRNSTADT, C., HOWELL, N., TORRONI, A. and VILLEMS, R. 2003. Origin and diffusion of mtDNA haplogroup X. *American Journal of Human Genetics*, 73.1178-90.

RIESE, TIMOTHY. 2001. *Vogul*. Munich: LINCOM Europa.

RIŠES, L.D. and CINCIUS, VERA I. 1952. Kratkij oček grammatiki èvenskogo (lamutskogo) jazyka. [A short grammar sketch of the Even (Lamut) language]. *Russko-Èvenskij Slovar'*, 693-777. Moscow: Gosudarstvennoe izdatel'stvo inostrannyx i nacional'nyx slovarej.

ROBBEETS, MARTINE IRMA. 2005. *Is Japanese related to Korean, Tungusic, Mongolic and Turkic?* Wiesbaden: Harrassowitz Verlag.

ROGERS, E.M. and SHOEMAKER, F.F. 1971. *Communication Innovations*. 2nd edition. New York: Free Press (cited from Milroy and Milroy 1985).

ROMANOVA, AGNIJA VASIL'EVNA. 1962. Oček tokkinskogo govora. [A sketch of the Tokko dialect.]. *Očerki tokkinskogo i tommotskogo dialektov*, ed. by Agnija Vasil'evna Romanova and Anna Nikolaevna Myreeva, 8-49. Moscow, Leningrad: Izdatel'stvo akademii nauk SSSR.

—. 1964. Govory ajano-majskix èvenkov. [The dialects of the Ajano-Maja Evenks.]. *Očerki učurskogo, majskogo i tottinskogo govorov*, ed. by Agnija Vasil'evna Romanova and Anna Nikolaevna Myreeva, 75-153. Moscow, Leningrad: Izdatel'stvo 'Nauka'.

ROMANOVA, AGNIJA VASIL'EVNA, MYREEVA, ANNA NIKOLAEVNA and BARAŠKOV, PETR PETROVIČ. 1975. *Vzaimovlijanie èvenkijskogo i jakutskogo jazykov*. [Mutual Influence of the Evenk and Yakut Languages]. Leningrad: Izdatel'stvo "Nauka" - Leningradskoe Otdelenie.

ROMBANDEEVA, EVDOKIJA IVANOVNA. 1973. *Mansijskij (Vogul'skij) jazyk*. [The Mansi (Vogul) language]. Moscow: Izdatel'svto 'Nauka'.

—. 1979. *Sintaksis mansijskogo (vogul'skogo) jazyka*. [Syntax of the Mansi (Vogul) language]. Moscow: Izdatel'stvo 'Nauka'.

ROOTSI, S., ZHIVOTOVSKY, L. A., BALDOVIC, M., KAYSER, M., KUTUEV, I. A., KHUSAINOVA, R., BERMISHEVA, M. A., GUBINA, M., FEDOROVA, S. A., ILUMAE, A. M., KHUSNUTDINOVA, E. K., VOEVODA, M. I., OSIPOVA, L. P., STONEKING, M., LIN, A. A., FERAK, V., PARIK, J., KIVISILD, T., UNDERHILL, P. A. and VILLEMS, R. 2007. A counter-clockwise northern route of the Y-chromosome

362

haplogroup N from Southeast Asia towards Europe. *European Journal of Human Genetics*, 15.204-11.

ROSS, MALCOLM. 1996. Contact-induced change and the comparative method: Cases from Papua New Guinea. *The Comparative Method Reviewed. Regularity and Irregularity in Language Change*, ed. by Mark Durie and Malcolm Ross, 180-217. New York, Oxford: Oxford University Press.

—. 1997. Social networks and kinds of speech-community events. *Archaeology and Language I. Theoretical and methodological orientations*, ed. by Roger Blench and Matthew Spriggs, 209-61. London, New York: Routledge.

—. 2001. Contact-Induced Change in Oceanic Languages in North-West Melanesia. *Areal Diffusion and Genetic Inheritance*, ed. by Alexandra Y. Aikhenvald and Robert M.W. Dixon, 134-66. Oxford: Oxford University Press.

—. 2003. Diagnosing prehistoric language contact. *Motives for Language Change*, ed. by Raymond Hickey, 174-98. Cambridge: Cambridge University Press.

RUHLEN, M. 1991. *A Guide to the World's Languages. Volume 1: Classification*. Stanford: Stanford University Press.

RYBATZKI, VOLKER. 2003. Middle Mongol. In: Juha Janhunen (ed), 57-82.

SADOCK, JERROLD M. and ZWICKY, ARNOLD M. 1996. Speech act distinctions in syntax. *Language Typology and syntactic description. Volume I: Clause structure*, ed. by Timothy Shopen, 155-96. Cambridge: Cambridge University Press.

