



HAL
open science

Sakha and Dolgan, the North Siberian Turkic languages

Brigitte Pakendorf, Eugenie Stapert

► **To cite this version:**

Brigitte Pakendorf, Eugenie Stapert. Sakha and Dolgan, the North Siberian Turkic languages. Robbeets, Martine & Alexander Savelyev. The Oxford Guide to the Transeurasian Languages, Oxford University Press, pp.430-445, 2020, 10.1093/oso/9780198804628.003.0027 . hal-02889684

HAL Id: hal-02889684

<https://hal.univ-lyon2.fr/hal-02889684v1>

Submitted on 4 Jul 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Chapter 26

Sakha and Dolgan, the North Siberian Turkic languages

Brigitte Pakendorf and Eugenie Stapert

Abstract

This chapter provides a brief structural overview of the North Siberian Turkic languages Sakha (also known as Yakut) and Dolgan. Both languages are spoken in the northeast of the Russian Federation: Sakha in the Republic Sakha (Yakutia) and Dolgan on the Taimyr Peninsula. These languages clearly fit the Turkic linguistic profile with vowel harmony, agglutinative morphology, SOV word order, and preposed relative clauses, but due to contact-induced changes there are considerable differences from other Turkic languages as well. Notable differences are the loss of the Turkic genitive and locative case and the development of a partitive and comparative case, as well as a distinction between an immediate and a remote imperative. Like other so-called Altaic languages, Sakha and Dolgan make widespread use of non-finite verb forms in subordination.

Keywords: Turkic, Sakha, Yakut, Dolgan, Taimyr, vowel harmony, partitive case, comparative case, remote imperative, contact-induced change

26.1 Introduction

Sakha—which is the endonym of the people and their language—is more commonly known in the literature as Yakut. It is the most numerous and sociopolitically dominant indigenous language in the Republic Sakha (Yakutia), and its geographical spread covers large parts of central and northeastern Siberia. Dolgan, on the other hand, is largely restricted to the southern edge of the Taimyr Peninsula and adjacent

northwestern Yakutia. On the Taimyr Peninsula Dolgans are settled primarily along the Kheta and Khatanga rivers, which flows northeast into the Laptev Sea.

<Insert Figure 26.1 here>

Figure 26.1 Map of eastern Siberia showing the areas where Sakha and Dolgan are spoken. © DDL

The 2010 Russian national census (Nacional'nyj sostav 2010: 12, 21) counted 478,085 Sakha and 7,885 Dolgans. These are still predominantly rural, though the Sakha are more urbanized: ~60% (284,834) of Sakha and ~77% (6,045) of Dolgans lived in rural settlements at last count. Among the Sakha, 406,126 (i.e. approximately 85%) claimed to speak Sakha in response to the census, with 428,280 (~90%) claiming a knowledge of Russian (Vladienie jazykami naseleniem 2010: 158, 152). Although there are thus still substantial numbers of Sakha speakers, there is a trend towards Russian dominance: in the 2002 census 93% of Sakha respondents claimed to know their heritage language and 87% claimed a knowledge of Russian. Nevertheless, in rural settlements Sakha is still quite healthy: here, it is the language of everyday use, and some older speakers and most pre-school children are monolingual (Pakendorf 2007: 2). Dolgan, in contrast, is clearly endangered: only 930 Dolgans (i.e. less than 12%) claimed to speak Dolgan in 2010 (Vladienie jazykami korennyx 2010: 2120). However, the sociolinguistic situation differs considerably from village to village. While Dolgan is still acquired by children in the remote villages of Syndassko and Popigaj, it is spoken fluently only by a few individuals over 60 years old in villages closer to the city of Dudinka (Stapert 2013: 29).

Both Sakha and Dolgan are written with a Cyrillic-based orthography with some additional graphemes for specific phones/phonemes. Sakha has been the object of extensive study, starting in the 19th century with the first reference grammar (Böhtlingk 1851) and a multi-volume dictionary (Pekarskij 1907–1930). Further landmarks are the two-volume reference grammar of the language (Ubrjatova et al. 1982; Ubrjatova et al. 1995) and the ongoing publication of the monumental Sakha explanatory dictionary (Slepcov 2004–2015). Among the key sources concerning Sakha language contact are Kałużyński (1962) on the Mongolian elements in Sakha and Romanova et al. (1975) on Sakha-Evenki contacts. Dolgan has been less widely studied; the most important grammatical descriptions of this language are Ubrjatova (1985) and Artem'ev (2001a, b), while Stachowski (1993, 1998) represents a rare lexicographical source.

This chapter¹ is largely based on our oral corpora of Sakha and Dolgan² (see Pakendorf 2007: 63–64 and Stapert 2013: 7–9 for details). These number approximately 29,400 and 17,700 words, respectively.

26.2 Historical connections: genealogy and contact

Both authors are sceptical with respect to the genealogical relationship of the Transeurasian languages. Nevertheless, if one considers Transeurasian or Altaic languages to be an areal grouping rather than a language family, then Sakha and Dolgan belong to this grouping.

Sakha and Dolgan are classified among the Northeastern Turkic languages, a heterogeneous group that is based on areal contiguity rather than on established genealogical unity (Johanson 1998a: 82–83). Within this grouping Johanson (1998a: 83) identifies the homogenous North Siberian group comprising Sakha and Dolgan,

and a heterogeneous South Siberian group. The North Siberian Turkic languages differ from their relatives in phonological and several structural features, as well as numerous lexical changes (Kałużyński 1962; Pakendorf 2007; Pakendorf and Novgorodov 2009a; Stapert 2013). Mutual intelligibility with other Turkic languages is thus severely restricted.

The two languages are so closely related that Dolgan is sometimes considered a dialect of Sakha (e.g. Voronkin 1999: 154); however, on the basis of sociocultural and political factors it can be considered a separate language (Artem'ev 2001a: 6; Stapert 2013: 62). Although Sakha and Dolgan are spoken over a vast geographical area, they are very homogenous, probably due to their recent spread, which is assumed to have taken place only in the late 17th and early 18th centuries (Dolgix 1960: 360; Wurm 1996: 971–972; Pakendorf et al. 2006: 348). Nevertheless, four dialect groups are distinguished for Sakha, based partly on geographic and sociocultural criteria (Voronkin 1999: 155): the central group, the Vilyuy group, the northwestern group (which in Voronkin's classification includes Dolgan), and the northeastern group (Voronkin 1999: 154–155). In Dolgan, an 'upriver' dialect (*üöhegi*) is distinguished from a "downriver" dialect (*allaragi*) (Stapert 2013: 27).

Sakha has undergone substantial contact influence in its history, both lexical and structural. Since Dolgan is derived from Sakha (Stapert 2013: 34–62), it shares some of these contact-induced changes, but has undergone additional changes under the influence of Evenki. During its history, Sakha has copied large amounts of lexical items from an unspecified Mongolic language. These include even basic lexical items such as kinship and body part terms, a large number of verbs (Pakendorf and Novgorodov 2009a: 509), as well as many derivational suffixes (Kałużyński 1962; Pakendorf 2015). In addition, Sakha has undergone structural changes under the

influence of Evenki (Pakendorf 2007: 95–270). Interestingly, as shown by molecular anthropological investigations (Pakendorf et al. 2006; Pugach et al. 2016: S3A), these contact-induced changes in Sakha took place largely in the absence of intimate physical contact (cf. Pakendorf 2007: 317–323).

Dolgan has copied some lexical items and has also undergone structural influence, including calques, from Evenki since its divergence from Sakha, and has also undergone Russian influence (Stapert 2013: 330–357). In contrast to the Sakha contact situation, the contact-induced changes in Dolgan went hand in hand with genetic admixture with both Evenks and Russians (Pugach 2016: S3A; cf. Stapert 2013: 69–76).

26.3 Phonology

26.3.1 Consonants

Both Sakha and Dolgan have 16 consonantal phonemes³ and three allophones (in brackets) shown in Table 26.1.

Table 26.1 Consonant phonemes

	Labial	Alveodent	Palatal	Velar	Uvular	Glottal
Plosive	p b	t d		k g		
Nasal	m	n	ɲ	ŋ		
Trill		r				
Fricative		s		(ʏ)	(χ)	(h)
Affricate			tʃ dʒ			
Approximant			j			
Lateral approx.		l				

In both languages there is variation between [dʒ], [dʲ], and [j], e.g. [dʒaχtar], [dʲaχtar], and [jaχtar] ‘woman’, with the fricative pronunciation being more frequent in Sakha and the palatalized stop more frequent in Dolgan. Similarly, [j], [jʲ], and [ɲ] occur interchangeably in a restricted number of words, e.g. [ajaχ] ~ [aʲjaχ] ~ [aɲaχ] ‘mouth’, [ije] ~ [iʲje] ~ [iɲe] ‘mother’. While in Sakha the glide is standardised as <й>, in Dolgan the palatal nasal fulfils this role, represented as <НЬ> in the orthography.

