Chapter 19

Nanai and the Southern Tungusic languages

Sofia Oskolskaya

Abstract

The chapter deals with an overview of the Southern Tungusic languages: Hezhe, Udihe, Oroch, Nanai, Ulcha and Orok. The sociolinguistic and dialectal situation is described as well as the history of the languages under discussion in respect of genealogy and contacts. The most specific features of this language group are observed alongside with the features shared with the other Transeurasian languages. An outline of the Nanai language is given as an example of one of the Southern Tungusic languages with a special focus on historical development of some Nanai peculiarities (derivation of negative forms, substitution of the finite forms by non-finite forms and others). Some of the Nanai features are discussed in comparison with other closely related languages.

Keywords: Nanai, Southern Tungusic languages, dialects, grammatical features

19.1 Introduction

The Southern Tungusic languages include Kilen, Udihe, Oroch, Nanai, Ulcha and Orok languages. The list is given according to Glottolog (Hammarström et al. 2018), where this group of languages is called Central Tungusic (while the term "Southern Tungusic" refers to Avrorin 1959). The Kilen language is also known as Hezhe with two dialects, Kilen and Hezhen (cf. Zhang 2013). I prefer to use Hezhe as it is more precise and unambiguous: it denotes only the language and not a Hezhe dialect. Orok has an alternative name Uilta (the self-designation of the Orok people), which has been admitted since 1991 (cf. Ikegami 1997; Tsumagari 2009b etc.). The Nanai language used to have an alternative name "Gold" before 1917. Although it is not in use any more, it has its relict in ISO-code, "gold".

The Southern Tungusic peoples inhabit basins of the rivers Amur, Sungari and Ussuri and the Sakhalin island on the territory of modern China and Russia (see Figure 19.1).

<Insert Figure 19.1 here>

Figure 19.1 The location of the Southern Tungusic peoples

All the Southern Tungusic languages are at different stages of endangerment. Speakers are bilingual, they speak their native Tungusic language, and Russian or Chinese, depending on the country they live in. Nanai seems to be the most living in this group of languages. According to my fielddata, there are not more than 300 speakers of Nanai nowadays: these are people who can fluently produce sentences in Nanai. Most of them are older than 1970 year of birth. The amount of the Nanai people in Russia is about 12 000 people according to the Russian population census (2010). The population of the Ulcha is much less—2765 (Russian population census 2010), there were about 100 speakers in 2010 (Kazama 2010b). The Udihe accounts for 1493 people, while there are only several speakers left (Perekhvalskaja 2012: 91). Among 295 Orok only 10 marked the Orok language as their native one in the census in 2010, which corresponds to the scholars' estimations. Thus, Tsumagari (2009b: 2) gives the amount of speakers between 16 and 25 referring to Bibikova (p.c.) and Ozolinja (2001) respectively. According to the Russian population census (2010), there are only 3 speakers of Oroch among 594 people. The field data of Perekhvalskaja are even more pessimistic: she points out that there are no Oroch speakers (Perekhvalskaja 2016a). So, the Oroch language may be considered extinct.

Although the population of Hezhe accounts for more than 5000 people (the Chinese census 2010), only few of them can speak Hezhe (Zhang 2013: 10).

The first linguistic works dedicated to the Southern Tungusic languages refer to the 19th century. They are mostly wordlists and dictionaries. There were no established writing systems until 1931 when an official alphabet based on Latin characters was developed for the majority of

the languages of Russian North. Then, Nanai and Udihe obtained their own Latin orthography, which was modified with account of the languages' properties by Avrorin and Schneider respectively (Onenko 1980: 6; Perekhvalskaja 2016b: 510). Later, in the end of 1930s, Nanai and Udihe writing systems were transformed to the Cyrillic systems as it was done for the most of other languages of the USSR. The Latin and Cyrillic writing systems were used in school books, books translated from Russian and in dictionaries. Nowadays, besides Nanai and Udihe, Ulcha has also its own orthography based on Cyrillic characters. It was compiled in 1990s (Sumbatova and Gusev 2016). Oroch, Orok and Hezhe languages do not have established writing systems.

The period of 1930s is also characterized by the beginning of intensive research of the languages under discussion when grammatical descriptions and vocabularies were published (see Petrova 1936, 1941, Snejder 1936 and others). The Nanai language has been the most popular among the scholars. An essential grammar was written by Avrorin (1959, 1961). Other grammar descriptions should be also mentioned: Petrova (1941, 1960), Kazama (2010a), Ko and Yurn (2011). Nanai dialects were described in Sunik (1958, Kur-Urmi dialect) and in Sem (1976, Bikin dialect). The most complete Nanai dictionaries, Nanai-German and Nanai-Russian, are Grube (1900) and Onenko (1980). Furthermore, a number of more particular investigations of Nanai have been made, such as Kim Cher Len (1968), Kile (1973), Avrorin (1981), Gerasimova (2006) and many others. Some scholars published their material—texts collected during fieldwork, among them: Avrorin (1986), Kile (1996), a collection of books with Nanai texts have been published by Kazama (1991, 1993 and others). Some of the grammar descriptions also contain sample texts in Nanai.

In Table 19.1, I present a list of principal grammar descriptions and dictionaries for Udihe, Ulcha, Orok, Oroch and Hezhe. Of course, the list is not exhaustive.

Table 19.1 Sources for the Southern Tungusic languages (excluding Nanai)

language	grammars	dictionaries and wordlists
Udihe	Nikolaeva and Tolskaya (2001), Kormushin (1998), Snejder (1936)	Udihe-Russian: Kjalundzjuga, Simonov (1998–1999), Girfanova (2001), Kormushin (1998)
Ulch	Petrova (1936), Sunik (1985), Kazama (2010b)	Ulch-Russian: Sunik (1985)
Orok	Tsumagari (2009b), Nakanome (1917/1928), Ikegami (1956, 1959, 1973), Petrova (1967)	Orok-Japanese: Ikegami (1997), Magata (1981)
Oroch	Cincius (1949a), Avrorin and Lebedeva (1968)	Oroch-Russian: Avrorin and Lebedeva (1978); Oroch-English: Schmidt (1928)
Hezhe	An (1986), Zhang et al. (1989), Zhang (2013)	Hezhe-Chinese: An (1986); Hezhe-English: Zhang et al. (1989)

19.2 Historical connections: genealogy and contact

The Southern Tungusic languages belong to the Tungusic family along with two other groups:

Northern Tungusic and Manchuric. Although the classification of the Tungusic languages is still

under debate, the Southern Tungusic group coincides with the branches of classifications

proposed by other scholars: with Goldi and Udihe subgroups of Southern/Manchu group in

Cincius (1949b), Benzing (1955a), with the Nanai and Udihe branches of Tungusic languages in

Ikegami (1974), with Amur subgroup of Tungusic group in Sunik (1997), with Central Tungusic group in Doerfer (1978a), with East Tungus languages in Vovin (1993c), see Whaley and Oskolskaya, this volume: Chapter 6 "The classification of the Tungusic languages" for further details.

The relation of the Tungusic family to other Transeurasian languages is much more questionable. It is obvious that Tungusic languages share a lot of common linguistic features with other Transeurasian languages, but the origin of this similarity remains disputable. As I do not know strong arguments for or against this relation, or probably I am not aware enough of the problem, I prefer not to take this or that side. Further discussion in the chapter will not bring up this problem.

Janhunen (2005b) suggests that first Manchuric (or Jurchen-Manchu) became a separate group about 2000 years ago and then other Tungusic groups (Nanai, Udihe and Evenki, according to Ikegami 1974) emerged between 2000 and 1000 years ago. The same age is proposed by Pevnov (2008). The Southern Tungusic languages form a separate group on the ground of phonological, morphological and lexical features.

Doerfer (1978a) develops his classification of Tungusic languages based on some phonological and morphological features. The terminology, which he uses, differs from ours: the Northern, Central and Southern groups in his paper correspond to the Northern, Southern and Manchuric groups in this volume. The most specific features of the Southern branch (the Central one in terms of Doerfer) are the following: i) the transition of Proto-Tungusic *-k- to k (Oroch, Kili[1]), g ~ ' (Udihe) or Ø (Nanai, Ulcha, Orok): e. g. *baka- 'to find' > baka- (Kili[2]), bʾa- (Udihe), baː- (Nanai, Ulcha, Orok); ii) the case system includes from 7 to 9 cases, which is significantly less than Northern Tungusic languages have (9–14 cases) and more than Manchu has (5 cases). Other phonological and morphological features that Doerfer investigates are shared either with Northern Tungusic or with Manchuric languages, which shows a transitional state of languages under discussion between these two groups. On the basis of selected features, Doerfer

(1978a: 5) suggests the following subdivision of the Southern Tungusic branch, as shown in

Figure 19.2 (marking the transitional position of some languages and groups)[3].