SAFRONOV, F. G. (ed.) 2000. *Ėnciklopedija Jakutii* [Encyclopedia of Yakutia]. Moscow: Jakutskij naučnyj centr sibirskogo otdelenija Rossijskoj akademii nauk, Akademija nauk Respubliki Saxa (Jakutija), OOO "Jakutskaja Ėnciklopedija".

SAMMALLAHTI, PEKKA. 1988. Historical Phonology of the Uralic Languages. *The Uralic Languages. Description, History and Foreign Influences*, ed. by Denis Sinor, 478-554. Leiden, New York, Copenhagen, Cologne: E.J. Brill.

SANŽEEV, GARMA DANCARANOVIČ. 1953. *Sravnitel'naja grammatika mongol'skix jazykov*. [A comparative grammar of the Mongolic languages]. Moscow: Izdatel'stvo Akademii Nauk SSSR.

—. 1964. *Sravnitel'naja grammatika mongol'skix jazykov: Glagol* [A comparative grammar of the Mongolic languages. The verb]. Moskva: Nauka.

SAUER, MARTIN. 1803. *Reise nach Siberien, Kamtschatka und zur Untersuchung der Mündung des Kowima-Flusses, der ganzen Küste der Tschutschken und der zwischen dem festen Lande von Asien und Amerika befindlichen Inseln. Auf Befehl der Kaiserin von Rußland, Catharina der Zweyten, in den Jahren 1785 bis 1794 unternommen vom Kapitän Joseph Billings und nach den Original-Papieren herausgegeben von Martin Sauer, Sekretär der Expedition*. Aus dem Englischen übersetzt. Berlin und Hamburg.

ŠČERBAK, ALEKSANDR MIXAJLOVIČ. 1970. *Sravnitel'naja fonetika tjurkskix jazykov*. [Comparative phonetics of the Turkic languages]. Leningrad: Izdate'stvo 'Nauka' - Leningradskoe otdelenie.

—. 1994. *Vvedenie v sravnitel'noe izučenie tjurkskix jazykov*. [An introduction to the comparative study of the Turkic languages]. St. Petersburg: 'Nauka'.

SCHÖNIG, CLAUS. [1987] 2003. Diachronic and areal approach to the Turkic imperative paradigm. *Electronic Journal of Oriental Studies*, VI.1-15.

—. [1988] 1990. Materialien zur Stellung des Lenatürkischen unter den Türksprachen. *Materialia Turcica*, ed. by Hermann Váry, 41-57. Bochum: Universitätsverlag Dr. N. Brockmeyer.

—. 1990. Classification problems of Yakut. *L'asie centrale et ses voisins*, ed. by Rémy Dor, 91-102. Paris: INALCO.

—. 1993. Türkische Sprachkontakte in Ostsibirien. *Sprach- und Kulturkontakte der türkischen Völker: Materialien der zweiten Deutschen Turkologen-Konferenz, Rauischholzhausen,13.-16. Juli 1990*, ed. by Jens Peter Laut and Klaus Röhrborn, 155-63. Wiesbaden: Harrassowitz Verlag.

—. 1997. A new attempt to classify the Turkic languages (1). *Turkic Languages*, 1.117-33.

—. 1998. Azerbaijanian. In: Lars Johanson and Eva Csató (eds), 248-60.

SCHWARTZ, MARIANNE and VISSING, JOHN. 2002. Paternal inheritance of mitochondrial DNA. *New England Journal of Medicine*, 347.576-80.

SEREBRENNIKOV, B.A. 1973. O pričinax prevraščenija načal'nogo s v h v baškirskom jazyke. [On the reasons for the change of initial s to h in Bashkir]. *Sovetskaja Tjurkologija*.10-15.

SEROŠEVSKIJ, VACLAV LEOPOL'DOVIČ. [1896] 1993. *Jakuty*. [The Yakuts]. Moscow: Rossijskaja političeskaja enciklopedija (ROSSPEN).

Severnaja Ėnciklopedija. 2004. Moscow: Evropejskie izdanija & Severnye prostory.

SEVORTJAN, E.V. 1974. *Ėtimologičeskij slovar' tjurskix jazykov. Obščetjurkskie i mežtjurkskie leksičeskie osnovy na glasnye*. [Etymological Dictionary of the Turkic Languages. Common Turkic and Interturkic Lexical Roots on Vowels].vol. 1. Moscow: Izdatel'stvo 'Nauka'.

—. 1978. *Ėtimologičeskij slovar' tjurskix jazykov. Obščetjurkskie i mežtjurkskie leksičeskie osnovy na bukvu B*. [Etymological Dictionary of the Turkic Languages. Common Turkic and Interturkic Lexical Roots on B].vol. 2. Moscow: Izdatel'stvo 'Nauka'.