Moreover, restrictions apply to the distribution of some phonemes: word-initially /r/ does not occur at all, /g/ occurs in only a very limited number of native words (only *gin* ‘do’ in Sakha, and a few more in Dolgan, e.g. *gini* ‘he’ and *gitta* ‘with’), and /ŋ/ occurs word-initially only in loanwords. Word-initial /ɲ/ is restricted to mostly onomatopoeic words of Turkic origin, whereas it is fairly common in words of Tungusic and Samoyedic origin, e.g. *noχffo* ‘hunchback’. The voiced consonants /b/, /d/, and /g/ do not occur in word-final position.

Prevocally [χ] occurs before the low back vowels /a/ and /o/ in both languages, whereas [k] is found before all other vowels, e.g. *χa:r* ‘snow’ vs. *kel* ‘come’. Post-vocally [χ] occurs after all low vowels in Sakha, whereas in Dolgan it is restricted to low back vowels. While this allophonic variation is represented graphemically in Sakha by <κ> for [k] and <x> for [χ], in the Dolgan orthography only <κ> is used. [s] occurs consistently in syllable-final position and before consonants, e.g. *mas* ‘wood’, *bus-put* ripen-PST.PTCP ‘ripened’. It changes to [h] in intervocalic position, e.g. *mas* > *mah-a* wood-POSS.3SG ‘his wood’. In word-initial position, [s] and [h] are possible in both spoken Dolgan and Sakha, but in writing, Sakha has standardised [s] with the grapheme <c>, whereas Dolgan uses [h] (<h>). In Sakha, [ɣ] occurs between all low vowels, as in *aya* ‘father’, *beyehe* ‘yesterday’, and is represented graphemically by

<ɣ>. In spoken Dolgan, [ɣ] occurs only after low back vowels, i.e. *aya*, but otherwise [g] is pronounced, e.g. *begehe*. Both [g] and [ɣ] are represented by the letter <r>.

Consonant assimilation (progressive, regressive, and bidirectional) takes place between adjacent consonants on morpheme boundaries, e.g. *at* ‘horse’ + *-LAR*⁴ ‘PL’ > *attar* ‘horses’. For a full overview see Stachowski and Menz (1998: 419–420). /p/, /k/ and /χ/ voice before vowels, whereas /t/ voices before vowels only when it is adjacent to /r/ or /l/.

26.3.2 Vowels

Sakha and Dolgan have 20 vowel phonemes: eight short vowels, eight long vowels, and four diphthongs. They can be categorised according to height, roundedness, and frontness (Table 26.2). Diphthongs behave phonologically according to their leading element and are therefore classified as high vowels. For all features minimal pairs are attested, e.g. *at* ‘horse’ *a:t* ‘name’, *it* ‘dog’ *i:t* ‘send!’, *ot* ‘grass’, *uot* ‘fire’.

Table 26.2 Vowel phonemes

	Low		High	
	Unrounded	Rounded	Unrounded	Rounded
Back	a a:	o o:	i i: ia	u u: uo
Front	ɛ ɛ:	ø ø:	i i: ie	y y: yø

Vowels are combined according to the principles of palatal and labial vowel harmony. Palatal vowel harmony is widespread among Turkic languages (Johanson 1998b: 32–34), and applies within lexical stems (e.g. *kirsa* ‘polar fox’, *kiehe* ‘evening’) as well as

across morpheme boundaries (e.g. *kirsa-lar* polar fox-PL, *kiehe-ler* evening-PL). While in Sakha this principle applies very strictly, Dolgan allows for certain exceptions, e.g. *harsierda* ‘morning’ vs. Sakha *sarsiarda*.

Labial vowel harmony is less common in Turkic languages (Johanson 1998b: 32–34) and also applies within stems and across morpheme boundaries (*oyo-lor* child-PL vs. *kirsa-lar* polar.fox-PL). The only exception occurs after the high rounded vowels /u/ and /y/, after which low vowels become unrounded, e.g. *tyrgeŋ* ‘speed’, *omuk-tar* foreigner-PL. Similarly, *muos* ‘antler’ becomes *muos-tar* antler-PL and not **muos-tor*, since in diphthongs the initial high element is leading.

26.3.3 Phonotactics

In native words the following syllable structures occur: V, VC, VCC, V:C, CV, CV:, CV:C, CVC, CVCC. Consonant clusters are prohibited word-initially, but word-finally they can occur in the combinations [rt], [rk], [lk], [lt], [mp], [ŋk].

Some stems containing a high vowel in the second syllable lose the high vowel in inflectional forms⁵, e.g. *tayis-* ‘exit’ becomes *taχs-ar* [exit-PRS⁶[PRED.3SG]] ‘she exits’. This also occurs in the passive, where the suffix *-(I)lln* changes into *-(I)ll*, e.g. *tut-ull-u-but-a* [build-PASS-EP-PST2-POSS.3SG] ‘it was built’.

26.4 Morphology

26.4.1 Nouns

Nouns have singular and plural number, whereby singular is unmarked, and plural is marked with a suffix (but see Section 26.5.2 for lack of plural agreement after numerals). Most common is *-LAr*, as in *hirga* ‘sleigh’ > *hirga-lar* ‘sleighs’. Some words are pluralised using the alternative suffix *-ttAr*⁷, as in *eme:χsi-tter* ‘old

women' < *eme:χsin* 'old woman'. In Sakha, this suffix is restricted to stems ending in -n, which typically refer to humans, but not always, e.g. *ati-ttar* < *atin* 'other', said in reference to 'other districts'. In Dolgan on the other hand, the association with humanness is stronger than the morphological restriction of ending in -n, e.g. *d'aχta-ttar* 'women' < *d'aχtar* 'woman' (Sakha *dʒaχtal-lar*), and *duoχtu-ttar* 'doctors' < *duoχtur* 'doctor' (copied from Russian). In addition, there are a few irregular plurals on -*ttAr* in both languages, e.g. *kī:s* 'girl' > *kīrgittar* 'girls'. Finally, the associative plural is derived using the proprietive suffix -*LA:χ*, which attaches only to names, e.g. *Jekim-ne:χ kīra uol-lara* [Jekim-PROPR small son-POSS.3PL] 'the youngest son of Jekim and family'.

Dolgan and Sakha have eight cases (Ubrjatova et al. 1982: 130; Ubrjatova 1985: 122): nominative, accusative, dative, partitive, ablative, instrumental, comitative, and comparative, which occur in a possessive and a non-possessive declension (for a complete overview of suffixes see Table 26.3). Sakha and Dolgan deviate from the Turkic pattern in that they have lost the genitive and locative case and have developed a partitive and comparative as well as a comitative, which differs from the instrumental. However, in Dolgan the status of the comitative is questionable (cf. Ubrjatova 1985: 122), since it occurs only once in the spoken corpus (vs. 50 times for Sakha). A much more common way to express a comitative meaning in Dolgan is the use of the postposition *gitta* 'with' (which governs accusative case) or the proprietive suffix -*LA:χ*.

26.4.2 Pronouns and possessive suffixes

Possessive suffixes, which are identical for Sakha and Dolgan (Table 26.3), are used to mark possession on nouns, and they play an important role in definiteness,

reference tracking in relative clauses, as well as in the verbal paradigm. In oblique cases, possessive formants for the singular differ from those in the nominative.

Possessive formants and case suffixes combine into the possessive declension represented in Table 26.3.

Table 26.3 Case and possessive suffixes

	Non-	Possessive					
	Posses sive	1SG	2SG	3SG	1PL	2PL	3PL
NO M	—	<i>-(I)m</i>	<i>-(I)ŋ</i>	<i>-(t)A</i>	<i>-BIt</i>	<i>-GIt</i>	<i>-LArA</i>
DAT	<i>-GA</i>	<i>-BAr</i>	<i>-GAR</i>	<i>-(t)IgAr/ -Ar</i>	<i>-BItIgAr</i>	<i>-GItIgAr</i>	<i>-LArIgAr</i>
ACC	<i>-(n)I</i>	<i>-BIn</i>	<i>-GIn</i>	<i>-(t)In</i>	<i>-BItIn</i>	<i>-GItIn</i>	<i>-LArIn</i>
PAR T	<i>-TA</i>	<i>-BInA</i>	<i>-GInA</i>	<i>-(t)InA</i>	<i>-BItInA</i>	<i>-GItInA</i>	<i>-LArInA</i>
ABL	<i>-(t)tAn</i>	<i>-BIttAn</i>	<i>-GIttAn</i>	<i>-(t)IttAn</i>	<i>-BItIttAn</i>	<i>-GItIttAn</i>	<i>-LArIttAn</i>
INS	<i>-InAn</i>	<i>-BInAn</i>	<i>-GInAn</i>	<i>-(t)InAn</i>	<i>-BItInAn</i>	<i>-GItInAn</i>	<i>-LArInAn</i>
COM 8	<i>-LI:n</i>	<i>-BInA:n/ -BInI:n</i>	<i>- GInA:n/ -GInI:n</i>	<i>-(t)InA:n/ -(t)InI:n</i>	<i>- BItInA:n/ -BItInI:n</i>	<i>- GItInA:n/ -GItInI:n</i>	<i>- LArInA:n/ -LArInI:n</i>
COM P	<i>- TA:yA r</i>	<i>- BInA:yA r</i>	<i>- GInA:y Ar</i>	<i>- (t)InA:yA r</i>	<i>- BItInA:y Ar</i>	<i>- GItInA:y Ar</i>	<i>- LArInA:yA r</i>

Predicative suffixes are used to mark nominal and verbal predicates (Table 26.4).