Southern branch (Northern branch → Manchuric branch)

    Southern-Eastern group (Southern branch → Northern-Western group)

      Oroch

      Udihe

    Southern-Western group (Southern branch → Manchuric group)

      Kili (Nanai → Evenki)

      Nanai

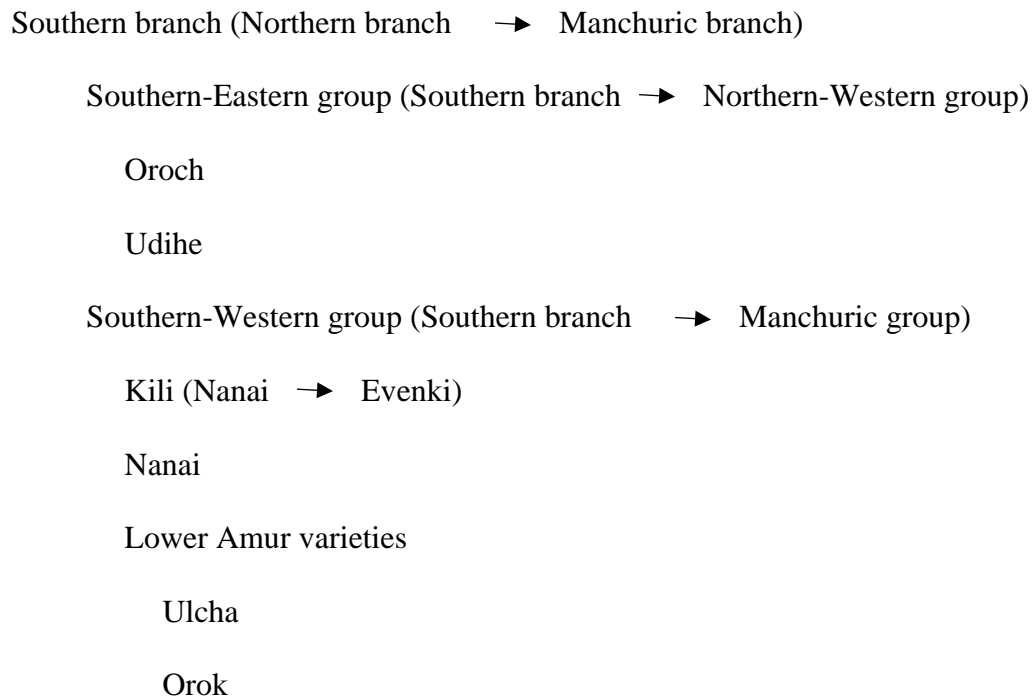   Lower Amur varieties

        Ulcha

        Orok

Figure 19.2 Subdivision of the Southern Tungusic branch

There is also a number of other morphological features, which are mostly specific for the

Southern Tungusic languages. Among them one could mention a destinative (or designative)

case (cf. Malchukov 2009: 644–646). This case marks the direct object intended for a particular

person. A beneficiant person is marked by a possessive suffix on the same noun (1).

(1)   Nanai (Avrorin 1986: 230)

    *təj   mapa   ńoanči   pokto-**go**-a-**či**   aŋgo-j-ni*

    that   old.man 3PL   way-DEST-OBL-3PL make-PRS-3SG

    'That old man makes a (hiking) trail for them.'

The example (1) can be compared to (2) where the same word *pokto* 'way' occurs only with the oblique case marker.

(2) Nanai (Avrorin 1986: 242)

  səːpə pokto-wa-ni   ičə-ʤəː-či=məː

  sable way-OBL[5]-3SG   see-FUT-2SG=PTCL

  'You will see a sable's trail.'

The sentence (2) does not have a destinative meaning 'you will see a trail for a sable'. A special destinative marker is attested in all Southern Tungusic languages and in Even (cf. Cincius 1949b: 256). Its origin remains unclear. Malchukov (2009: 645–646) proposes two possible explanations. One explanation implies that the case marker relates to the verb *ga-* 'to take'. Another explanation refers to a partitive marker, which was reinterpreted as a destinative one in the context of possessive markers. It may be supported by the use of an "indefinite" accusative marker in Evenki: its basic meaning is partitive but it functions as a destinative marker when acquiring possessive suffixes (Nedjalkov 1997: 147–148).

Another specific morphological feature of the languages under discussion is the repetitive verbal marker that marks a repetitive or a restitutive action (3).

(3) Nanai (Oskolskaya, fieldnotes)

  a. *ʤi-či-ni*

    come-PST-3SG

    'he came'

  b. *ʤi-**ʤu**-hə-ni*

    come-REP-PST-3SG

    'he came back'

Although the repetitive marker is attested both in Southern and Northern Tungusic languages, its type-frequency, i.e. the amount of stems that can acquire the suffix, differs much. In all the Southern Tungusic languages and in Negidal, a Northern Tungusic language spoken in the region of Lower Amur (see Pakendorf and Aralova, this volume: Chapter 18), it is highly productive and it is able to attach to most of the verbal stems, while in Evenki and Even it occurs with a very restricted number of verbs (see Stoynova and Oskolskaya 2014 for more details). It allows to assume that the distribution of the repetitive marker has been gradually reduced in most of the Northern Tungusic languages.

Vovin (1993c) showed the lexical peculiarity of the Southern Tungusic languages: he developed the classification of the Tungusic languages using M. Swadesh 100 word list for 10 Tungusic languages. According to his data, Udihe, Oroch, Orok, Ulcha and Nanai form the East Tungus group[6]. The lexical data of M. Swadesh 100 word list allow to make a further subdivision: the East Tungus divide into Orok-Nanai subgroup and Udihe and Oroch languages. Orok-Nanai subgroup can be divided into Nanai subgroup which includes Nanai and Ulcha languages and the Orok language (see Figure 19.3).

Proto-East Tungus

Proto-Orok-Nanai

Proto-Nanai

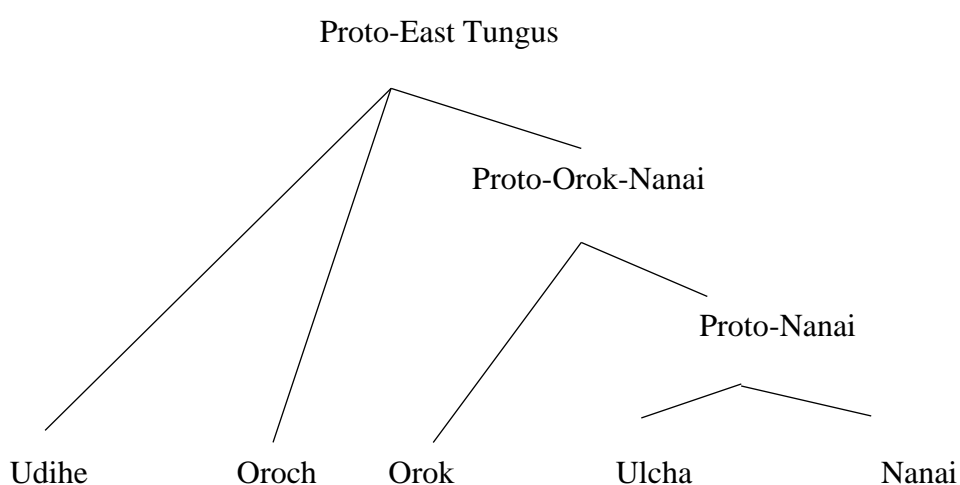Udihe   Oroch   Orok   Ulcha   Nanai

Figure 19.3 Family tree of the East Tungus languages (part of the family tree for the Tungusic languages (Vovin 1993c: 113)

Some of the Southern Tungusic languages have dialects, which could be regarded as transitional lects between languages of the group or to the languages of other Tungusic groups. As it was mentioned above, Hezhe—a Southern Tungusic language spoken in the Northern-East of China—has two dialects: Hezhen and Kilen. Scholars used to consider Hezhe as a Nanai dialect (Avrorin 1959 and others). Nowadays, it is often described as a separate language (Kazama 1996). The Hezhe language is closely related to Udihe. Udihe has mainly two dialects: the Southern and the Northern dialects (Nikolaeva and Tolskaya 2001: 6–7). The Southern dialect is spoken in the Primorje region of Russia in the area of Bikin and Iman rivers, the Northern dialect is spoken in the Khabarovsk region of Russia in the area of the Khor, Anyui, and Samarga rivers. In 1930s, the Khor Udihe was chosen as a standard literary variant of Udihe. The Oroch language had three dialects: Tumnin, Khungari and Khadja (the names are given following the river names, i.e. the area where dialects were spoken). The Khadja dialect was influenced by Udihe, while Tumnin was influenced much by Negidal (Perekhvalskaja 2016a: 339) or by the Gorin dialect of Nanai (Lebedeva 1997: 215). The Khungari dialect was the most distinct one.