SIMČENKO, JU.B. 1980. Rannie ėtapy ėtnogeneza narodov ural'skoj jazykovoj sem'i Zapoljar'ja i Pripoljar'ja Evrazii. [The early stages of the ethnogenesis of the Uralic peoples of Northern Eurasia]. In: Il'ja Samojlovič Gurvič (ed), 11-27.

SIMONCSICS, PÉTER. 1998. Kamassian. In: Daniel Abondolo (ed), 580-601.

ŠIROBOKOVA, NATAL'JA NIKOLAEVNA. 1977. Ob otnošenii jakutskogo jazyka k jazykam drevnetjurkskix pamjatnikov. [On the relationship of Yakut to the language of the Old Turkic remains]. *Issledovanija po jazykam narodov Sibiri - sbornik naučnyx trudov*, ed. by Elizaveta Ivanovna Ubrjatova, 108-16. Novosibirsk: Sibirskoe Otdelenie AN SSSR - Institut Istorii, Filologii i Filosofii.

—. 1980. O jakutsko-mongol'skix kontaktax. [On the Yakut-Mongolic contacts]. *Narody i Jazyki Sibiri*, ed. by Elizaveta Ivanovna Ubrjatova, 140-47. Novosibirsk: Izdatel'stvo Nauka, Sibirskoe otdelenie.

—. 2005. *Otnošenie jakutskogo jazyka k tjurkskim jazykam južnoj Sibiri*. [The relationship of the Yakut language to the Turkic languages of Southern Siberia]. Novosibirsk: 'Nauka'.

SKORIK, PETR JAKOVLEVIČ. 1961. *Grammatika čukotskogo jazyka. Čast' pervaja: fonetika i morfologija immenyx častej reči*. [A grammar of Chukchi. Part 1: phonetics and morphology of the nominal parts of speech]. Moscow, Leningrad: Izdatel'stvo akademii nauk SSSR.

SKRIBNIK, ELENA KONSTANTINOVNA. 1988. *Polipredikativnye sintetičeskie predloženija v burjatskom jazyke*. [Polypredicative synthetic sentences in Buryat]. Novosibirsk: Izdatel'stvo 'Nauka', Sibirskoe otdelenie.

—. 2001. Variation of noun phrase markers in Siberian languages. In: Winfried Boeder and Gerd Hentschel (eds), 345-64.

—. 2003. Buryat. In: Juha Janhunen (ed), 102-28.

—. 2004. Sprachen Sibiriens: Sprachkontakte und Sprachgeschichte, Sprachgeburt und Sprachtod. *Sprachtod und Sprachgeburt*, ed. by Peter Schrijver and Peter-Arnold Mumm, 141-70. Bremen: Hempen Verlag.

SLEPCOV, P.A. (ed.) 1972. *Jakutsko-Russkij Slovar'*. [Yakut-Russian dictionary]. Moscow: Izdatel'stvo "Sovetskaja Ènciklopedija".

SOLOMINA, OLGA and ALVERSON, KEITH. 2004. High latitude Eurasian paleoenvironments: introduction and synthesis. *Paleogeography, Paleoclimatology, Paleoecology*, 209.1-18.

SOMFAI KARA, DÁVID. 2002. *Kazak*. Munich: LINCOM EUROPA.

—. 2003. *Kyrgyz*. Munich: LINCOM EUROPA.

SOTAVALTA, ARVO and HALÉN, HARRY. 1978. *Westlamutische Materialien*. Mémoires de la Société Finno-Ougrienne, vol. 168. Helsinki: Suomalais-Ugrilainen Seura.

Sovremennyj kazaxskij jazyk. Fonetika i Morfologija. [The modern Kazakh language. Phonetics and Morphology.]. 1962. Alma-Ata: Izdatel'stvo Akademii Nauk Kazaxskoj SSR.

SPULER, BERTOLD. 1966. Geschichte Mittelasiens seit dem Auftreten der Türken. *Handbuch der Orientalistik. Erste Abteilung, Fünfter Band, Fünfter Abschnitt: Geschichte Mittelasiens*, ed. by Bertold Spuler, 123-310. Leiden, Köln: E.J. Brill.

STACHOWSKI, MAREK. 1998. An example of Nganasan-Dolgan linguistic contact. *Turkic Languages*, 2.126-29.

STACHOWSKI, MAREK and MENZ, ASTRID. 1998. Yakut. In: Lars Johanson and Eva Csató (eds), 417-33.

STASSEN, LEON. 1985. *Comparison and Universal Grammar*. Oxford: Basil Blackwell (cited from Heine 1997).