They are the same for both languages.

Table 26.4 Predicative suffixes

Person	Predicative suffix
1SG	<i>-BIn</i>
2SG	<i>-GIn</i>
3SG	-
1PL	<i>-BIt</i>
2PL	<i>-GIt</i>
3PL	<i>-LAr</i>

Sakha and Dolgan distinguish three persons in singular and three in plural for the free personal pronouns. Every person can be inflected for all cases, using the case suffix of the non-possessive declension, but for 1SG and 2SG a different stem is used in inflected forms. Since for the dative case yet another stem is used, all three stems are given in Table 26.5⁹.

Table 26.5 Personal and possessive pronouns and personal inflectional stems

Person	Free pronoun	Inflectional stem	Dative	Possessive
1SG	<i>min</i>	<i>mi:gi-/ minigi-</i>	<i>miɣe/ minieke</i>	<i>miene/miniene</i>
2SG	<i>en</i>	<i>eʒigi-/ enigi-</i>	<i>eʒiɣe/ enieke</i>	<i>eʒiene/eniene</i>
3SG	<i>kini/gini</i>	-	<i>kinieɣe/ ginieke</i>	<i>kiene~kiniene/giniene</i>
1PL	<i>bihigi</i>	-	<i>bihiɣe/ bihieke</i>	<i>bihiene~bihiettere</i>
2PL	<i>ehigi</i>	-	<i>ehieɣe/ ehieke</i>	<i>ehiene/ehiene~ehiettere</i>
3PL	<i>kiniler/giniler</i>	-	-	<i>kiennere/giniettere</i>

Dolgan and Sakha have a special dual inclusive pronoun *bihikki* referring to ‘you and me’. There are no special politeness forms, and 2SG is used for all addressees regardless of age or social status.

Possessive pronouns can be used to reinforce possessive marked nouns (Table 26.5). Typically, these possessive pronouns are postposed, e.g. *mun’ax-pit bihiene* [meeting-1PL our] ‘our meeting’. However, in Dolgan they can be used as a modifier and occur in prenominal position:

(1) Dolgan (Story_Syn_marriage_APC_138)

onton miniene tohoyo-m ba:r
 then my birth.pole-POSS.1SG EXIST
 ‘...and then there is my birth pole’

The reflexive pronoun *beje* ‘self’ in combination with possessive suffixes is used with a purely reflexive meaning; it can also function as an emphatic marker (2).

(2) Dolgan (Story_Syn_Syndassko_PPK_26)

oskuola-ta huox-tar uraha die-ge üöret-el-ler
 school.R-POSS.3SG NEG-PL pole house-DAT teach-PRS-PRED.3PL
beje-lere tut-an-nar
 self-POSS.3PL build-SEQ.CVB-PRED.3PL

‘There was no school, they taught us in a tent that they had put up themselves.’

Demonstrative pronouns represent three degrees of distance: proximal, medial and distal (Table 26.6). In emphatic use a prefix *h(V)-* is added to the demonstrative e.g.

hu-bu [EMPH-this] ‘just this’, *ha-manna* [EMPH-here] ‘just here, exactly here’, *h-ol* [EMPH-that] ‘just that’. This form is particularly frequent in western Yakutia (3).

(3) Sakha (IvaP_039)

h-onnuk hu-bu nehiliek-ke üs ere ial ba:r

EMPH-such.a EMPH-this nasleg-DAT three only family EXIST

‘In this nasleg [administrative unit] there are only three families like that.’

Table 26.6 Demonstrative pronouns

	Nominative	Inflectional stem	Dative
Proximal	<i>bu</i>	<i>man-/man-~mun- ~bun-</i>	<i>maniaxa / maniga</i>
Medial	<i>iti</i>	<i>itin-</i>	<i>itiniexe / itinige</i>
Distal	<i>ol</i>	<i>on-</i>	<i>onuoxa / onuga</i>

Qualitative demonstratives (‘in this way’, ‘in that way’) are derived from the demonstratives with the adverbializing suffix *-Lik*, e.g. *man-nik* ‘in this way, like this’, *itin-nik* ‘in this way, like this’, *on-nuk* ‘like that’.

(4) Dolgan (Story_tundra_Syn_PPK_19)

on-nuk üöske:-bip-pit bihigi maɲnaj

this-ADV develop-PST2-1PL 1PL at.first

‘Like this we grew up at first.’

Quantitative demonstratives are derived using the suffix *-ččA* > *bačča* ‘this much’, *itičče* ‘this much’, *oččo* ‘that much’. *Oččo* is extending its meaning to ‘then, at that time’.

(5) Sakha (XatR_070)

en bačča ies-te:χ-χin dien, ol kihieχe oruobuna
 2SG this.much debt-PROPR-PRED.2SG say.SEQ.CVB that person.DAT exact.R
a:γ-an bier-er e-bit
 read-SEQ.CVB BEN-PRS.PTCP be-PST2[PRED.3SG]

‘... "you owe me this much", he read the exact (amount) to the person, it seems.’

Interrogative pronouns are largely the same in Sakha and Dolgan (Table 26.7). While most interrogative pronouns are unchangeable, *kim* ‘who’ and *tuoχ* ‘what’ have a special status since they can be case marked and pluralised. For this, Sakha uses the associative plural *-LA:χ* with *kim* (when asking about a plurality of humans) and the plural suffix *-LAr* with *tuoχ* (when asking about a plurality of things), whereas in Dolgan *kim* occurs with either *-LA:χ* or *-LAr*; the former refers to persons (6a) and the latter to things (6b).

Table 26.7 Interrogative pronouns

Pronoun	Meaning
<i>kim</i>	who
<i>tuoχ</i>	what
<i>χahan</i>	when
<i>χanna</i>	where
<i>χajdaχ</i>	how

<i>χas</i>	how many
<i>töhö</i>	how much
<i>χaja</i>	which, how
<i>χaja:-</i>	do what (Sakha only)

(6) a. Dolgan (Story_Syn_Holiday_MSA_21)

kim-ne:χ ba:l-lar ke tïa-γa

who-PROPR EXIST-PL CONTR tundra-DAT

‘Who all was there in the tundra?’

b. Dolgan (Story_Vol_LKS_194)

onno kim-ner-i, χonserva-lar-ï egel-iek-tere

then who-PL-ACC canned.food.R-PL-ACC bring-FUT-POSS.3PL

‘There they will bring what, canned food.’

Another important function of *kim* and *tuoχ* is their use as filler words. In this capacity, they may serve as a nominal or verbal stem, and can be declined and inflected according to the full morphological paradigm. In Sakha *tuoχ* is mostly used, whereas in Dolgan *kim* fulfils this function (6b).

Negative pronouns are the same as interrogative pronouns, with the addition of a particle and in combination with a negative predicate. In Sakha, either the particle *da* or *dayani* can be used, while in Dolgan, only *da* is used.

(7) Sakha (XatR_269)

kim da uor-bat

who PTCL steal-NEG.PRS[PRED.3SG]

‘Nobody stole.’

Indefinite pronouns are formed by interrogative pronouns in combination with an indefinite particle, whereby a distinction is made between the indefinite specific particle *ere* (Dolgan also *ire*) (8a), and the indefinite non-specific article *eme* (8b). According to the grammars (Ubrjatova 1985: 107; Artem’ev 2001b: 146), in Dolgan *dayani* is also used in this function, yielding the meaning ‘whoever, whatever’, but it does not occur as such in the corpus.

(8) a. Sakha (XatR_008)

ol_a:ta xahan ere bu dojdu leŋkir oju:r buol-an
 that.means when PTCL this land unbroken forest AUX-SEQ.CVB
tur-but e-bit
 stand-PST.PTCP be-PST2[PRED.3SG]

‘That means that at some time this country was covered by unbroken forest.’

b. Sakha (BesP_205)

bayar, üörem-mit-im e-bit-e
 perhaps learn-PST.PTCP-POSS.1SG be-PST.PTCP-POSS.3SG
[buol-lar] hin tugu eme büit-er-iem
 AUX-COND[PRED.3SG] MDL what.ACC PTCL end[INTR]-CAUS-FUT.1SG
[e-t-e]
 be-PST1-POSS.3SG

‘Perhaps if I’d gone to school, I would have finished something [a college].’

26.4.3 Numerals

The cardinal numerals are cognate to forms found in other Turkic languages, with a decimal-based system. The numerals from one to ten, as well as twenty, thirty, hundred and thousand (a Russian copy) are expressed by separate lexical items (Table 26.8). The decimals from 40 upwards are multiplicative and the other numerals are additive, e.g. 40 = *tüört uon* ‘four ten’, 50 = *bies uon* ‘five ten’, 11 = *uon bi:r* ‘ten one’, 12 = *uon ikki* ‘ten two’, etc.