The Nanai subdivision into dialects varies from one scholar to another. Here, I give the dialectal subdivision based on Doerfer (1978a: 23–24) as it corresponds to my own fielddata in many respects: Bikin / Ussuri, Sikači-Aljan, Najkhin / Torgon, Dzhuen, Bolon, Gorin (listed from South-West to North-East). The Bikin (or Ussuri) dialect is spoken on the border of the Primorje and the Khabarovsk regions of Russia along the river Bikin. It is closely related to Udihe. The Sikači-Aljan dialect is spoken in the Sikači-Aljan village, which is situated by the side of the Amur river not far from Khabarovsk. It shares some features with Hezhe. Najkhin (or Torgon), Dzhuen and Bolon dialects are located in the basin of the Amur river between cities of Khabarovsk and Komsomolsk. The Najkhin dialect is spoken in several villages and towns settled along the Amur river. In 1930s, Najkhin was used for developing a standard literary variant of Nanai. The Dzhuen dialect is spoken in the village Dzhuen at the lake Bolon and it is

in the best condition among Nanai dialects at the moment (some Dzhuen inhabitants still regularly use their native language in everyday life). The Bolon dialect is spoken mostly in the villages Achan (at the lake Bolon) and Nergen (on Amur). The Gorin dialect is located in the village Kondon at the basin of the Gorin river and shares some linguistic features with the Negidal language.

A special rather disputable case is Kur-Urmi (or Kili). Some scholars used to consider it as a Nanai dialect (Avrorin 1959; Sunik 1958). An independent language status of Kur-Urmi was discussed by Doerfer (1973a, 1975a). Pevnov (2008) and Janhunen (2005b: 42) suggest that Kur-Urmi is a mixed language based on Evenki (lexical and phonological features) and Nanai (morphological and syntactic features).

Previously, the Ulcha language was also regarded as a Nanai dialect because it shares a lot of linguistic features with Nanai (see Petrova 1936), but now it is considered to be a separate language (see Sunik 1985) due to some linguistic features and self-identification of the people. Scholars did not subdivide Ulcha into dialects. Nevertheless, my fielddata show that there might be two dialects under question: the so-called Lake and the River dialects referring to the Ulcha who live at the lake Udyl and at the Amur river respectively. The dialects are certainly intelligible but have some slight differences, at least in vocabulary. The Lake dialect seems to have been influenced by Negidal or Evenki. Still, the question of Ulcha dialects should be investigated more thoroughly.

The Orok language is spoken in the Sakhalin island and includes two dialects: the Northern one and the Southern one (Tsumagari 2009b: 1).

The Southern Tungusic languages as well as other Tungusic languages share some linguistic features with other Transeurasian languages. Below there is a list of prototypical Transeurasian features which characterize the Southern Tungusic languages (the list is based on Robbeets 2017g).

Phonological features

i. The languages under discussion are not tonal. The vocal system of Udihe includes laryngealized vowels. Janhunen (1999) suggests that they could appear under influence of the tonal Chinese language. Anyway, Udihe does not have a complex tonal system.

ii. Vowel harmony is observed in all the Southern Tungusic languages. However, some stipulations can be made for every language concerning its realization (see Section 3.4 for Nanai vowel harmony).

iii. Initial *r-* is absent in all the languages under discussion except for recent borrowings.

iv. Initial consonant clusters are also absent.

Morphological features

i. Udihe, Oroch and Hezhe have different forms for inclusive and exclusive 1st person plural pronouns. A distinction of this kind is also attested in Northern Tungusic languages and in Turkic, Mongolic and Korean languages as well.

ii. The morphology of the Southern Tungusic languages is agglutinative: morphemes have one-by-one relationship with their meanings.

iii. Inflectional as well as derivational morphology is completely suffixing.

iv. The Southern Tungusic languages display *mi~Ti* opposition in 1st vs. 2nd singular pronouns, that is labial nasal *m* is observed in 1st person singular pronominal paradigms while apical or palatal obstruent *t, c, s* etc. is observed in 2nd person singular pronominal paradigms. The only stipulation that should be done is that the nominative form of 1st singular pronoun in all the Southern Tungusic languages except for Nanai is *bi* rather than *mi*. Still, the oblique case form is *min-* in all the languages.

v. A secondary oblique stem of personal pronouns is derived with dental nasal *-n-*: e.g. Na. *si:* [2SG.DIR] – *sin-* [2SG.OBL].

Syntactic features

i. A basic word order is SOV (subject – object – verb) as in the most languages of the area. However, it is not strict in the Southern Tungusic languages.

ii. Modifiers normally precede heads (genitive – noun / adjective – noun).

iii. Converbs are the main means for clause-chaining. They are used to mark subordination and to link clauses, see example (5).

(5) Udihe (Nikolaeva and Tolskaya 2001: 740)

in'əi-wə tindaŋi-gə-si   ŋənəː-ti   caihi

dog-ACC let-PRF-PTCP.SS   go.PST-3PL further

'Having loosened the dogs they went further.'

iv. The Southern Tungusic languages express a possession basically with a locative existential construction, see example (6).

(6) Udihe (Nikolaeva and Tolskaya 2001: 622)

min-du käŋa jäː-ni   biə

me-DAT deer antlers-3SG be.PRS.HAB

'I have deer's antlers.'

v. Ablative case form is used for marking the standard of comparison in predicative comparative constructions, see example (7).

(7) Udihe (Nikolaeva and Tolskaya 2001: 180)

ussuri   biki-digi   suŋta

Ussuri   Bikin-ABL deep

'The Ussuri is deeper than the Bikin.'

This is the case that can be observed in Udihe and Hezhe. In other Southern Tungusic languages the state of affairs is a bit more complicated. Thus, in Nanai and Oroch the general ablative case marker is *-ʤi(ə)ʤi* while the standard of comparison is marked by *-du(j)*. The suffix *-duj* formally coincides with the combination of dative and reflexive markers—*-du-j*. But etymologically, it originates from the ablative marker **-duki* which reduced its functions to the comparative one (Avrorin 1959: 185–186).

In Orok and Ulcha the standard of comparison is marked by *-ʤi* which formally coincides with the instrumental case marker. However, the instrumental case marker in Ulcha and Orok also serves as an ablative one (Petrova 1936: 29; 1967: 50), see example (8).

(8)   Orok (Petrova 1967: 50, glosses are mine—*S. O.*)

Noːni duku-ʤi   ŋənə-hə-ni.

3SG   house-INS   go-PST-3SG

'He went away from his house.'

Petrova (1936: 29) points out that coincidence of the suffix *-ʤi* in ablative-comparative function with the instrumental marker may be accidental: in fact, ablative-comparative *-ʤi* may be a reduced variant of the mere ablative case marker *-ʤiʤi*.

Thus, we can assume that the standard of comparison in Southern Tungusic languages is marked by a suffix that is somehow related to the ablative function.

Grammaticalization features

i. A common feature of the Southern Tungusic languages is the use of historically non-finite verb forms in the position of finite verbs (i.e. as a head of a sentence), that is called direct insubordination (Robbeets 2017g: 603). Thus, Orok form *ŋənə-hə-ni* [go-PST-3SG] from example (7) can also occur as a past participle and serve for characterizing a nominal phrase. Etymologically, a participial function was primary.

ii. The negative verb *e- or a verbal negative suffix grammaticalized from the negative verb are generally used for expressing verbal negation (e.g. Na. *əčiə ńiru-ə* [NEG.COP write-NEG] '(he) did not write' vs. *ńiru-ə-si-ńi* [write-NEG-PRS-3SG] '(he) does not write'). Actually, the negation system of the Southern Tungusic languages is rather complicated. Some of the lects display analytical negative constructions which are based on the use of the existential negative marker:

(9) Nanai, Sikachi-Aljan dialect (Stoynova and Oskolskaya 2015)

  a. *mi   čisəniə   dʒobo-m(=da)   aba-(i)*

   1SG   yesterday   work-CVB.SIM.SG=PTCL   NEG-1SG

   'I didn't work yesterday.'

  b. *kalgama   aba*

   yeti   NEG

   'Yeti does not exist.'