—. 2005. Comparative constructions. In: Martin Haspelmath, Matthew Dryer, David Gil and Bernard Comrie (eds), 490-93.

STATISTICA for Windows (Volume III): *Statistics II,* 2nd Edition. 1995. Tulsa: StatSoft, Inc.

STEINITZ, WOLFGANG. [1937] 1980. Xantyjskij (ostjackij) jazyk. [The Khanty (Ostyak) language]. *Wolfgang Steinitz. Ostjakologische Arbeiten Band IV*, 5-71. Budapest: Akadémiai Kiadó.

STOLZ, CHRISTEL and STOLZ, THOMAS. 1996. Funktionswortentlehnung in Mesoamerika. Spanisch-amerindischer Sprachkontakt (Hispanoindiana II). *Sprachtypologie und Universalienforschung*, 49.86-123.

STOLZ, THOMAS. 1996. Some instruments are really good companions - some are not. On syncretism and the typology of instrumentals and comitatives. *Theoretical Linguistics*, 23.113-200.

STOLZ, THOMAS and STOH, CORNELIA. 2001. Wenn Komitative Instrumentale sind und umgekehrt. In: Winfried Boeder and Gerd Hentschel (eds), 387-411.

STOLZ, THOMAS, STROH, CORNELIA and URDZE, AINA. 2005. Comitatives and instrumentals. In: Martin Haspelmath, Matthew Dryer, David Gil and Bernard Comrie (eds), 214-17.

STONE, A. C. and STONEKING, M. 1998. mtDNA analysis of a prehistoric Oneota population: implications for the peopling of the New World. *American Journal of Human Genetics*, 62.1153-70.

SUNIK, OREST PETROVIČ. 1948. O possessivnyx affiksax i roditel'nom padeže v tunguso-man'čžurskix jazykax. [On the possessive affixes and the genitive case in the Tungusic languages.]. *Jazyk i myšlenie*, XI.283-91.

—. 1962. *Glagol v tunguso-man'čžurskix jazykax*. [The verb in the Tungus-Manchu languages]. Moscow, Leningrad: Izdatel'stvo Akademii nauk SSSR.

—. 1968. Tunguso-man'čžurskie jazyki (vvedenie). [Tungus-Manchu languages (introduction)]. *Jazyki narodov SSSR: Mongol'skie, tunguso-man'čžurskie i paleoaziatskie jazyki*, 53-67. Leningrad: Izdatel'stvo ‚Nauka' Lenigradskoe otdelenie.

—. 1985. *Ul'chskij jazyk. Issledovanija i materialy*. [Ulča. Study and materials]. Leningrad: Nauka.

SVANTESSON, JAN-OLOF. 2003. Khalkha. In: Juha Janhunen (ed), 154-78.

TEKIN, TALÂT. 1976. The representation of Proto-Turkic medial and final /s/ in Yakut. *Central Asiatic Journal*, XX.110-14.

TENIŠEV, E.R. (ed.) 2001. *Sravnitel'no-istoričeskaja grammatika tjurkskix jazykov. Leksika*. [A comparative-historical grammar of the Turkic languages. Lexicon]. Moscow: 'Nauka'.

—. (ed.) 2002. *Sravnitel'no-istoričeskaja grammatika tjurkskix jazykov: regional'nye rekonstrukcii*. [A comparative-historical grammar of the Turkic languages: regional reconstructions]. Moscow: 'Nauka'.

TEREŠČENKO, NATAL'JA M. 1973. *Sintaksis samodijskix jazykov. Prostoe predloženie*. [The syntax of the Samoyedic languages. Simple clauses]. Leningrad: Izdatel'stvo 'Nauka', Leningradskoe otdelenie.

—. 1979. *Nganasanskij jazyk*. [The Nganasan language.]. Leningrad: Izdatel'stvo "Nauka", leningradskoe otdelenie.

TEREŠKIN, NIKOLAJ I. 1961. *Očerki dialektov xantyjskogo jazyka*. [Sketches of Khanty dialects.]. Moscow, Leningrad: Izdatel'stvo Akademii Nauk SSSR.

THOMASON, SARAH GREY. 2001. *Language Contact. An Introduction*. Edinburgh: Edinburgh University Press.

—. 2003. Contact as a source of language change. *The Handbook of Historical Linguistics*, ed. by Brian D. Joseph and Richard D. Janda, 687-712. Malden, Mass.: Blackwell Publishers. (Reprinted as a paperback in 2005).

THOMASON, SARAH GREY and KAUFMAN, TERRENCE. 1991. *Language Contact, Creolization, and Genetic Linguistics. Berkeley*, Los Angeles, Oxford: University of California Press. (Originally published in 1988).