Table 26.8 Numeral lexemes

1	<i>bi:r</i>	6	<i>alta</i>	20	<i>sü:rbe / hü:rbe</i>
2	<i>ikki</i>	7	<i>sette / hette</i>	30	<i>otut</i>
3	<i>üs</i>	8	<i>ayıs / agıs</i>	100	<i>sü:s / hü:s</i>
4	<i>tüört</i>	9	<i>toyus / togus</i>	1000	<i>tı:hınča</i>
5	<i>bies</i>	10	<i>uon</i>		

Ordinal numerals are derived with the suffix *-Is*, e.g. *ikk-is* ‘second’, *üh-üs* ‘third’, *törd-üs* ‘fourth’, etc. This suffix can also attach to the interrogative pronoun *χas* to derive ‘the how many-th’ (9a). There are separate lexemes for ‘first’: *mañnajgı* (9b) shared by both languages, or *bastakı* in Sakha.

(9) a. Sakha (Efmy_079)

biligin χah-ıs-ka *üören-er=ij*
 now how.many-ORD-DAT learn-PRS[PRED.3SG]=Q
 ‘In which (grade) is he now?’

b. Dolgan (Story_Vol_LKS_007)

öjdü:-bün, *mañnajgı* *učitel-ım* *Anna* *Grigorievna*
 remember.PRS-PRED.1SG first teacher.R-POSS.1SG Anna Grigorievna

köl-lör-ör *risunok-tar-ï*

look-CAUS-PRS[PRED.3SG] drawing.R-PL-ACC

‘I remember my first teacher Anna Grigorievna shows drawings...’

Approximative numerals are derived from decimals with the suffix *-čA*, e.g. *uon-čA* ‘about ten’, *sü:rbe-če* ‘about 20’, *otut-čA* ‘about 30’, etc., with the allomorph *-čeke* occurring with ‘hundred’, i.e. *sü:s-čeke* ‘about hundred’. Distributive numerals are derived with *-LI*: (10a), and numeral adverbs with a meaning ‘number of times’ are derived with the suffix *-TA*, e.g. *ikki-te* ‘twice’, *üs-te* ‘three times’, etc. Collective numerals are derived with *-IA*n (10b), which is shortened to *-IA* before predicative suffixes.

(10) a. Dolgan (Story_Xet_Kopytka_APF_51)

uon kino-nu köl-lör-d-üm ikki-li: kino-nu
ten film.R-ACC look-CAUS-PST1-POSS.1SG two-DISTR film.R-ACC
brigada-ya a:jï
bigade.R-DAT every

‘I showed 10 films, two films in every brigade.’

b. Sakha (Efmy_216)

barï da ba:r buol-al-lar, uon behien, ki:r-en
all PTCL EXIST AUX-PRS-PRED.3PL ten five.COLL enter-SEQ.CVB

‘And even all of them sometimes are here, all fifteen come by.’

Fractions exist only in Sakha; they are expressed with complex constructions, in which the denominator occurs first, followed by *gim-mit* [do-PST.PTCP] ‘done’ and the

numerator, which takes the case-marking required by the syntactic function of the constituent, e.g. *üs güm-mit bi:r-in* [three do-PST.PTCP one-ACC.3SG¹⁰] ‘one third[ACC]’, *tüört güm-mit bi:r-in* [four do-PST.PTCP one-ACC.3SG] ‘one fourth[ACC]’, *üs güm-mit ikki-tiger* [three do-PST.PTCP one-DAT.3SG] ‘two thirds[DAT]’. ‘One half’ is expressed in both Dolgan and Sakha with the lexeme *añar*.

26.4.4 Property words

Adjectives share many properties with nouns: they are often used as substantives (especially when they refer to humans), e.g. *ba:j* ‘rich’ and ‘rich person’, and they can take possessive and case marking (cf. Johanson 1998b: 39). Conversely, nouns can behave as modifiers in izafet constructions, e.g. *ta:s dʒie* ‘stone house’. However, the adverbializing suffix *-lik* attaches only to property words, e.g. *ulaɣan-nik* ‘a lot’ < *ulaɣan* ‘big’. Furthermore, there are suffixes that specifically derive adjectives, such as *-GI*, which derives adjectives from adverbs with a temporal meaning, e.g. *bilir-gi* ‘previous’ (<*bilir* ‘in the past’). As in other Turkic languages (Johanson 1998b: 39–40), partial reduplication can derive an intensive form of adjectives, e.g. *kıp-kıra* [INTS~small] ‘tiny’ (<*kıra* ‘small’), *ɣap~ɣara* [INTS~black] ‘very black’ (<*ɣara* ‘black’).

26.4.5 Verbs

There are two main paradigms of subject agreement markers in Sakha and Dolgan, which are identical to the nominal possessive and predicative suffixes (see Tables 26.3 and 26.4 above). The choice of paradigm is determined by the tense, with the present tense and resultative post-terminal past taking predicative person suffixes, while the simple past, post-terminal past, imperfective past, and future take possessive

suffixes. In addition, subordinate predicates take different case-marked possessive suffixes, and the person marking of predicates in realis conditional protases is historically derived from possessive-marked partitive suffixes.

Verbal negation is mostly suffixal, with the exception of the future tense. This is negated with *suoya/huoya*, e.g. *hit-iaŋ huoya* [lie-FUT.2SG NEG] ‘you won’t lie’.

26.4.5.1 Tense

There are seven different tenses in the indicative in Sakha and Dolgan: present, future, and five past tenses with different temporal, epistemic or aspectual values. The affirmative present tense takes predicative subject agreement marking, but lacks a unified tense marker. The 1st and 2nd person are formed with the markers *-A* or *-I:*, which historically derive from the simultaneous converb. *-A* is used with stems ending in a consonant, whereas *-I:* is used with stems ending in a long vowel or a diphthong, with the long high vowel replacing the final vowel of the stem, e.g. *bar-* ‘go’ > *bar-a-bin* [go-PRS-PRED.1SG] ‘I go’, and *hana:-* ‘think’ > *hani:-bin* [think.PRS-PRED.1SG] ‘I think’. The 3rd person is formed with the marker *-Ar* or *-I:r*, which show the same distribution, e.g. *bar-ar* [go-PRS[PRED.3SG]] ‘s/he goes’ and *hani:r* [think.PRS[PRED.3SG]] ‘he thinks’. The negative forms take the negation marker *-BA*, the origin of which is found in the negative present participle, e.g. *bar-bap-pin* [go-NEG.PRS-PRED.1SG] ‘I don’t go’, *bar-bat-tar* [go-NEG.PRS-PRED.3PL] ‘they don’t go’.

The affirmative future tense is formed with the suffix *-IAχ* (historically the future participle) and takes possessive subject agreement marking. The singular forms have two variants, full and contracted, e.g. *kör-üöy-üŋ* [see-FUT-POSS.2SG] = *kör-üöŋ* [see-FUT-POSS.2SG] ‘you will see’, *hiđđ-iaŋ-a* [go-FUT-POSS.3SG] = *hiđđ-ia* [go-

FUT.POSS.3SG] ‘she will go’, *kel-ieχ-tere* [come-FUT-POSS.3PL] ‘they will come’. In spontaneous speech the contracted forms are vastly more frequent than the full forms.

The simple past is marked by the suffix *-T* and takes possessive subject agreement marking (with the exception of 3PL, which takes the predicative agreement suffix); it is negated with the suffix *-BA*. Examples are *bar-d-īm* [go-PST1-POSS.1SG] ‘I went’, *bar-d-a* [go-PST1-POSS.3SG] ‘s/he went’, *taxsi-ba-t-īm* [go.out-NEG-PST1-POSS.1SG] ‘I didn’t go out’.

Two distant past tenses are formed with the markers *-BIt* (affirmative) and *-BAAtAχ* (negative), which are historically derived from the affirmative and negative past participle, respectively. Note that since the precise TAM reading of these distant past tenses is obtained only through the choice of possessive or predicative subject agreement, we here gloss these suffixes simply as PST2 (to contrast with the simple past suffix *-T*, glossed as PST1, and the imperfect past discussed below, glossed as PST3). The witnessed past takes possessive person marking, e.g. *kergen taxs-an bar-bīt-īm* [spouse exit-SEQ.CVB go-PST2-POSS.1SG] ‘I married’, *kel-betey-e* [come-NEG.PST2-POSS.3SG] ‘he didn’t come’. The unwitnessed past takes predicative subject agreement, e.g. *üleli: bar-bīt* [work.SIM.CVB go-PST2[PRED.3SG]] ‘she went to work’, *bier-betex* [give-NEG.PST2[PRED.3SG]] ‘she didn’t give’, *χahan kör-bük-kün=ij* [when see-PST2-PRED.2SG=Q] ‘when did you see?’.