Still, the relicts of the negative verb *e- are attested in all Southern Tungusic languages.

A number of features are not typical for most of Transeurasian languages (primarily for Turkic and Mongolic). First, the vowel harmony in Southern Tungusic languages as well as in other Tungusic languages is usually described as a height harmony, while in Turkic languages and in Western Mongolic languages, such as Oirat and Kalmyk, a palatal harmony is attested. The Khalkha language displays a tongue root retraction system. However, Ko (2012) suggested that the Tungusic vowel harmony is also retracted tongue root based.

The features mentioned above as specific for Southern Tungusic languages can be also regarded as atypical across other Transeurasian languages, i. e. the use of a destinative case marker and a repetitive verbal marker. Another Tungusic specific morphological feature is the

occurrence of two imperative forms: immediate imperative and distant imperative, see example (10).

(10) Nanai (Oskolskaya, fieldnotes)

  a. *nasaptom-ba-ja baːg-u*

    glasses-OBL-1SG find-IMP

    'Find my glasses (begin to search right now)'

  b. *mi   nasaptom-bi   huədə-uči-ə-jə   baːgo-hari*

    1SG   glasses-REFL.SG   lose-CVB.COND2-OBL-1SG   find-IMP.DIST

    'If I lose my glasses, find (it).'

Two kinds of imperative forms are attested in all Southern (except for Oroch) and Northern Tungusic languages. A special distant imperative form seems to be absent in other groups of Transeurasian languages.

   Some other features atypical for most of Transeurasian languages can be described as negative—the absence of this or that feature in Southern Tungusic. A special genitive case marker is absent in the languages under discussion. The Manchuric languages is the only Tungusic group where genitive case is attested. Other Transeurasian languages have genitive in their case systems.

   The Tungusic languages (and particularly Southern Tungusic) do not display serial verb constructions. The number of auxiliary verbs in Southern Tungusic is usually restricted to *bi- 'to be', *oː- 'to become, to do' and ta- 'to do' while in most of other Transeurasian languages (Turkic, Mongolic, Japonic and Koreanic) verbs of position, motion and some others are also often grammaticalized into auxiliary ones.

   The Tungusic languages might have had contacts with neighboring languages even at the proto-language stage that must reflect in modern Southern Tungusic languages. I will follow

Janhunen (2013/2016) who describes contacts between Tungusic and non-Tungusic languages. The Proto-Tungusic language might have interacted with Proto-Mongolic, Proto-Amuric (the protolanguage of Nivkh), Proto-Japonic and Proto-Koreanic. However, it is very difficult to reconstruct the influence of these protolanguages on each other because there is not enough evidence for explaining some similar features or lexical items as a result of language contacts, or a case of genetic relationship.

In the later period Tungusic languages (and Southern languages on later stages) interacted with Nivkh. At least some vocabulary items mostly connected with maritime topic were borrowed from Nivkh. It is also assumed that Ulcha is a variety which was formed in the Nivkh-Nanai contact zone, see Janhunen (2013/2016: 254–255). Orok also displays a number of Nivkh loanwords which are absent in other Tungusic languages. Beyond that, Orok borrowed a number of lexical cultural items from Ainu.

During last centuries the Southern Tungusic languages interacted with Russian and Chinese intensively. Russian has had large influence on all language levels (i.e. vocabulary, grammar and phonology) of all languages under discussion but Hezhe. At the same time, the Tungusic languages have also influenced Russian dialects of the Far East with regard to borrowed vocabulary and some calqued grammatical constructions. The Chinese language influenced much Hezhe with regard to lexicon and grammar. Some cultural loanwords from Chinese are observed in Nanai and Udihe due to the trade relations. An interesting case of Chinese and Southern Tungusic interference is the language of Taz people. This is a population which inhabits the area close to the border of China and Russia. It was formed as a result of intensive contacts between the Chinese and Udihe. Therefore, the Taz language is based on Chinese with a great impact of Udihe, see Janhunen (2013/2016: 260).

The Southern Tungusic languages had also contacts with other groups of Tungusic languages. Thus, Nanai obtains some Manchu lexical and derivational items that are hardly a result of genetic relations. Nanai has also some Mongolian loanwords, which are likely to be borrowed

via Manchu rather than directly from Mongolian. Two other cases of inter-Tungusic contacts have already been mentioned above: Kur-Urmi, which may be considered as a mixed Evenki-Nanai language, and Gorin Nanai dialect, which was influenced by Negidal. It is assumed that the Gorin Nanai are the Evenki by origin who were assimilated with the Nanai. It could explain a number of Northern Tungusic features on different language levels of the Gorin dialect.

19.3 Phonology

19.3.1 Consonants

The Nanai consonant system includes 19–22 phonemes depending on a dialect.

Table 19.2 Consonant system of Nanai (based on Moroz 2013)

	labial		alveolar		palatal		velar	
	-v	+v	-v	+v	-v	+v	-v	+v
plosive	p	b	t	d			k	g
fricative	(ɸ)		s				x	
affricates			(ts)	(dz)	tɕ	dʑ		
nasal	m		n		ɲ		ŋ	
approximant	w		l, lʲ		j			
trill			r (rʲ)					

Table 19.2 demonstrates a consonant system of the Najkhin dialect which was used for developing a standard variant of Nanai. Phonemes attested in other dialects but not in standard Nanai are put in parentheses, e.g. (rʲ). All the phonemes are given in IPA format. The phoneme

/ɸ/ occurs in Ussuri dialect and partially in Sikachi-Aljan, the phonemes /ts, dz/ occur in Sikachi-Aljan and partially in Najkhin although it was changed to /tɕ, dʑ/ in the standard Nanai in respect of Lower Amur Nanai dialects. The phoneme /rʲ/ is attested in Dzhuen dialect. The Ussuri dialect does not display the phoneme /p/.

The velar consonants /k, g, x/ occasionally have phonetic uvular realization [q, ɢ, χ] which does not have a phonemic status as well as palatal realization of all the consonants but /t, d, n, l/. In position of palatalisation /tɕ, dʑ, ɲ, lʲ/ are used instead of /t, d, n, l/. All of them are proved to be phonemes in Najkhin Nanai (see Moroz 2013). In Dzhuen dialect /rʲ/ has also developed a phonemic status.

19.3.2 Vowels

The Nanai vocalic system comprises 6 short and 6 long vowels (see Table 19.3).

Table 19.3 Nanai vowels

	front	central	back
close	i iː		u uː
close-mid	e eː		
open-mid		ə əː	o oː
open		a aː	

The relative position of the vowels in a cell in the Table 19.3 corresponds to their average values of F1 and F2. In modern Nanai, vowels have a distinctive feature of length only in the first syllable of a word. Short vowels occasionally have a nasal phonetic realization at the end of a word (e.g., [sikʷuⁿ] 'new'). However, it is not regular and it is often realized as a combination of the vowel and /n/ or just as a vowel without nasalization: [sikʷun] or [sikʷu]. So nasal vowels cannot be considered as separate phonemes in modern Nanai.

The Nanai vocalic system also includes a number of diphthongs: *iə, iu, ea, eo, əi, oe, ui, ae, əu, ao, oa, uə*. In modern oral speech, diphthongs tend to reduce to the combination of a short vowel and a consonant: [ʲə, ʲu, ʲa, ʲo, əj, oj, uj, aj, ew, aw, wa, o] correspondingly. Nevertheless, there are some lexical items where the diphthongs still remain even in fluent oral speech, e.g., [ʤuər] 'two'.

19.3.3 Syllable structure

In Nanai, there are four possible syllable structures: CV, V, VC and CVC. V refers to a short, long vowel or to a diphthong. All four types of syllables are attested at the beginning of words: *a-ma* V-CV 'father', *og-da* VC-CV 'boat', *pok-to* CVC-CV 'way', *ga-ra* CV-CV 'branch'. Middle and final syllables can have only two types of structure: CV and CVC: e.g., *o-mol* V-CVC 'belt'. Only nasal and approximant consonants and occasionally a trill consonant /r/ may occur in the final position of final syllables (the exception is made for ideophones and recently borrowed words which allow more various syllable structures).