THURSTON, WILLIAM R. 1987. *Processes of Change in the Languages of North-Western New Britain*. Pacific Linguistics Series B, vol. 99. Canberra: Department of Linguistics, Research School of Pacific Studies, The Australian National University.

TIMBERLAKE, ALAN. 1975. Hierarchies in the genitive of negation. *The Slavic and East European Journal. Special Issue: Soviet-American Russian Language Contributions.*, 19.123-38.

TODAEVA, BULJAŠ XOIČIEVNA. 1986. *Dagurskij jazyk*. [The Dagur language]. Moscow: Izdatel'stvo 'Nauka'.

TORRONI, A., SCHURR, T. G., CABELL, M. F., BROWN, M. D., NEEL, J. V., LARSEN, M., SMITH, D. G., VULLO, C. M. and WALLACE, D. C. 1993. Asian affinities and continental radiation of the four founding Native American mtDNAs. *American Journal of Human Genetics*, 53.563-90.

TRAUZETTEL, ROLF. 1986. Die Yüan-Dynastie. In: Michael Weiers (ed), 217-82.

TRUDGILL, P. 1986. *Dialects in Contact*. Oxford: Blackwell (cited from Ross 1997).

TSUMAGARI, TOSHIRO. 1997. Linguistic Diversity and National Borders of Tungusic. *Northern Minority Languages. Problems of Survival. Papers Presented at the Eighteenth Taniguchi International Symposium: Division of Ethnology*, ed. by Hiroshi Shoji and Juha Janhunen, 175-86. Osaka: Osaka: National Museum of Ethnology.

—. 2003. Dagur. In: Juha Janhunen (ed), 129-53.

TUGOLUKOV, VLADILEN ALEKSANDROVIČ. 1980. Ètničeskie korni tungusov. [The ethnic roots of the Tungus.]. In: Il'ja Samojlovič Gurvič (ed), 152-76.

—. 1985. *Tungusy (ėvenki i ėveny) Srednej i Zapadnoj Sibiri*. [The Tungus (Evenks and Evens) of Central and Western Siberia]. Moscow: Izdatel'stvo 'Nauka'.

TUZUN, E., SHARP, A. J., BAILEY, J. A., KAUL, R., MORRISON, V. A., PERTZ, L. M., HAUGEN, E., HAYDEN, H., ALBERTSON, D., PINKEL, D., OLSON, M. V. and EICHLER, E. E. 2005. Fine-scale structural variation of the human genome. *Nature Genetics*, 37.727-32.

UBRJATOVA, ELIZAVETA IVANOVNA. 1956. Vzaimodejstvie jazykov na materiale vzaimootnošenij jakutskogo i ėvenkijskogo jazykov. [The reciprocal influence of languages exemplified by Yakut and Evenki]. *Doklady i soobščenija instituta jazykoznanija AN SSSR*, 9.85-92.

—. 1960. *Opyt sravnitel'nogo izučenija fonetičeskix osobennostej jazyka naselenija nekotoryx rajonov Jakutskoj ASSR*. [An attempt at a comparative study of the phonetic peculiarities of the inhabitants of some districts of the Yakut ASSR]. Moscow: Izdatel'stvo Akademii nauk SSSR.

—. 1966. O jazyke dolgan. [On the Dolgan language]. J*azyki i fol'klor narodov sibirskogo Severa*, ed. by Valentin Aleksandrovič Avrorin, 41-68. Moscow, Leningrad.

—. 1972. Kratkij grammatičeskij očerk jakutskogo jazyka. [A short grammar sketch of Yakut]. *Jakutsko-Russkij Slovar'*, ed. by P.A. Slepcov, 569-606. Moscow: Izdatel'stvo "Sovetskaja Ènciklopedija".

—. 1976. *Issledovanija po sintaksisu jakutskogo jazyka II: Složnoe predloženie, kniga pervaja*. [A study of Yakut syntax II: complex clauses, first book.]. Novosibirsk: Izdatel'stvo "Nauka", Sibirskoe otdelenie.

—. 1985a. *Istoričeskaja grammatika jakutskogo jazyka*. [A historical grammar of the Yakut language]. Yakutsk: Jakutskij gosudarstvennyj universitet.

—. 1985b. *Jazyk noril'skix dolgan*. [The language of the Norilsk Dolgans.]. Novosibirsk: Izdatel'stvo 'Nauka', Sibirskoe Otdelenie.

VAJDA, EDWARD J. 2004. *Ket*. LINCOM Europa.

VAKHTIN, NIKOLAI. 1992. The Yukagir Language in Sociolinguistic Perspective. *Linguistics,* 1.47-82.