In Sakha, the imperfect past has two affirmative variants, synthetic and analytic, with no difference in meaning, although the analytic construction—which is the sole means of expressing this tense in Dolgan—predominates in the oral corpus. The synthetic construction consists of the suffix *-Ar* (historically the present participle) taking possessive person marking, and the analytic construction consists of the lexical verb in participial form plus the auxiliary *e-* in simple past, e.g. *bar-ar-īm* [go-PST3-

POSS.1SG] ‘I went (frequently, often)’, *ölör-ör e-t-im* [kill-PRS.PTCP AUX-PST1-POSS.1SG] ‘I used to kill/I killed many’. The negative exists only as an analytic construction, e.g. *hugaha:-bat e-t-im* [come.close-NEG.PTCP AUX-PST1-POSS.1SG] ‘I didn’t come close’.

Finally, there is a remote past with plusquamperfect (11a) and resultative (11b) readings. Formally, it is parallel to the imperfect past, with the lexical verb carrying the suffix *-BIt* (historically the past participle) plus the auxiliary *e-* in simple past.

(11) a. Sakha (IvaP_018)

onu Muxta tiksi-ge bi:r-de bar-a hüdždž-a:rī gim-mit-im
 PTCL Muxta jetty-DAT one-MULT go-SIM.CVB IPFV-PURP do-PST2POSS.1SG
min inni-ber bar-bit boluot-tan tüöt kī:s öl-ön
 1SG front-DAT.1SG go-PST.PTCP raft.R-ABL four girl die-SEQ.CVB
χa:l-bit e-t-e, boluot adždžan-an
 RES-PST.PTCP AUX-PST1-POSS.3SG raft.R damage-SEQ.CVB

‘Once I was going to go to the Muxta jetty, and four girls who had gone on the raft ahead of me had died when their raft broke.’

b. Sakha (Efmy_318)

biligin da:nī učaska-ya bar-bit-tara üs oyo
 now PTCL quarter.R-DAT go-PST2-POSS.3PL three child
χon-o bar-bit e-t-e
 spend.night-SIM.CVB go-PST.PTCP AUX-PST1-POSS.3SG

‘Even now three children went to the quarter (outside the village), they have gone to spend the night.’

26.4.5.2 Mood

Sakha and Dolgan have a large number of moods. Neither the indicative nor the hortative or immediate imperative have a dedicated mood suffix, but both the indicative and the imperative distinguish tense—the indicative seven, as described in the preceding section, and the imperative two: an immediate imperative and a remote imperative. The jussive is marked by the suffix *-TIn* and takes predicative subject agreement marking, and the singular immediate imperative is expressed by the bare verb stem (Table 26.9). The hortative marks a distinction between minimal and augmented inclusive using the plural imperative marker *-(I)η*. The distinction between an immediate and remote imperative is likely to be the result of Evenki influence (Pakendorf 2007: 208–241).

Table 26.9 Hortative, imperative, and jussive suffixes

		Affirmative			Negative		
		SG	MIN.IN	AUGM.I	SG	MIN.INC	AUGM.IN
			CL	NCL		L	CL
Hortative		<i>-I:m</i>	<i>-IAχ</i>	<i>-IAyIη</i>	<i>-(I)m-I:m</i>	<i>-(I)m-IAχ</i>	<i>-(I)m-IAyIη</i>
Imperative	<i>Immediate</i>	<i>-∅</i>		<i>-(I)η</i>	<i>-mA-∅</i>		<i>-mA-(I)η</i>
	<i>Remote</i>	<i>-A:r-</i> <i>∅</i>		<i>-A:r-(I)η</i>	<i>-(I)m-A:r-</i> <i>∅</i>		<i>-(I)m-A:r-</i> <i>(I)η</i>
Jussive		<i>-TIn</i>		<i>-TIn-LAr</i>	<i>-BA-TIn</i>		<i>-BA-TIn-</i> <i>LAr</i>

The necessitive, which has been copied into Lamunkhin Even (Pakendorf 2009, 2014b), is formed with a bimorphemic suffix *-IAχ-LA:χ* plus predicative person marking, and the past necessitive is expressed by addition of the auxiliary *e-*. Another Sakha mood that has been copied not only into Lamunkhin Even, but also into Učur Evenki (Myreeva 1964: 51) is the assertive-presumptive marked by the suffix *-TAχ* and possessive subject agreement marking. Although this can mark probability, in spoken discourse it predominantly expresses a strong, emphatic assertion (Pakendorf 2009: 90–92).

The voluntative-potential, marked by *-(A:)jA* plus predicative person marking, expresses possibility and hope and can have an apprehensive meaning. The 2SG/PL affirmative forms have developed into a remote prohibitive (Pakendorf and Schalley 2007) in both Sakha and Dolgan (12).

(12) Dolgan (Story_Syn_marriage_APC_144)

onno di:-bin tohoyo-bun tüh-er-e:je-git
 then say.PRS-PRED.1SG birth.pole-ACC.1SG fall-CAUS-VOLPOT-2PL
 ‘Then I say: “Don’t knock over my birth pole”.’

The subjunctive mood is expressed by an analytic construction with the lexical verb carrying the future participle and the auxiliary in simple past tense; in Sakha, subject agreement is marked variably on the lexical verb (cf. 14a) or on the auxiliary, whereas in Dolgan it is marked only on the auxiliary (13).

(13) Dolgan (Story_Xet_Bear_TJP_074)

de onton dal’she bar-iaχ e-t-e kihi-ler

PTCL then further.R go-FUT.PTCP AUX-PST1-POSS.3SG human-PL

die-lerin alǰat-iaχ e-t-e

house-ACC.3PL destroy-FUT.PTCP AUX-PST1-POSS.3SG

‘He would have gone further, would have destroyed people’s houses.’

There are two different conditional forms that mark a difference between unreal and real conditions. Unreal conditions are expressed analytically, with the lexical verb carrying the present or the past participle plus possessive person marking and the auxiliary *buol-* carrying the conditional suffix *-TAr* (14a). Real conditions are expressed by the suffix *-TAr* plus conditional person marking (historically, possessive-marked partitive case endings which themselves developed out of the Old Turkic locative (cf. Pakendorf 2007: 148–149 and references therein)). This is frequently used in Dolgan (14b), and it also (and more commonly) functions as the predicate of temporal adverbial clauses in both languages.

(14) a. Sakha (XatR_095)

ari:j čepčeki-tik ep-pit-iŋ buol-lar

a.bit light-ADV say-PST.PTCP-POSS.2SG AUX-COND[PRED.3SG]

bayar il-iam e-t-e die-bit

perhaps take-FUT.1SG AUX-PST1-POSS.3SG say-PST2[PRED.3SG]

‘If you’d said a bit less, I might have taken it, he said.’

b. Dolgan (Story_cure_Syn_IMA_20)

onton buollayına χara bīar-iŋ iarij-day-ina buo ...

then PTCL black liver-POSS.2SG be.ill-COND-3SG PTCL

‘And then if your liver hurts...’

There are furthermore two moods with epistemic functions, the presumptive and the assertive. The non-past and past presumptive are expressed by the present or past participle, respectively, carrying possessive person marking plus the auxiliary *buol-* marked for future third singular, e.g. *kü:t-er-e buol-uo* [wait-PRS.PTCP-POSS.3SG AUX-FUT.3SG] ‘he is probably waiting’, *isti-bik-kit buol-uo* [hear-PST.PTCP-2PL AUX-FUT.3SG] ‘you probably heard’. In addition, presumption can be expressed with the particle *ini* ‘probably’. The affirmative assertive is marked by the suffix *-I:hI* plus predicative person marking (15a); the negative is expressed by the contracted form of the third singular future *-IA* plus the negative element *suoχ/huoχ* carrying predicative subject agreement (15b). These forms express a strong assumption or decision.

(15) a. Sakha (YmyE_158)

oččoyo bar-i:hi-gin

in.that.case go-ASS-PRED.2SG

‘In that case you can go of course,...’

b. Sakha (XatR_082)

huoχ il-ia huoχ-pun dien

NEG take-FUT.3SG NEG-PRED.1SG say.SEQ.CVB

‘Saying no, I won't take it,...’

26.4.5.3 Aspect/Aktionsart

As is common in Turkic languages (Anderson 2004; Johanson 1998b: 42), aspect and aktionsart in Sakha and Dolgan are mainly expressed via auxiliary verb constructions

(16), in which the lexical verb carries a converb marker, while TAM and subject agreement are expressed on the auxiliary. The auxiliaries also function as independent lexical verbs (Table 26.10).

(16) Dolgan (Story_Syn_Holiday_MSA_04)

onton bar-a tur-a-bit
 then go-SIM.CVB DUR-PRS-1PL
 ‘then we go [for a long time]’

Table 26.10 Common auxiliary verb constructions (constructions not found in the Dolgan corpus are italicized in the column “aspectual meaning”)

auxiliary	lexical meaning	SEQ.CVB	SIM.CVB	aspectual meaning
<i>bar-</i>	‘go’	+		sequential
			+	inchoative
<i>er-</i>	‘be’	+		inchoative
		+		continuous
<i>is-</i>	‘walk’	+		continuous
<i>kebis- / ke:s-</i>	‘throw’	+		perfective
<i>kel-</i>	‘come’	+		perfective
<i>oyus-</i>	<i>‘beat’</i>		+	<i>accelerative</i>
<i>sirüt-/hirüt</i>	‘go/be/wander’		+	imperfective
<i>tayıs-</i>	‘exit’	+		<i>durative</i>
<i>tur-</i>	‘stand’	+		resultative
			+	durative
<i>tüs-</i>	‘fall’		+	attenuative

<i>χa:l-</i>	‘stay’	+		perfective
--------------	--------	---	--	------------

While in Dolgan auxiliary verb constructions with *hirit-* largely still express a meaning of motion, in Sakha they have purely aspectual meaning, as seen by their occurring with stative verbs such as ‘live’, ‘be called’, ‘be seen’, or ‘have a population’.