19.3.4 Morphophonology

Nanai displays height vowel harmony. All vowels are divided into two series:

I: *e e: a a: o o:*

II: *i i: ə ə: u u:*

The general rule is that one word form can contain vowels of one series. However, the rule often gets broken in final syllable where vowels *i i: u u:* are regularly used instead of *e e: o o:*, even in words with *e – a – o* series: e. g. *ogda-do-a-ni* [boat-DAT-OBL-3SG] 'in his boat'. Though it does not occur with the vowels *a a: ~ ə ə:*. The vowels *i i:* are also often attested in middle syllables instead of *e e::* e. g. *ʤari-ni* [sing.PRS-3SG] '(s)he sings'. Therefore, the first syllable appears to be the only position where phonemes *e e:* are realized constantly. Avrorin (1957: 58–60) suggests that the change of vowels in final syllable takes place due to the vowel reduction.

Furthermore, the final vowel is often fully reduced in oral speech: *bi-či-ni* [be-PST-3SG] '(s)he was' > *bi-či-n'*.

Both regressive and progressive assimilation takes place in Nanai. Consonants *p n g* are subject to regressive assimilation (e.g. *gup-* 'to burn out' > *gub-dʒərə* [burn.out-FUT] '(it) will burn out'), while consonants *w d dʒ* are subject to progressive assimilation (e.g. *geol* 'paddle' + -*wa* [OBL] > *geol-ba* 'paddle-OBL'). In some cases, a regressive-progressive assimilation occurs: *exan* 'cow' + -*wa* [OBL] > *exam-ba* [cow-OBL].

Another morphophonological feature attested in Nanai is a metathesis of consonants. It happens when a stem ending with -*p* attaches a suffix beginning with *k* or *ŋ*: *dʒəp-* 'to eat' + -*kin* [PST] > *dʒəkpin* [eat.PST].

19.3.5 Suprasegmentals

The stress in Najkhin Nanai does not serve as a distinctive feature. It falls on the last syllable, which is usually phonetically longer than previous syllables though there is no phonemic opposition between short and long vowels in the last syllable. The sound intensity decreases from the first syllable to the last one. Thus, Najkhin stress relates to a pitch accent rather than to a dynamic accent. However, the state of affairs varies across dialects. In Gorin dialect the last syllable is not so emphasized. Syllables with etymologically long vowels appear to be more intensive though in the words without long vowel the last syllable remains accented.

19.4 Morphology

19.4.1 Inflectional morphology of nouns

Nanai nouns can be marked for number, case, and possessiveness. Case system includes nominative, oblique (or accusative), destinative, instrumental, dative, locative, directive and ablative cases. Number system includes singular and plural forms. Possessive suffixes mark the

1st, 2nd, 3rd person singular or plural. There is no distinction of exclusive and inclusive forms for 1st person plural.

Tables 19.4–6 demonstrate two types of Nanai nominal paradigms: a non-possessive paradigm, i.e. without possessive markers, a part of the possessive paradigm and a reflexive paradigm.

Table 19.4 A non-possessive paradigm of *mapa* 'a bear'

	SG	PL
NOM	*mapa* bear	*mapa-sal* bear-PL
OBL (ACC)	*mapa-wa* bear-OBL	*mapa-sal-ba* bear-PL-OBL
INS	*mapa-ǯi* bear-INS	*mapa-sal-ǯi* bear-PL-INS
DAT	*mapa-du* bear-DAT	*mapa-sal-du* bear-PL-DAT
LOC	*mapa-la* bear-LOC	*mapa-sal-dola* bear-PL-LOC
DIR	*mapa-či* bear-DIR	*mapa-sal-či* bear-PL-DIR
ABL	*mapa-ǯeaǯi* bear-ABL	*mapa-sal-ǯeaǯi* bear-PL-ABL

Table 19.5 A personal-possessive paradigm of *mapa* 'a bear' (1sg and 2sg)

	1SG	2SG

	SG	PL	SG	PL
NOM	*mapa-i* bear-1SG	*mapa-sal-bi* bear-PL-1SG	*mapa-si* bear-2SG	*mapa-sal-si* bear-PL-2SG
OBL (ACC)	*mapa-e-wa* bear-1SG-OBL	*mapa-sal-be-wa* bear-PL-1SG-OBL	*mapa-wa-si* bear-OBL-2SG	*mapa-sal-ba-si* bear-PL-OBL- 2SG
DEST	*mapa-go-e-wa* bear-DEST-1SG-OBL	*mapa-sal-go-e-wa* bear-PL-DEST-1SG- OBL	*mapa-go-a-si* bear-DEST- OBL-2SG	*mapa-sal-go-a- si* bear-PL-DEST- OBL-2SG
INS	*mapa-ʤe-e-wa* bear-INS-1SG-OBL	*mapa-sal-ʤe-e-wa* bear-PL-INS-1SG-OBL	*mapa-ʤe-a-si* bear-INS-OBL- 2SG	*mapa-sal-ʤe- a-si* bear-PL-INS- OBL-2SG
DAT	*mapa-do-e-wa* bear-DAT-1SG-OBL	*mapa-sal-do-e-wa* bear-PL-DAT-1SG- OBL	*mapa-do-a-si* bear-DAT-OBL- 2SG	*mapa-sal-do-a- si* bear-PL-DAT- OBL-2SG
LOC	*mapa-la-e-wa* bear-LOC-1SG-OBL	*mapa-sal-dola-e-wa* bear-PL-LOC-1SG- OBL	*mapa-la-si* bear-LOC-2SG	*mapa-sal-dola- si* bear-PL-LOC- 2SG
DIR	*mapa-če-e-wa* bear-DIR-1SG-OBL	*mapa-sal-če-e-wa* bear-PL-DIR-1SG- OBL	*mapa-če-a-si* bear-DIR-OBL- 2SG	*mapa-sal-če-a- si* bear-PL-DIR- OBL-2SG

ABL	*mapa-ʤeaʤe-e-wa* bear-ABL-1SG-OBL	*mapa-sal-ʤeaʤe-e-wa* bear-PL-ABL-1SG-OBL	*mapa-ʤeaʤe-a-si* bear-ABL-OBL-2SG	*mapa-sal-ʤeaʤe-a-si* bear-PL-ABL-OBL-2SG

Table 19.6 A reflexive paradigm of *mapa* 'a bear'

	REFL.SG		REFL.PL	
	SG	PL	SG	PL
NOM	—	—	—	—
OBL (ACC)	*mapa-i* bear-REFL.SG	*mapa-sal-bi* bear-PL-REFL.SG	*mapa-wari* bear-REFL.PL	*mapa-sal-bari* bear-PL-REFL.PL
DEST	*mapa-go-i* bear-DEST-REFL.SG	*mapa-sal-go-i* bear-PL-DEST-REFL.SG	*mapa-go-ari* bear-DEST-REFL.PL	*mapa-sal-go-ari* bear-PL-DEST-REFL.PL
INS	*mapa-ʤe-i* bear-INS-REFL.SG	*mapa-sal-ʤe-i* bear-PL-INS-REFL.SG	*mapa-ʤe-ari* bear-INS-REFL.PL	*mapa-sal-ʤe-ari* bear-PL-INS-REFL.PL
DAT	*mapa-do-i* bear-DAT-REFL.SG	*mapa-sal-do-i* bear-PL-DAT-REFL.SG	*mapa-do-ari* bear-DAT-REFL.PL	*mapa-sal-do-ari* bear-PL-DAT-REFL.PL
LOC	*mapa-la-i* bear-LOC-REFL.SG	*mapa-sal-dola-i* bear-PL-LOC-REFL.SG	*mapa-la-ri* bear-LOC-REFL.PL	*mapa-sal-dola-ri* bear-PL-LOC-REFL.PL
DIR	*mapa-če-i* bear-DIR-REFL.SG	*mapa-sal-če-i* bear-PL-DIR-REFL.SG	*mapa-če-ari* bear-DIR-REFL.PL	*mapa-sal-če-ari* bear-PL-DIR-REFL.PL

ABL	*mapa-ʤeaʤe-i* bear-ABL- REFL.SG	*mapa-sal-ʤeaʤe-i* bear-PL-ABL- REFL.SG	*mapa-ʤeaʤe-ari* bear-ABL- REFL.PL	*mapa-sal-ʤeaʤe-ari* bear-PL-ABL-REFL.PL

The data in the tables above show some peculiarities of the nominal declension. The destinative case can occur only with a possessive suffix, which marks beneficiary of the situation (see the discussion in Section 19.2 above). The order of suffixes in personal-possessive forms depends on the person of possessor: it is root-(PL)-POSS-OBL for 1st person both singular and plural and root-(PL)-OBL-POSS for 2nd and 3rd person singular and plural. OBL refers to the oblique case marker, which marks the direct object in accusative forms both in simple and personal-possessive paradigms and functions as an additional suffix in oblique case forms. In reflexive forms it is absent. The use of oblique case marker is a specific feature of Nanai, it is not attested in other Southern Tungusic languages and even in peripheral Nanai dialects. Its origin remains unclear (see the discussion of this problem and some speculations on a possible grammaticalization way in Avrorin 1948: 233–255; Oskolskaya 2015).