VAN COETSEM, FRANS. 1988. *Loan Phonology and the Two Transfer Types in Language Contact*. Dordrecht et al.: Foris Publications (cited from Winford 2005).

—. 2000. *A General and Unified Theory of the Transmission Process in Language Contact*. Monographien zur Sprachwissenschaft. Heidelberg: Winter.

VAN DER AUWERA, JOHAN, DOBRUSHINA, NINA and GOUSSEV, VALENTIN. 2005. Imperative-Hortative systems. In: Martin Haspelmath, Matthew Dryer, David Gil and Bernard Comrie (eds), 294-97.

VASILEVIČ, GLAFIRA MAKAR'EVNA. 1969. *Èvenki. Istoriko-ètnografičeskie očerki (XVIII-načalo XX v)*. [The Evenks. Historical-ethnographical sketches (17th to the beginning of the 20th century)]. Leningrad: Izdatel'stvo ‚Nauka' Leningradskoe otdelenie.

VAXTIN, NIKOLAJ BORISOVIČ. 2001a. *Jazyki narodov severa v XX veke. Očerki jazykovogo sdviga*. [The languages of the peoples of the north in the 20th century. Sketches of language shift]. St Petersburg: Evropejskij Universitet v Sankt-Peterburge.

—. 2001b. Imperative sentences in Asiatic Eskimo. In: Viktor Xrakovskij (ed), 129-44.

VIETZE, HANS-PETER. 1988. *Lehrbuch der mongolischen Sprache*. Leipzig: VEB Verlag Enyzklopädie.

VOLODIN, ALEXANDER PAVLOVIČ. 1976. *Itel'menskij jazyk*. Leningrad: Izdatel'stvo 'Nauka', Leningradskoe otdelenie.

VORONKIN, MIXAIL SPIRIDONOVIČ. 1999. *Dialektnaja sistema jazyka Saxa*. [The dialect system of the Saxa language]. Novosibirsk: 'Nauka'. Sibirskaja izdatel'skaja firma RAN.

WEIERS, MICHAEL. 1977. Typen verbaler Personendifferenzierung im Mongolischen. *Central Asiatic Journal,* 21.295-325.

WEIERS, MICHAEL (ed). 1986. *Die Mongolen. Beiträge zu ihrer Geschichte und Kultur*. Darmstadt: Wissenschaftliche Buchgesellschaft.

—. 1986a. Das Khanat Tschaghatai. In: Michael Weiers (ed), 290-99.

—. 1986b. Die goldene Horde oder Das Khanat Qyptschaq. In: Michael Weiers (ed), 345-78.

—. 1986c. Die Mongolen im Iran. In: Michael Weiers (ed), 300-44.

—. 1986d. Zur Herausbildung und Entwicklung monoglischer Sprachen. In: Michael Weiers (ed), 29-69.

—. 1986e. Von Ögödei bis Möngke – Das mongolische Großreich. In: Michael Weiers (ed), 192-216.

—. 2003. Moghol. In: Juha Janhunen (ed), 248-64.

WEINREICH, URIEL. 1953. *Languages in Contact - Findings and Problems*. New York: Publications of the Linguistic Circle of New York - Number 1.

WERNER, HEINRICH. 1997. *Die ketische Sprache*. Wiesbaden: Harrassowitz Verlag.

—. 2005. *Die Jennisej-Sprachen des 18. Jahrhunderts*. Veroeffentlichungen der Societas Uralo-Altaica, vol. 67. Wiesbaden: Harrassowitz Verlag.

WHALEY, LINDSAY J., GRENOBLE, LENORE A., LI, FENGXIANG. 1999. Revisiting Tungusic classification from the bottom up: a comparison of Evenki and Oroqen. *Language*, 75/2.286-321.

WINFORD, DONALD. 2005. Contact-induced changes. Classification and processes. *Diachronica*, 22:2.373-427.

WITSEN, NICOLAES. [1692] 1705. *Noord en Oost Tartarye, ofte Bondig Ontwerp Van eenige dier Landen en Volken, welke voormaels bekent zijn geweest. Bemessens verscheide tot noch toe onbekende, en meest nooit voorheen beschreve Tartersche en Nabuurige Gewesten, Landstreken, Steden, Rivieren, en Plaetzen, in de Noorder en Oosterlykse gedeelten van Asia en Europa, Zoo buiten en binnen de Rivieren Tanais en Ob, als omtrent de Kaspische, Indische-Ooster, en Zwarte Zee gelegen; gelijk de Landschappen Niuche, Dauria, Jesso, Moegalia, Kalmakkia, Tangut, Usbek, en Noorde-Persie, Turkestan, Georgia, Mengrelia, Cirkassia, Crim, Astakkia, Altin, Tingoesia, Siberia, Samojedia, en andere aen hunne Tzaersche Majesteiten Kroon gehoorende Heerschappyen: Verdeeld in twee Stukken, Met der Zelver Land-Karten: mitsgaders, onderscheidene Afbeeldingen van Steden, Drachten, enz. Zedert naeuwkeurig onderzoek van veele Jaren, en eigen ondervindinge ontworpen, beschreven, geteekent, en in't licht gegeven door Nicolaes Witsen.* Tweede Druk. Amsterdam: François Halma, Boekverkooper op de Nieuwen-dyk.