The unaccomplished aspect is expressed by analytical constructions consisting of the lexical verb carrying the simultaneous converb suffix and the lexeme *ilik* carrying subject agreement marking; in the past tense, this combines with the past tense-marked auxiliary *e-*, e.g. *as-ti: ilik-pin* [food-VBLZ.SIM.CVB not.yet-PRED.1SG] ‘I have not yet cooked’, *kel-e ilik-terine* [come-SIM.CVB not.yet-COND.3PL] ‘when they had not yet come’, *taxs-a ilik e-t-im* [go.out-SIM.CVB not.yet AUX-PST1-POSS.1SG] ‘I hadn’t gone out yet’.

There are also some morphologically coded aspect/aktionsart forms: the habitual, the attenuative/accelerative, and the multiplicative. The habitual is formed with the suffix *-(A:)ččI* and predicative subject agreement marking. While this suffix also functions as a nominalizer in Sakha (cf. Section 26.4.6.2), in Dolgan it has only aspectual use and occurs very frequently (Stapert 2013: 215–224). The suffix *-MAχtA:* is described as an accelerative (Ubrjatova et al. 1982: 279), but it has an attenuative meaning in the Sakha corpus; this does not occur in the Dolgan oral corpus. The multiplicative derives verbs expressing actions that are performed multiple times (17a) or on multiple objects (17b). In Sakha there are three allomorphs of this suffix: *-(I)tALA:*, *-(I)ALA:*, and *-tA:*, with *-(I)ALA:* being restricted to the western dialects. In contrast, in Dolgan *-(I)tALA:* is absent.

(17) a. Dolgan (Story_Vol_ANS_113)

kühün buollayına kī:l it-īali:-bit
in.autumn PTCL wild.reindeer shoot-MULT.PRS-1PL

‘In autumn we shoot wild reindeer.’

b. Sakha (MatX2_027)

ol ebiet buh-ar-aŋ-ŋin aha-t-tali:-gin
that dinner.R ripen-CAUS-SEQ.CVB-PRED.2SG eat-CAUS-MULT.PRS-PRED.2SG

‘So you prepare their dinner and feed them (the multiple reindeer herders).’

26.4.5.4 Non-finite verbal morphology

Sakha and Dolgan make widespread use of converbs and participles. Common converbs are the simultaneous converb *-A*, the sequential converb *-(A)n*, the privative converb *-BAkkA*, the purposive converb *-A.rI*, the immediate precedence converb *-A:t* and the negative converb *-(I)mInA*. Sakha and Dolgan stand out among the Turkic languages in that these converbs (excepting the privative converb) can take person marking, even though all but the immediate-precedence converb occur with coreferential main clause subjects (Pakendorf 2007: 273–276; cf. (2) above and (27) below). The most commonly used converbs are the simultaneous converb *-A* and the sequential converb *-(A)n*, which not only function in auxiliary verb constructions (see Table 26.10), but also frequently express a simultaneous and a sequential action, respectively.

Sakha and Dolgan have a present, past, and future participle (*-Ar*, *-BI*, and *-IAχ*, respectively), which synchronically function widely as indicative tense markers (see Section 26.4.5.1), but which also serve as the predicates of relative clauses (Section 26.5.4). The present and past participle have negative counterparts (*-BA*t and *-BA*tAχ,

respectively). The habitual suffix *-(A:)ččI* can also function in nominal modification (18), although this might be obsolete in Sakha, where in the oral corpus this suffix occurs only as a finite verb or a nominalizer.

(18) Dolgan (Story_tundra_Syn_PPK_ 55)

iti olor-o:čču oyo-lor-um bari-ta taba üöreg-iger
 this live-HAB child-PL-POSS.1SG all-POSS.3SG reindeer education-DAT.3SG
bar-bit-tara
 go-PST2-POSS.3PL

‘My children who lived there all went to study for reindeer (veterinary medicine)’

26.4.6 Derivational morphology

26.4.6.1 Verbal derivation

Like other Turkic languages (Johanson 1998b: 42), Sakha and Dolgan have a morphological causative, passive, reflexive, and reciprocal. The passive of consonant-final stems is derived with *-(I)In*, that of vowel-final stems with *-(I)n*. The latter suffix is homophonous with the reflexive suffix while the reciprocal is derived with *-(I)s*.

Sakha and Dolgan share the causative suffixes *-TAr* and *-(I)Ar*, which occur after consonants, and *-T*, which occurs after vowels and the glide. In Sakha there is a further allomorph *-InnAr* not found in Dolgan. Although this formally looks like a combination of the reflexive suffix *-(I)n* with the causative *-TAr*, synchronically the reading is one of simple causatives, e.g. *itird-ïnnar-d-im* [sneeze-CAUS-PST1-1SG] ‘I made (this person) sneeze’. Causatives can be derived from both intransitive and

transitive stems and can also combine, e.g. *büt-er-ter-bit-im* [end[INTR]-CAUS-CAUS-PST2-1SG] ‘(I) let (her) finish’.

The verbal evaluative *-A:χtA:* expresses compassion towards the subject of the verb (19a). In Dolgan, it also has an augmentative function when combined with *dien* (19b). In addition, the habitual, attenuative/accelerative, and multiplicative discussed in Section 26.4.5.3 are derivational suffixes.

(19) a. Sakha (RaxA_050)

ki:r-en *bar-an* *mama-m* *χahi:ta:-t-a*
 enter-SEQ.CVB SEQ-SEQ.CVB mama.R-POSS.1SG scream-PST1-POSS.3SG
ïta:-t-a *u:* *oyo-m* *öl-ö:χtö:-bü* *di:*
 cry-PST1-POSS.3SG ooh child-POSS.1SG die-EVAL-PST2[PRED.3SG] ASS
die-t-e
 say-PST1-POSS.3SG

‘After going in my mother screamed, she cried, “ooh, my poor child has died”, she said.’

b. Dolgan (Story_Xet_Bear_TJP_053)

bari-tin *arij-d-a* *maltaj-an* *prjamo*
 all-ACC.3SG open-PST1-POSS.3SG have.broad.face-SEQ.CVB directly.R
a:η-ηa *kel-e:kti:r* *die-n* *maltaj-an*
 door-DAT come-EVAL.PRS[PRED.3SG] say-SEQ.CVB have.broad.face-SEQ.CVB
tur-a:χti:r *die-n*
 stand-EVAL.PRS[PRED.3SG] say-SEQ.CVB

‘He opened everything, with his big face right in the door, he comes in enormously, broad faced stands up gigantically.’

The most productive verbalizing suffix by far in Dolgan and Sakha is the suffix *-LA:*, which derives verbs from any part of speech with a very wide range of meanings. Thus, it can derive verbs from practically any noun, e.g. *üle* ‘the work’ > *üle-le:-* ‘to work’; from adjectives, e.g. *ira:s* ‘clean’ > *ira:s-ta:-* ‘to clean’; from numerals, e.g. *uon* ‘ten’ > *uon-na:-* ‘do ten times’; *ikkis* ‘second’ > *ikkis-te:-* ‘do for the second time’; or from pronouns, e.g. *tuox* ‘what’ > *tuox-ta:-* ‘do what’ (Sakha hesitantive), *kim* ‘who/what’ > *kim-ne:-* ‘do what’ (Dolgan hesitantive), *χanna* ‘where’ > *χanna-la:-* ‘go where’. In Dolgan, most loanverbs from Russian are integrated with the help of this suffix, e.g. *zvonit-ta:-* < *zvonit* ‘to phone’, *zamečaj-da:-* < *zamečat* ‘notice’.

26.4.6.2 Nominal derivation

Among the most productive nominalizing suffixes in Sakha and Dolgan is *-SIt*, which derives nouns referring to a person who performs an action on the object designated by the base noun, e.g. *bul-čut* ‘hunter’ < *bult* ‘prey’, *inaχ-sit* ‘milker’ < *inaχ* ‘cow’, *üle-hit* ‘worker’ < *üle* ‘work’. Among the suffixes deriving nouns from verbs both Sakha and Dolgan have *-I:*, e.g. *χa:y-ī:* ‘prison’ < *χa:y-* ‘lock up’. Furthermore, in Sakha but not Dolgan the suffix *-(A:)hIn* derives abstract nouns from verbs e.g. *batta:-* ‘to push, squeeze’ > *batta:hin* ‘pressure’. Lastly, the suffix *-(A:)ččI*, which expresses habitual aspect (see above), also functions as an agent nominalizer in Sakha (but not in Dolgan), e.g. *kömölöh-ö:ččü* ‘helper’ < *kömölös-* ‘to help’.