19.4.2 Inflectional morphology of pronouns

The Nanai language exhibits personal, reflexive, demonstrative and interrogative pronouns. Table 19.7 demonstrates three case forms of personal pronouns, i. e. just a part of the paradigm.

Table 19.7. Personal pronouns (NOM, ACC and DAT)

	1SG	1PL	2SG	2PL	3SG	3PL
NOM	*mi*	*buə*	*si*	*suə*	*ńoa-ni*	*ńoan-či*
ACC	*mimbiwə*	*bumbiwə*	*simbiwə*	*sumbiwə*	*ńoam-ba-ni*	*ńoam-ba-či*
DAT	*min-du*	*buŋ-du*	*sin-du*	*sun-du*	*ńoan-do-a-ni*	*ńoan-do-a-či*

The morphology of 1st and 2nd person pronouns differs much from the one of 3rd person pronouns. The 1sg pronoun has *mi* in nominative form, which is a specific feature of Nanai and Kur-Urmi: other Tungusic languages display *bi* for nominative of 1st sg pronoun. Nanai does not have different forms for inclusive or exclusive 1st pl pronoun. Accusative forms of 1st and 2nd person pronouns are irregular and include a syllable *bi* of unclear origin. The oblique stems of 1st and 2nd person pronouns end with *-n* (see the dative forms in the Table 19.7). The 3rd person pronouns include endings *-ni* and *–či,* which coincide with the personal-possessive markers, cf. *ogda-ni* [boat-3SG] 'his boat'. However, they have another function in personal pronouns: they refer to the referent itself rather than to its possessor. The declension of 3rd person pronouns corresponds to the personal-possessive noun declension including the use of a special oblique case marker.

The reflexive pronoun is *məpi* (*mənə* serves as a reflexive-possessive pronoun). Demonstrative pronouns exhibit two-partial opposition: *əj* 'this' for close distance and *təj* 'that' for far distance. The basic interrogative pronouns are *xaj* 'what?' (for inanimate objects and animals) and *uj* 'who?' (for persons). Other demonstrative and interrogative pronouns are mostly etymologically derived from these four pronouns with the means of case markers, e.g. *əj-du* [this-DAT] 'here (lit. in this (place))'.

19.4.3 Numerals

The Nanai number system is decimal. First ten cardinal numerals as well as numerals for 20, 30, 40, 50, 100 and 1000 are primary, non-derivative. The numerals for 60, 70, 80, 90 are derived with the suffix *-EŋO*: *nada-eŋgo* [seven-DEC] '70'. All other cardinal numerals consist of the combination of primary numerals: *ǯoan əmun* [ten one] '11'.

The suffix *-tO* (*-tOŋgA* or *-tOŋgAsA*) marks the entire set of objects: *duin-tu* [four-UNI] 'all four (objects)'. Numerals with accusative, or oblique, marker *-wA* denote approximateness:

xorem-ba naj-sal [20-OBL person-PL] 'about 20 people'. General cardinal numerals can also derive ordinal numerals with suffix *-čEA* (e.g. *elan-čea* [three-ORD] 'third', 'first' has suppletive forms *boŋgo* and *ʤuluj*), numerals meaning the amount of days with classifier *-ltA* (*nada-lta* [seven-DAY] 'seven days'), collective numerals with instrumental case suffix and plural reflexive marker *-ʤEAri* (*ʤoan-ʤe-ari* [ten-INS-REFL.PL] 'ten together (adv.)'), restrictive numerals with the suffix *-rOkAn* (*dui-rukən* [four-RESTR] 'only four together (adv.)') and distributive numerals by means of reduplication (*ʤuər-ʤuər* [two-two] 'two by two').

19.4.4 Property words

Basic adjectives describing quality, color, size etc. of an object are morphologically primary, they do not have any special suffixes of adjectivization. Adjectives can be derived from adverbs by attaching the auxiliary verb in present tense form *bi* '~is': *nam bi* [comfortable (adv.) be.PRS] 'comfortable (adj.)'. The suffix *-ktO* forms verbal adjectives with a resultative meaning: *jakči-* 'to lock' > *jakči-kto* 'locked'. The suffix *-psi* derives adjectives from verbs with the meaning of a feeling or perception: *ŋəːlə-* 'to fear' > *ŋəːlə-psi* 'frightful'. These and some other suffixes are not highly productive. The adjectives of size etymologically include suffix *-mi,* which is nowadays not productive at all: *dara.mi* 'wide'.

   A number of other suffixes such as *-kAn* or *-kO* express the grade of the attribute: *nuči-ku* [little-DIM] 'very little'.

   Nanai adjectives do not have any inflectional morphology.

19.4.5 Inflectional morphology of verbs

The most neutral verbal forms which occur in realis mood for factual statements in modern Nanai originate from participle forms. Nowadays, they function both as finite forms and as participles. Table 19.8 demonstrates the mood and tense forms of the Nanai verb *niru-* 'to write'.

Table 19.8 Moods and tenses of *niru-* 'to write' (only one person-number form is given)

form	example	glossing
INDICATIVE MOOD		
Past (= past participle)	*niru-hə-ni*	write-PST-3SG
Imperfect	*niru-j bi-či-(ni)*	write-PRS be-PST-(3SG)
Present (= present participle)	*niru-j-ni*	write-PRS-3SG
Prospective	*niru-j*	write-PRS
Future	*niru-ʤərə*	write-FUT
Future II	*niru-ʤə=mə*	write-FUT=PTCL.EMPH
ASSERTIVE MOOD		
Present	*niru-əm-bi*	write-ASS.NPST-1SG
Past	*niru-kə-i*	write-ASS.PST-1SG
OPTATIVE	*niru-ŋə*	write-OPT
IMPERATIVE		
Canonical imperative		
Neutral imperative	*niru-ru*	write-IMP
Distant imperative	*niru-həri*	write-IMP.DIST
Hortative	*niru-gitə*	write-HORT.SG
Jussive	*niru-gi-ni*	write-JUSS-3SG
SUBJANCTIVE MOOD	*niru-mčə*	write-SBJV
CONVERBS		
Simultaneous	*niru-mi*	write-CVB.SIM.SG
Non-simultaneous	*niru-rə*	write-CVB.NSIM
Conditional I	*niru-pi*	write-CVB.COND.SG

Conditional II	*niru-uči-ə-ni*	write-CVB.COND2-OBL-3SG
Purpose	*niru-gu-j*	write-CVB.PURP-REFL.SG
IMPERSONAL FORM	*niru-u-ri*	write-IMPS-PRS

Aspect has not been grammaticalized into a separate grammatical category in Nanai, i.e. most of the verbal forms do not express any aspectual meaning, or, in other words, they can occur in different aspectual contexts. Though some tense forms include aspectual component in their semantics. While past tense can refer both to progressive and completed events in the past, imperfect indicates first of all a habitual action in the past which usually does not take place at present time. Alongside with other forms and expressions, imperfect can also mark a progressive action in past (11).

(11) a. Nanai (Avrorin 1986: 232)

jajka:n   *more-e-wa-ni*   *a:mda-mi*   **ʒari-i**   **bi-či-ni**

bird   cry-PRS-OBL-3SG   imitate-CVB.SIM.SG   sing-PRS   be-PST-3SG

'He used to sing imitating a bird voice'.

b. Nanai (Oskolskaya, fieldnotes)

tawa-či   *ʒi-čin-du-j-ə*   *tawa*   *ələ*   **gup-či**   **bi-čin[7]**

fire-DIR   come-PST-DAT-1SG-OBL   fire   already   go.out-PRS   be-PST

'When I came to the fire, the fire was going out.'