WOODING, S., OSTLER, C., PRASAD, B. V., WATKINS, W. S., SUNG, S., BAMSHAD, M. and JORDE, L. B. 2004. Directional migration in the Hindu castes: inferences from mitochondrial, autosomal and Y-chromosomal data. *Human Genetics*, 115.221-9.

WURM, STEPHEN A. 1996a. Siberia: 1650-1950 ethnic and linguistic changes. In: Stephen A. Wurm, Peter Mühlhäusler and Darrell T. Tryon (eds), 969-74.

—. 1996b. Indigenous lingue franche and bilingualism in Siberia (beginning of the 20th century). In: Stephen A. Wurm, Peter Mühlhäusler and Darrell T. Tryon (eds), 975-78.

WURM, STEPHEN A., MÜHLHÄUSLER, PETER, and TRYON, DARRELL T. (eds). 1996. *Atlas of Languages of Intercultural Communication in the Pacific, Asia, and the Americas*. Berlin, New York: Mouton de Gruyter.

XARITONOV, L.N. 1947. *Sovremennyj jakutskij jazyk. Čast' pervaja: fonetika i morfologija*. [The modern Yakut language. Part I: phonetics and morphology]. Yakutsk: Naučno-Issledovatel'skij Institut jazyka, literatury i istorii JaASSR.

XASANOVA, MARINA and PEVNOV, ALESANDR. 2003. *Mify i skazki negidal'cev*. [Myths and tales of the Negidals]. Kyoto: Nakanishi.

XRAKOVSKIJ, VIKTOR (ed). 2001. *Typology of Imperative Constructions*. Munich: Lincom Europa.

Y_CHROMOSOME_CONSORTIUM. 2002. A nomenclature system for the tree of human Y-chromosomal binary haplogroups. *Genome Research*, 12.339-48.

ZERJAL, T., WELLS, R. S., YULDASHEVA, N., RUZIBAKIEV, R. and TYLER-SMITH, C. 2002. A genetic landscape reshaped by recent events: Y-chromosomal insights into central Asia. *American Journal of Human Genetics*, 71.466-82.

ZERJAL, T., DASHNYAM, B., PANDYA, A., KAYSER, M., ROEWER, L., SANTOS, F. R., SCHIEFENHÖVEL, W., FRETWELL, N., JOBLING, M. A., HARIHARA, S., SHIMIZU, K., SEMJIDMAA, D., SAJANTILA, A., SALO, P., CRAWFORD, M. H., GINTER, E. K., EVGRAFOV, O. V. and TYLER-SMITH, C. 1997. Genetic relationships of Asians and Northern Europeans, revealed by Y- chromosomal DNA analysis. *American Journal of Human Genetics*, 60.1174-83.

370

SUMMARY

This study analyses the prehistory of a northeastern Siberian population, the Sakha, from both a molecular-genetic and a linguistic perspective. The Sakha, who are a Turkic-speaking group of cattle- and horse-breeders, migrated to the Lena river from an area further to the south several hundred years ago. This migration brought the ancestors of the Sakha into contact with populations speaking different languages and with different subsistence patterns. The aim of this investigation is to elucidate the extent to which the Sakha interacted with the indigenous populations of the territory that they migrated to, both from the point of view of genetic admixture and from the point of view of language contact.

The results of this study show that the Sakha were in contact with two different groups during their history: with speakers of a Mongolic language and with speakers of Evenki. The contact with the Mongolic-speaking group can be shown to have taken place during the period of the Mongol Empire, when the Mongols ruled over large tracts of Eurasia. During this time, the Sakha copied a large number of lexical items from Mongolic, possibly due to the social and political prestige of the Mongolic-speaking group.