Sakha has copied the diminutive suffix *-čA:n* from Evenki, e.g. *uol-ča:n-iŋ* [boy-DIM-POSS.2SG] ‘your little boy’. In addition, both Dolgan and Sakha have copied the emphatic-diminutive suffix *-kA:n* from Evenki, e.g. *kira-ka:n* ‘tiny’ < *kira* ‘small’.

Dolgan uses a combination of the two with the demonstratives *bu* ‘this’ and *iti* ‘that’, turning them into emphatic adverbs, e.g. *bu-ka:-ča:n* ‘just like this’ < *bu* ‘this’.

26.4.6.3 Other derivational suffixes

Adjectives can be derived from nouns with the proprietive *-LA:χ*, e.g. *öj-dö:χ* ‘clever’ < *öj* ‘mind’. The suffix *-(I)msAχ* derives adjectives with a meaning of loving something, having a habit, e.g. *et-imseχ* ‘loving meat’ < *et* ‘meat’, while the suffix *-TA:yI* derives adjectives referring to a place from nouns and adverbs e.g. *tia-ta:yï* ‘someone/something in the taiga’ < *tia* ‘taiga’, *ira:χ-ta:yï* ‘someone far away’ < *ira:χ* ‘far’, i.e. the tsar.

The most productive suffix to derive adverbs from adjectives is *-Llk*, e.g. *üčügej-dik* ‘well’ < *üčügej* ‘good’. Finally, the suffix *-LI:* derives nouns referring to ethnolinguistic groups or their languages, e.g. *haχa-li:* ‘in Sakha’.

26.5 Syntax

26.5.1 The clause

The basic word order in Sakha is SOV. Deviations are possible, but carry pragmatic meaning (cf. Stapert 2013: 244–246). Converbs precede the final finite verb or auxiliary, adjectives, numerals, and demonstratives precede nouns, and the basic order within noun phrases is demonstrative-numeral-adjective-noun. This holds for Dolgan as well, but there is significantly more freedom in word order, probably due to contact with Russian (Stapert 2013: 262).

The particle *duo* marks yes/no questions, while the enclitic *=Ij* attaches to the predicate in content questions (20a, b).

(20) a. Dolgan (Story_Syn_Holiday_MSA_41)

die-tiger kel-te¹¹ duo

house-DAT.3SG come-PST2.POSS.3SG Q

‘Did he come home?’

b. Sakha (RaxA_275)

onu min Innokentij Vlasevič-tan ijüt-a-bin

PTCL 1SG Innokentij Vlasevich-ABL ask-PRS-PRED.1SG

en Kuoma-ni χahan kör-bük-kün=üj

2SG Kuoma-ACC when see-PST2-PRED.2SG=Q

‘So I asked Innokentij Vlasevich, “when did you see Kuoma?”.’

As is common in Turkic languages (Johanson 1998b: 41), in Sakha and Dolgan ‘being’ is expressed by the addition of predicative suffixes to the predicate nominal in the present tense (21) or with the inflected copula *e-* in the past tense. The predicative suffix is zero for 3SG, and 3PL is marked by the simple plural (cf. Table 26.4).

(21) Sakha (RaxA_251)

beje-ŋ em-čik-kin, buol-bat duo

self-POSS.2SG medicine-NMLZ-PRED.2SG AUX-NEG.PRS[PRED.3SG] Q

‘...you yourself are a healer, aren't you?’

In contrast to what is found in Turkic (Johanson 1998b: 56–57), the basic means to express ‘having’ in Sakha and Dolgan makes use of the proprietive suffix *-LA:χ* with the possessor cross-referenced with predicative suffixes (22a). The absence of having is expressed as in other Turkic languages, with the negative existential item *suoχ/huoχ*

and the person who lacks being cross-referenced with possessive suffixes on the lacking entity in Sakha (22b), and on *huoχ* in Dolgan (22c). The common Turkic means of expressing ‘having’ with possessive-marked possessum and the existential noun *ba:r* is rare, though found in the speech of individual speakers. Dolgan additionally makes use of constructions in which dative-marked pronominal possessors combine with the existential noun; this is rare in Sakha. A further means to express ‘having’ in Sakha is with the suffix *-LAn* (22d), which arguably grammaticalized out of the common verbalizer *-LA:* and the reflexive suffix *-(I)n*.

(22) a. Sakha (Chir_114)

oyonńor-do:χ-χun *duo*
 old.man-PROPR-PRED.2SG Q
 ‘Do you have a husband?’

b. Sakha (Chir_116)

huoχ *oyonńor-um* *huoχ*
 NEG old.man-POSS.1SG NEG
 ‘No, I don’t have a husband.’

c. Dolgan (Story_tundra_Syn_PPK_15)

ah-a *huoχ-put*
 food-POSS.3SG NEG-1PL
 ‘We don’t have food.’

d. Sakha (RaxA_271)

če, *üčügej-dik* *olor,* *die-t-e, [...]* *elbeχ*
 PTCL good-ADV sit[IMP.SG] say-PST1-POSS.3SG much
oyo-lon

child–ACQUIS[IMP.SG]

‘Well, live well, he said ...have lots of children,...’

26.5.2 The nominal group

It is in the case system that Sakha and Dolgan most notably differ from their Turkic relatives (cf. Pakendorf 2007: 94–207). The most salient differences concern the lack of a genitive and locative case and the existence of the partitive and comparative case as well as a separate comitative case.

In contrast to other Turkic languages, not only subjects of main clauses and generic indefinite direct objects, but also subjects of relative clauses and possessors stand in the unmarked nominative case. A remnant of the old Turkic genitive is found, however, in sequences of possessive NPs with a 3SG possessor (Pakendorf 2007: 99). The accusative case marks definite and specific indefinite direct objects, while the partitive marks partially affected or indefinite direct objects in the imperative mood. This case is more widespread in Dolgan than in Sakha, occurring even with verbs in the indicative mood (23).

(23) Dolgan (Story_Syn_Hunt_SEK_07)

onton ke atin istoria-ta keps-iem

then CONTR other story.R-PART tell-FUT.1SG

‘And then I will tell another story.’

Unlike other contemporary Turkic languages, the dative in Sakha and Dolgan conflates dative, locative and allative functions and marks recipients, addressees of speech, stative location, and the direction of movement. The ablative marks the source

of motion as well as a temporal starting point. In contrast to other “Transeurasian” languages, it does not mark the standard of comparison, since this function is carried by the dedicated comparative case. The instrumental marks instruments, means, adverbials of time, manner, cause, and space and is distinct from the comitative, which in Sakha is used to express a joint action of two animate referents of equal standing as well as the inclusive coordination of two NPs. This case is very rare in Dolgan, which instead makes heavy use of the postposition *gitta* ‘with’.

There is no plural agreement after numerals, e.g. *bies oyo* [five child] ‘five children’, *alta ki:s* [six girl] ‘six girls’, although there is plural agreement after nouns marked with the associative plural (24). Adjectives do not agree with their head nouns in either case or number.

(24) Dolgan (Story_Syn_Holiday_MSA_23)

Nata-la:χ huoχ e-t-iler

Natasha-PROPR NEG be-PST1-PRED.3PL

‘Natasha and her family weren’t there.’

As is common in the “Altaic” languages, the possessum agrees with the possessor in number and person. These possessive suffixes also function to cross-reference the subject in certain TAM forms (see Section 26.4.5). There is no formal distinction in the 3rd person between a single possessor with plural possessum (‘his horses’ *at-tar-a* [horse-PL-POSS.3SG]), a plural possessor with a single possessum (‘their horse’ *at-tara* [horse-POSS.3PL]), or a plural possessor with plural possessum (‘their horses’ *at-tara* [horse-PL.POSS.3PL]).

There are no morphological means to mark topic or focus in Sakha or Dolgan. Rather, a variety of particles are used, such as *buollayīna* or *ke* (the latter only in Dolgan, cf. 6a or 23), or the converb of the generic verb of speech *dien* (cf. Matic' and Pakendorf 2013: 383).

26.5.3 The verbal group

Predicates agree with their subjects in number and person, but as mentioned above there is no plural agreement after numeral-marked NPs.

The overt causee of causatives derived from intransitive verbs stands in the accusative case. Causatives derived from transitive verbs rarely occur with an overt causee in oral speech; but where they do, this mostly carries accusative marking (25), although dative marking also occurs.

(25) Sakha (XatR_126)

onu kenniki manna ostuoruja-tin il-an
 PTCL afterwards here history.R-ACC.3SG take-SEQ.CVB
balī:ha arχī:ba-tin irit-tar-bip-pit
 hospital.R archive.R-ACC.3SG scrutinize-CAUS-PST2-1PL

‘Afterwards in order to take his (medical) history we made the hospital archives scrutinize (everything),’

The agent of passive constructions is nearly always omitted, but where present it carries dative case marking:

(26) Dolgan (Story_Xet_Family_SNB_09)

kogda min oččuguj er-dek-pine kihi-ler-ge
 when.R 1SG tiny be-TEMP-COND.1SG human-PL-DAT
i:t-illi-bit-im
 raise-PASS-PST2-POSS.1SG
 ‘When I was small I was raised by (other) people.’