Assertive forms are etymologically finite words. Nowadays, they are mostly only used in special contexts. They appear to mark the involvement of the person who is in charge of assertion. Thus, in example (12a) the speaker makes a statement (so, he is in charge of assertion) and he speaks about himself, so he is involved in a situation which he describes. In example (12b) the speaker is also in charge of assertion, but he is not involved in the described situation, therefore the

indicative present is used. In the example (12c) it is supposed that it is the addressee who is responsible for assertion and he is supposed to be involved in a described situation, therefore the occurrence of assertive form is possible.

(12) Nanai (Oskolskaya, fieldnotes)

    a. *mi   ʤobo-am-bi*

       1SG   work-ASS.NPST-1SG

       'I am working.'

    b. *ńoani   ʤobo-j-ni*

       3SG   work-PRS-3SG

       'He is working.'

    c. *si   huktə-i   selko-ka-si?*

       2SG   tooth-REFL.SG clean-ASS.PST-2SG

       'Have you cleaned your teeth?'

As a result, it is not surprising that 1st and 2nd person assertive forms are attested much more frequently than 3rd person assertive forms. Still, 1st and 2nd person indicative forms occur more frequently than the corresponding assertive forms at least in the available text collections. The cognates of assertive forms have a bit different functions in other Tungusic languages (see Malchukov 1999).

    Future form is morphologically derived from assertive non-past, but unlike assertive mood it does not include a modal component of assertor's involvement—it is the basic form for marking an event in future. Therefore, it refers to indicative mood.

    The form of future II occurs rather occasionally. It is derived with the future suffix -*ʤA* (without -*rA*) and clitic =*mA*. Future II implies inevitability of the future event.

Prospective form is morphologically a bare participle form, i.e. without person-number markers, which is used as a head of a finite form. Perspective indicates an event that is going to happen in the near future (13).

(13) Nanai (Oskolskaya, fieldnotes)

semata ələ~ələ   uːn-ʤi

snow   soon~soon melt-PRS

'The snow is about to melt.'

Canonical imperative, i.e. a category which implies commands and refers to 2nd person, includes two morphologically different forms: a neutral and a distant one. The neutral form may be used in all the occurrences where a command or a request takes place. The distant imperative can be used instead of the neutral form in the case when it is supposed that an addressee should perform an action after another event rather than just after the request (14).

(14) Nanai (Oskolskaya, fieldnotes)

poja mi   nasaptom-bi   huədə-uči-ə-j-ə   baː-go-hari

kid 1SG glasses-REFL.SG lose-CVB.COND2-OBL-1SG-OBL find-REP-IMP.DIST

'Baby, when I lose my glasses, please, find them.'

Hortative and jussive forms occur very occasionally when 1st or 3rd person correspondingly are encouraged to perform an action.

The subjunctive mood marks an unreal event. Usually it is used in the main clause of conditional sentences.

There is also a special impersonal suffix *-(w)O* which could refer to the derivational verbal suffixes with modal meanings as it often expresses obligation or possibility: e.g. *halače-o-* [wait-

IMPS-] 'one should wait'. However, an impersonal stem has only present and past forms as distinct from other stems with derivational suffixes, which normally exhibit a full paradigm. Therefore, an impersonal form is usually considered as a special participle form (along with present and past participles).

Apart from the present and past forms, which can function as participles, converbs can also head a non-finite clause. As a rule, simultaneous converb refers to an action that is performed at the same time with the action indicated by the finite verb form. Non-simultaneous converb refers to a preceding action. Two conditional converbs occur in more or less identical contexts and they mark the condition or sometimes time of the event expressed in finite clause. They differ in syntactic properties (see Section 19.5.3). The converb of purpose indicates the purpose of an action expressed in the main clause.

Most of the verbal forms acquire person-number markers. There are two series of them: the first one I is historically nominal and coincides with the person-number markers attached to nouns, the other one II is purely verbal, see Table 19.9.

Table 19.9 Person-number markers

	I	II
1SG	*-i / -bi*	*-i / -bi*
2SG	*-si*	*-či / -si*
3SG	*-ni*	Ø
1PL	*-pu*	*-pu*
2PL	*-su*	*-su*
3PL	*-či*	*-l*

The markers of the series I attach to present, past, imperfect, optative, and subjunctive forms while suffixes of the series II attach to future, future II and both assertive forms.

A system of verbal negation in Nanai is very complicated and asymmetrical comparing to the assertive verbal forms. Most of the negative constructions contain an item *ə-,* which originates from a special negative verb. It does not function as a verb with full paradigm in Nanai, though Nanai represents different stages of grammaticalization of the negative verb into a negative suffix. Thus, the most common negative construction referring to past is analytical and includes a negative particle: e. g. *əčiə niru-ə* [PST.NEG write-NEG] 'did not write' where *əčiə* is etymologically a past form of the negative form. The suffix *-ə* in this past negative construction has the same origin as the assertive mood marker **-ra*, but on a synchronic level they differ due to morphological and syntactic features. The most common negative form for present is synthetic: *niru-ə-si-ni* [write-NEG-PRS-3SG] 'does not write', which is etymologically derived from a combination **niru-ə ə-si-ni* [write-NEG NEG-PRS-3SG]. Another past negative construction, which is frequently attested in the Sikachi-Aljan dialect, contains an existential negative item *aba* rather than any relics of the negative verb **ə-*: *niru-mi=də aba* [write-CVB.SIM.SG=PTCL.EMPH NEG.EXIST] 'did not write'.

19.4.6 Derivational morphology

Nanai is a very rich language in respect of derivational affixes. Here I list only some of them. The full list of the affixes and their detailed description can be found in Avrorin (1961: 22–65). Nominal derivational suffixes can be divided into two groups: nominalization suffixes which derive noun from other parts of speech (e.g. *bota-mʤi* [fish(v.)-AG] 'a fisherman') and suffixes which change the meaning of a noun (e.g. *sole-kan* [fox-DIM] 'a little fox').

The basic verbalizing suffix, which is highly productive, is *-lA*. It derives verb stems from nominals or from borrowed words. Alongside with *-lA*, the suffix *-čI* can verbalize nouns or adjectives. Then these suffixes can also express perfective and imperfective aspectual semantics

respectively (e.g. *ajakta-la-* [angry-VBLZ-] 'to become angry' vs. *ajakta-če-* [angry-IPFV-] 'to be angry').

Furthermore, Nanai displays a number of markers that occur between the verbal root and inflectional markers of tense and person-number. Their semantics is connected with modality, voice or aspect. Some of them are highly productive so that they can be considered derivational suffixes only with stipulations. Voice markers are causative *-wAn* (*sea-wan-* [eat-CAUS-] 'to feed'), decausative *-p* (*huədə-p-* [lose-DECAUS-] 'to get lost'), reciprocal *-mAčE* (*sore-mače-* [fight-RECP-] 'to fight with each other'). Modal suffixes include debitive *-gElA* (*ga-gela-* [buy-DEB-] 'must buy'), desiderative *-jčA* (*mora-jča-* [cry-DES-] 'want to cry') and some other. The most productive aspectual markers are repetitive *-gO* (*niru-gu-* [write-REP-] 'to write once again to rewrite'), inchoative *-lO* (*niru-lu-* [write-INCH-] 'to start writing') and distributive *-ktA* (*nə:-ktə-* [put-DISTR-] 'to put things separately').

19.5 Syntax

19.5.1 The clause

Basic word order in Nanai is SOV (subject – object – verb) and dependents usually precede heads. However, the word order is not strict and can be disrupted due to pragmatic reasons.

The word order remains the same in questions, interrogative words take the position *in situ*, i.e. in the place where a supposed phrase would be in the relevant assertive sentence rather than at the beginning of a question (15).

(15)   Nanai (Oskolskaya and Stoynova, fieldnotes)

   mapa haj-du   bi-i-ni?

   bear   what-DAT   be-PRS-3SG

   'Where is the bear?'

Nanai is a pro-drop language, i.e. pronominal subjects can be omitted while the subject referent is clear due to person-number markers on a verb form (16).

(16) Nanai (Oskolskaya and Stoynova, fieldnotes)

(*mi*) *jada-ham-biə!*

1SG   get.tired-PST-1SG

'I am tired!'