In contrast, the contact with the Evenks led to the introduction of a number of schematic copies, but only a relatively small amount of substance copies from Evenki into Sakha. The schematic copies from Evenki are the loss of the Genitive case, the use of the Partitive case to mark indefinite direct objects, the retention of the distinction between Comitative and Instrumental, the development of a Distant Future Imperative, and the extension in use of the possessive suffixes to non-possessive functions. The nature of the copies from Evenki implies that the Sakha were dominantly bilingual in Evenki. The genetic results, however, show no evidence of male admixture from Evenks, and little evidence for female admixture, although this cannot be excluded. Thus, the genetic analyses indicate that there was no shift of entire Evenk communities to the Sakha language and identity. The schematic copies from Evenki found in Sakha may provide evidence of female-biased intermarriage with Evenks; although there is no conclusive evidence for this in the genetic results, this cannot be excluded on the basis of the mtDNA analyses alone. Possibly, however, the schematic copies entered the language through frequent social interaction of Evenks and Sakha during the initial period after the Sakha migrated to the Lena. As shown by the genetic analyses, the immigrating Sakha ancestors consisted of a very small group of men; these may initially have been dependent on communication with the indigenous Evenks, until they had fully adapted to the new environment.

SAMENVATTING IN HET NEDERLANDS

Dit proefschrift onderzoekt de prehistorie van de Sakha (een bevolkingsgroep in noordoost Siberië) vanuit zowel een moleculair-genetisch als een linguïstisch perspectief. De Sakha, een Turkssprekende groep koeien- en paardenhouders, zijn een aantal eeuwen geleden naar de Lena getrokken vanuit een meer naar het zuiden gelegen gebied. Deze migratie bracht de voorouders van de huidige Sakha in contact met bevolkingsgroepen die andere talen spraken en andere levensgewoonten hadden. Het doel van dit onderzoek is inzicht te krijgen in de mate waarin de Sakha in contact traden met the oorspronkelijke bevolking van het gebied waar zij naartoe trokken, zowel vanuit het gezichtspunt van genetische vermenging alsook vanuit het perspectief van taalcontact.

Dit onderzoek toont aan dat de Sakha in de loop van de geschiedenis contact hebben gehad met twee verschillende groepen: enerzijds met sprekers van een Mongoolse taal, en anderzijds met sprekers van het Evenki. Het contact met een Mongools sprekende bevolking kan gedateerd worden in de periode van het Mongoolse Rijk. De Mongolen heersten toen over grote delen van Eurazië. In deze periode ontleenden de Sakha een grote hoeveelheid woorden uit het Mongools, waarschijnlijk vanwege het sociale en politieke prestige van de Mongools sprekende populatie.

In tegenstelling daarmee veroorzaakte het contact met de Evenki wel een aantal structurele ontleningen, maar slechts een relatief kleine hoeveelheid lexicale ontleningen vanuit het Evenki in het Sakha. De structurele ontleningen vanuit het Evenki zijn het verlies van de genitief, het gebruik van de partitieve naamval voor de markering van een onbepaald lijdend voorwerp, het behoud van een onderscheid tussen een comitatief en een instrumentalis, de ontwikkeling van een imperatief voor de verre toekomst, en de uitbreiding van het gebruik van possessieve suffixen naar niet-possessieve functies. De aard van ontleningen vanuit het Evenki suggereren dat de Sakha in meerderheid tweetalig waren in Evenki.

De resultaten van het genetische onderzoek echter tonen geen evidentie voor vermenging met mannelijke Evenken, en maar weinig evidentie voor vrouwelijke vermenging, alhoewel dat laatste niet uitgesloten kan worden. Deze genetische resultaten geven aan dat er hoogstwaarschijnlijk geen complete groepen Evenken de taal en identiteit van de Sakha hebben aangenomen. De structurele ontleningen vanuit het Evenki in het Sakha echter kunnen wijzen op gemengde huwelijken met vrouwelijke Evenken. Alhoewel er geen definitief bewijs hiervoor gevonden kan worden in de genetische resultaten, kan het ook niet worden uitgesloten op basis van de uitgevoerde analyses van mtDNA.

De structurele ontleningen zijn waarschijnlijk in het Sakha terecht gekomen door regelmatig contact tussen de Evenken en de Sakha in de beginperiode vlak nadat the Sakha naar de Lena waren getrokken. De genetische analyses laten zien dat de immigrerende voorouders van de huidige Sakha bestonden uit een kleine groep mannen. Deze kleine groep was waarschijnlijk afhankelijk van de communicatie met de lokale Evenken totdat ze geheel aangepast waren aan hun nieuwe omgeving.

CURRICULUM VITAE

Brigitte Pakendorf was born on March 6th, 1970 in Johannesburg, South Africa. She finished her Master's degree in Biological Anthropology at the University of Hamburg in 1996, and obtained her Ph.D. degree from the Faculty of Biology at the University of Hamburg in 2001. In 2001 she began the research for her Linguistics Ph.D. thesis at the Department of Linguistics and the Department of Evolutionary Genetics at the Max Planck Institute of Evolutionary Anthropology in Leipzig, where she currently holds the position of leader of the independent junior scientists group on 'Comparative Population Linguistics'.