In the oral corpora, the reflexive has mainly “reflexive-possessive” or “reflexive-benefactive” meanings (cf. V. Nedjalkov 2003: 39–41) and hardly any simple reflexive meanings. This might be due to the nature of the corpora, which cover a rather restricted set of genres. The reciprocal functions as a true reciprocal, but also has “assistive”, “comitative” and “sociative” meanings (cf. V. Nedjalkov 2003: 30–32). Verbal predicates are mostly negated with a negative suffix (cf. Section 26.4.5); nominal constituents are negated with *suox* (cf. 22b above).

26.5.4 Complex sentences

There are various means to coordinate NP constituents (cf. Pakendorf 2007: 177–179 and Stapert 2013: 276–293 for examples): with the proprietive suffix *-LA:χ*; with the postposition *kitta/gitta* ‘with’ (this is the predominant means in Dolgan); with the conjunction *uonna* ‘and’ (Sakha only) or *onton* ‘and’ (Dolgan only); with the comitative case if two human subjects are being coordinated; with repetition of the plural suffix when proper nouns are being coordinated, or with repetition of *ikki* ‘two’—though this construction is archaic.

Main clauses (i.e. clauses with independent finite verbs) can be coordinated asyndetically, via mere juxtaposition, or with overt conjunctions, such as *uonna* in Sakha and *onton* in Dolgan (cf. Stapert 2013: 274–277). A further frequent means of

clausal coordination is by way of converbs—sometimes extending to quite long chains (e.g. Pakendorf 2007: 277, ex.114)—with only one final finite verb (27).

(27) Sakha (Mat1_033)

onton üöreχ-pin bīray-an kel-en bar-am-mīn
 then studies-ACC.1SG throw-SEQ.CVB come-SEQ.CVB SEQ-SEQ.CVB-PRED.1SG
χatta:n üören-em-min vsjō.taki medučilise-ni
 repeat.SEQ.CVB learn-SEQ.CVB-PRED.1SG after.all.R medical.institute.R-ACC
büt-er-bit-im
 end[INTR]-CAUS-PST2-POSS.1SG

‘Then after I gave up studying I came here, but then went back and finished the medical institute after all.’

As is common in the Transeurasian languages, subordination in Sakha and Dolgan largely makes use of non-finite predicates. However, temporal clauses can be expressed as finite clauses, with the predicate carrying the temporal-conditional mood marker (itself originally a participle) plus dedicated subject agreement suffixes. When the subject of the temporal clause is not coreferential with the subject of the main clause, the subordinate predicate is expressed by dative-marked present or past participles, with possessive suffixes referencing the person and number of the subordinate clause subject. In contrast, when the predicate of the main and the temporal clause are coreferential, the subordinate predicate is expressed by sequential or simultaneous converbs. The sequential converb also expresses the predicate of causal clauses.

As in other Transeurasian languages, the predicate of complement clauses in both Sakha and Dolgan is an accusative-marked participle, with possessive suffixes indexing the subject of the complement clause. Purpose constructions in Sakha and Dolgan are expressed by the dedicated purposive converb *-A:rI* (for coreferential purpose clauses) or accusative- and dative-marked participles (cf. Stapert 2013: 297–299). There are two further purpose constructions, however, in which the languages differ: in Sakha the converb of the generic verb of speech *dien* is commonly used to mark purpose clauses, as is widespread in Turkic languages (cf. Matic’ and Pakendorf 2013: 376–377). This construction is entirely lacking in Dolgan; instead, more than half the instances of purpose clauses in the corpus are expressed with the Russian subordinator *čtoby* in addition to the accusative-marked participle (Stapert 2013: 300–301).

Although all the Transeurasian languages make use of prenominal participial relative clauses, there is substantial variation as to the locus and extent of person and case agreement across languages (cf. Pakendorf 2012 for details). Sakha and Dolgan belong to the type called “head-marked” in Pakendorf (2012): in non-subject relative clauses the subject of the relative clause is indexed on the head noun with possessive suffixes (28).

(28) Dolgan (Story_Syn_Syndassko_PPK_28)

onton tup-put die-lere ba:r e-t-e
 then build-PST.PTCP house-POSS.3PL EXIST be-PST1-POSS.3SG

‘Then there was the house that they built.’

In Sakha, relativization on a wide range of constituents is possible, comprising subject, direct object, oblique, possessor, and locative. In Dolgan, in contrast, non-subject relative clauses occur rarely in spontaneous speech, and possessive person marking does not always agree with the number and person of the subject of the relative clause. Instead, 3SG marking appears to be used as a default (Stapert 2013: 311–315).

26.6 Lexicon

Sakha and Dolgan share a large number of cognates with other Turkic languages, especially in the basic lexicon. Nevertheless, both languages have adopted very many lexical copies, mainly from Mongolic and Russian, and to a lesser extent from Evenki (Kałużyński 1962; Pakendorf and Novgorodov 2009a; Romanova et al. 1975; Stapert 2013: 117–163).

The Mongolic copies in Sakha are found not only in cultural domains such as livestock terminology and terms pertaining to law and order, but even in basic lexicon such as kinship and body part terminology (Pakendorf and Novgorodov 2009b). Like the other minority languages of Siberia, Sakha has copied extensively from Russian, especially with respect to terms denoting a modern, Europeanized life-style. However, even some terms denoting natural phenomena entered the Sakha lexicon from Russian (Pakendorf 2007: 296–297; Pakendorf and Novgorodov 2009b). Interestingly, notwithstanding the evidence for structural influence from Evenki on Sakha (Pakendorf 2007: 303–305), the lexical impact of Evenki on Sakha is rather small, with only 35 items of Evenki origin being found in literary Sakha (Romanova et al. 1975: 163–166). Sakha dialects, however, have adopted more copies from Evenki, especially those spoken in areas where Sakha have adopted reindeer herding.

Compared to Sakha, Dolgan has undergone considerable semantic changes in its lexicon, especially in the domains of body part and kinship terminology, as well as having copied more items from Russian and Evenki (Stapert 2013: 128–129). Some of the semantic changes in Dolgan are likely to have occurred under influence from Evenki, such as the extension of the word *ulluŋ* ‘sole’ to also denote ‘foot’, or the restructuring of kinship terminology (Stapert 2013: 131–142).

¹ Eugénie Stapert compiled the information on Dolgan and wrote the sections on phonology and on morphology of nouns and pronouns, and Brigitte Pakendorf compiled the information on Sakha and wrote the rest of the chapter.

² We are grateful to the Wenner-Gren Foundation, the Max Planck Society (via the Department of Linguistics and the MPRG on Comparative Population Linguistics at the MPI for Evolutionary Anthropology, both now defunct), and to the INEL-project funded by the Union of the German Academies of Sciences and Humanities, for funding the field trips to compile and check these corpora, and to all the Sakha and Dolgan individuals who contributed to our work. BP is furthermore grateful to the LABEX ASLAN (ANR-10-LABX-0081) of Université de Lyon for its financial support within the program "Investissements d'Avenir" (ANR-11-IDEX-0007) of the French government operated by the National Research Agency (ANR).

³ In Section 26.3, we use IPA for transcription. Correspondences with the transcription adopted throughout the rest of the chapter are (where different): $n=\acute{n}$, $\text{ʃ}=\check{c}$, $d_z=d_3$, $i=\acute{i}$, $\varepsilon=e$, $\text{ø}=\ddot{o}$, $y=\ddot{u}$.

⁴ Throughout the chapter, morphemes in isolation are presented in their underlying form, with capital letters indicating phonemes that undergo changes according to vowel harmony and consonant assimilation rules.

⁵ This phenomenon is also found in other Turkic languages (Johanson 1998b: 31–32) but in these cases the elision of the high vowel occurs only before vowel-initial suffixes.

⁶ Note that in Sakha and Dolgan most finite tense markers are transparently derived from participles and are identical with these in form. However, our glosses reflect their synchronic function, so that we gloss the same form either as a tense marker or as a participle.

⁷ This suffix was derived from the Mongolic plural marker *-t*, which occurred after a wordfinal *-n*. The suffix was copied from Mongolic into Sakha, and was reinforced by the Turkic plural suffix *-LAR* (Pakendorf 2015: 167–168).

⁸ Possessive declension only in Sakha.

⁹ Throughout the chapter, when Sakha and Dolgan formally differ, the Sakha forms are presented first and are separated from the Dolgan forms by a slash.

¹⁰ Note that for reasons of space we use the truncated gloss “case.person.number” to designate case-marked possessive suffixes. Thus, ACC.3SG is to be read as “possessive 3SG in accusative case”.

¹¹ Note that *-te* is the contracted form of *-bit-e*; the 3SG simple past tense of *kel-* would be *kel-l-e*.