There is no clear distinction between coordination and subordination of clauses. Syntactically all non-final clauses are subordinate as their heads are expressed by non-finite verb forms (converbs or bare participles). However, these clauses can relate to the final clause as they are coordinate clauses in semantic respect (17).

(17) Nanai (Avrorin 1986: 229)

sənə-rəː	*hadʒom-bi*	*baːradʒego-raː,*	*pokto-i*
wake.up-CVB.NSIM	clothes-REFL.SG	put.on-CVB.NSIM	way-REFL.SG
ele-mi	*dəruː-hə-ni*		
stand.up-CVB.SIM.SG	begin-PST-3SG		

'He woke up, put on his clothes and went on his (hunting) path.'

In example (17), if the clauses change their place, the meaning of the whole sentence will be different: 'He put on his clothes, woke up and went on his path'. Such a property is typical for coordination rather than for subordination, though actions are expressed by converbs, i.e. syntactically subordinate forms.

This type of clause chaining is quite common in Nanai, and non-simultaneous converbs occur especially frequently in narrative chains such as (17).

Existential constructions are expressed with the verb *bi-* 'be' including locative, identification, possession and other types of existential constructions. There is no special verb for marking possession. The verb *bi-* is sometimes omitted in sentences referring to present: e.g. *Mədur daːi.* [dragon big] 'A dragon is big' vs. *Mədur daːi bi-či-ni* [dragon big be-PST-3SG] 'A dragon was big'.

19.5.2 The nominal group

Nanai exhibits a nominative-accusative alignment with zero-marking of a subject both of transitive and intransitive predicates and oblique case marking of an object. The direct object is, though, often expressed by a nominative form without any case markers, which is connected with low referential and indefinite status, inanimacy or with position in information structure of the sentence (18).

(18)   Nanai (Oskolskaya and Stoynova, fieldnotes)

 əm ʤoː baː-ha-ni

 one house find-PST-3SG

 '(he) found a house'.

More or less, the same parameters influence the plural marking, which is also optional (19).

(19)   Nanai (Oskolskaya and Stoynova, fieldnotes)

 kəči *waː-o-r* *aso* *ərdəŋgə* *bi-ə-si*

 carp kill-IMPS-PRS not.very interesting be-NEG-PRS

 'It is not very interesting to fish carps'.

In example (19), it is supposed to be plural as one normally catches several carps while fishing.

There are no special possessive forms of personal pronouns. The possession is always marked on the possessed with person-number suffixes. Still a pronoun referring to a possessor can be also expressed in its nominative form: *mi: piktə-i* [1SG child-1SG] 'my child'.

Nanai does not display special grammaticalized means for marking topic or focus. Nevertheless, there is a number of discourse particles with a great variety of functions. Some of them can serve for underlining a discourse status (20).

(20) Nanai (Oskolskaya and Stoynova, fieldnotes)

təj	*neka-sal=tani*	*bun-du*	*gəsə*	*bi-či-či*
that	Chinese-PL=PTCL	1PL-DAT	together	be-PST-3PL

'(We, the Nanai and the Chinese, used to live together.) Those Chinese lived at our place together (with us).'

In example (20), a particle *=tAni* marks topic or given information. Although it is not obligatory to mark the topic, it never occurs with nominal phrases referring to focus or new information.

Another very frequent particle, *=dA*, is more or less neutral in respect of topic-focus marking, but it exhibits a wide range of syntactic functions. Thus, it can be used for contrasting the nominal phrases or as a connective (21).

(21) Nanai (Oskolskaya and Stoynova, fieldnotes)

*ənim=**də***	*sagʒi*	*bi-či,*	*amen=**da***	*sagʒi*	*bi-či-ni*
mother=PTCL.EMPH	old	be-PST	father=PTCL.EMPH	old	be-PST-3SG

'Both mother was old and father was old.'

This particle is also used for deriving indefinite pronouns which can occur in negative contexts: *xaj-wa=**da** ičə-ə-si* [what-OBL=PTCL.EMPH see-NEG-PRS] '(he) does not see anything'.

19.5.3 The verbal group

Nanai verbs can have from zero (*tugdə-j-ni* [rain(v)-PRS-3SG] 'it rains') to at least three

arguments. The voice marker changes the valency of the verb. Thus, the causative marker

introduces a new participant into a situation, a causer, who takes a syntactic position of subject (a

boy in example 22b), while cause, which is a subject of the caused action, takes a syntactic

position of direct object (a dog in example 22a). The direct object of the caused action (water in

example 22) remains at the same syntactic position.

(22) Nanai (Oskolskaya, fieldnotes)

    a. *Enda muə-wə   ome-ha-ni.*

       dog   water-OBL drink-PST-3SG

       'A dog drank water.'

    b. *Naondʒokan   enda-wa muə-wə   ome-waŋ-ke-ni.*

       boy   dog-OBL water-OBL drink-CAUS-PST-3SG

       'A boy gave water to the dog.'

Decausative and reciprocal suffixes are much less productive. Still, they also influence the

argument structure of a verb. Both of them decrease the number of arguments. In the case of

verbs with decausative marker there is no agent of situation ("we" in example 23) while a

semantic undergoer (a girl) takes a subject syntactic position.

(23) Nanai (Oskolskaya, fieldnotes)

    a. *Buə arčokam-ba   huədə-hə-pu.*

       1PL   girl-OBL   lose-PST-1PL

       'We lost the girl.'

b. *Arčokan   huədə-kpi-ni.*

girl   lose-DECAUS.PST-3SG

'The girl got lost.'

In the case of verbs with reciprocal marker, there is a plural subject and all participants are agents and undergoers of the situation at the same time. In example (24), it is true about each boy that he (agent) hit another one and he (patient) was hit by another one at the same time.

(24) Nanai (Oskolskaya, fieldnotes)

ʤuər naonʤoka-sal sore-mače-ha-či.

two   boy-PL   fight-RECP-PST-3PL

'Two boys fought with each other.'

Converbs differ in respect of subject referents. The subject of the simultaneous, non-simultaneous, conditional I or purpose converb and the subject of the finite verb form should be the same, while the subjects of the conditional II converb and the finite verb are always different. It is the main property that allows to divide the use of conditional I and conditional II converbs.

19.6 Lexicon

The basic Nanai vocabulary is in general common with other Tungusic languages. Apart from nouns, verbs and adjectives, Nanai displays a large class of ideophones, the words expressing manner, color, size, shape, sound, smell and other characteristics of an object or an action. Nanai ideophones exhibit specific features such as reduplication, combination with auxiliary words and some others: e.g. *koah-koah bi* [hard be.PRS] 'hard (adj)'.

  The Nanai dialects differ in respect of borrowed vocabulary. Thus, in the Bikin dialect a lot of Manchu and Chinese loanwords are observed: *emaha* 'fish' < Ma. *nimaha*, *naịna* 'grandmother'

< Ch. *naina*. The Gorin dialect borrowed some items from Evenki or Negidal: *ʤuki* 'otter' < Neg. *ʤukin*. All the Nanai dialects have loanwords from Manchu (*ehan* 'cow'), Mongolian (*moren* 'horse'), Chinese (*duduəsə* 'potato') and Russian (*repa* 'turnip').

Nanai exhibits a tendency to synthetism when analytical constructions are turning into single form: e.g. *niru-ə-si-ni* [write-NEG-PRS-3SG] 'does not write' < **niru-ə ə-si-ni* [write-NEG NEG-PRS-3SG]. Derivational markers tend to lexicalize, i.e. to fuse with the root and to lose their productivity. For instance, probably there used to be a suffix *-gi* which used to derive transitive verbs from intransitive ones and thus reflected some modern verbs: *ʤəgʤi-* 'burn (tr)' < **ʤəgdə-gi-* [burn(in)-TR-].

Acknowledgments

---

[1] Kili is also known as Kur-Urmi, which used to be considered as a dialect of Nanai.

[2] In Oroch, 'to find' will be *baː-* (Avrorin and Lebedeva 1978: 165) rather than *baka-* which was expected to be from (Doerfer 1978a: 11). Still, it just proves the classification proposed by Doerfer.

[3] Here and elsewhere the names of languages are given in Glottolog variants, the names of three main Tungusic groups are given as follows: Manchuric, Southern and Northern; all other names are taken from the cited sources.

[5] A direct object can be marked by the oblique case marker (2) while other oblique case forms are marked by a combination of two markers: a mere case + oblique case, e. g. DEST + OBL as in example (1), see Section 19.4.1 of the chapter for more details.